VIT | Vidyalankar Institute of Technology
Accredited A+ by NAAC

ISSN : 2278 - 1722

Applied Engineering and Technologies

VOLUME - 11, ISSUE 1

June - 2024

CONTENTS

Paper Id	Title	Page No.
AET 001	Secureye	1
AET 002	Leveraging LSTM Ensembles for Enhanced Remaining Useful Life Prediction of Filter Mats	8
AET 003	Enhancing Urban Infrastructure Management through AI-Powered Defect Detection and Reporting Systems	14
AET 004	Automatic Timetable Generator using Genetic Algorithm	19
AET 005	Video Surveillance for Violence Detection	23
AET 006	CropNexus: Your All-in-One Crop Management Solution	28
AET 007	Sign Language Recognition System for Deaf and Mute People	33
AET 008	AI-based tool for Preliminary Diagnosis of Dermatological Manifestations	38
AET 009	Plant Disease Detection	47
AET 010	Blockchain-based Copyright Detection System: A Decentralized Approach to Protecting Intellectual Property	51
AET 011	Allocatrix: Revolutionizing Investment Strategy Optimization	57
AET 012	Saga – Gauging Seconds	62
AET 013	Car crash Detection and Assistance	65
AET 014	Performance Analysis of 16 Bit RISC Processor Using Vedic Multiplier	68
AET 015	Cricket Data Analytics: IPL Score Prediction	78

Continue on back cover page...



Published by

Vidyalankar Institute of Technoogy

Wadala, Mumbai, Maharashtra

www.vit.edu.in

VIT

Vidyalankar Institute of Technology







About VIT

Vidyalankar - A Sanskrit word which reveres 'Vidya' (Knowledge) as a precious 'Alankar' (Ornament), the essence being that knowledge is the true ornament of a progressive mind. A trusted and prestigious education specialist, the Vidyalankar Group of Educational Institutes continues to spread the radiance and the bright sparkle of knowledge and education far and wide.

Vidyalankar Institute of Technology (VIT) is an Autonomous Engineering degree college approved by the AICTE, the DTE (Maharashtra), and affiliated to the University of Mumbai. Institute is accredited A+ by NAAC and all UG Programs of VIT are NBA Accredited. From June 2022 onwards VIT is an Autonomous institute affiliated with University of Mumbai.

VIT offers undergraduate programs in Information Technology, Computer Engineering, Electronics and Computer Science, Electronics & Telecommunication Engineering, and Biomedical Engineering, with an average annual intake of 660 students. The Institute also provides postgraduate programs in Electronics & Telecommunication, Computer Engineering, and Management Studies, along with doctoral (Ph.D.) programs in Computer Engineering and Electronics & Telecommunication Engineering

EDITORIAL BOARD

Editor-in-Chief Dr. Sangeeta Joshi, Principal, VIT

Editorial Advisory Board

Dr. Varsha Bhosale-Vice Principal
Dr. Saurabh Mehta -Chief Academic Officer
Dr. Pravin Annadate-Chief Quality Officer
Dr. Mandar Sohani-Director IQAC
Dr. Arun Chavan- Head, Electronics and Computer Science
Dr. Gajanan Nagare- Head, Biomedical Engineering
Dr. Girish Gidaye- Head, Electronics and Telecommunication Engineering
Dr. Vidya Chitre-Head, Information Technology
Dr. Ravindra Sangale- Head Computer Engineering
Prof. Sonaali Borkar, In-Charge First Year Engineering
Dr. Sheetal Mapare-Convener, R&D Committee
Dr. Dhananjay Patel-Convener-Institute InnovationCell

Editorial Team

Dr. Akshatha Bhat
Dr. Sanjeev Dwivedi
Dr. Amit Nerurkar
Prof. Akhil Masurkar
Prof. Manoj Suryawanshi
Prof. Rajashree Soman
Prof. Satendra Mane
Prof. Kanchan Dhuri
Prof. Amit Aylani
Prof. Bhavika Khatri
Prof. Meenakshi Nadula
Dr. Suja Jayachandran
Prof. Geetha Narayanan

Foreword

We are delighted to present the eleventh edition of Applied Engineering and Technologies, the technical journal published by Vidyalankar Institute of Technology (VIT). The primary goal of this journal is to provide a platform for participants from academic, industrial, and research backgrounds to exchange ideas and effectively present and share their research projects.

At VIT, we have always encouraged our students and faculty not only to excel in disseminating current knowledge but also to explore new horizons and engage with technology in broader, innovative contexts.

The current edition of Applied Engineering and Technologies is centered around the theme: "Machine Intelligence for Sustainable Technology and Health." This theme reflects the dynamic convergence of Biomedical Engineering, Information Technology, Computer Engineering, Electronics, and Electronics and Communication Engineering—fields that are increasingly leveraging machine intelligence to solve complex, real-world problems.

This year's edition proudly features 47 research articles, each meticulously prepared by student groups and rigorously reviewed by an editorial team comprising experienced faculty members. Special attention was given to maintaining academic integrity, with all submissions screened through advanced plagiarism detection software to ensure originality and authenticity.

Among the highlights are articles on AI-powered defect detection for urban infrastructure management and Data-driven investment strategy optimization, showcasing the breadth and depth of research conducted by our students. These contributions reflect the potential of machine learning and deep learning applications in creating smarter, more sustainable systems across diverse sectors. Our target readers include undergraduate and postgraduate students, as well as emerging professionals from industry and academia. We hope this journal acts as a catalyst for meaningful intellectual exchanges, helping students and professionals alike to sharpen their research skills and grow into informed, capable engineers and technologists.

As technology continues to evolve at a rapid pace, formal research at the Institute remains a cornerstone of our academic philosophy. Through research, our faculty and students enrich their knowledge base, gain valuable insights, and become better equipped to meet future challenges. Initiatives like "Applied Engineering and Technologies" move us closer to realizing our vision of being a center of excellence in innovative, research-driven education, in alignment with the Vision and Mission of VIT.

The success of this initiative relies heavily on your engagement. We look forward to your valuable feedback and continued support to help us make each edition even better.

Editorial Team

AET 001

Secureye

Prajwal Vaidya
Dept. of Computer Engineering
Vidyalankar Institute of Technology
Wadala, Mumbai, India
prajwal.vaidya@vit.edu.in

Niraj Koli
Dept. of Computer Engineering
Vidyalankar Institute of Technology
Wadala, Mumbai, India
niraj.koli@vit.edu.in

Aryan Kokabankar
Dept. of Computer Engineering
Vidyalankar Institute of Technology
Wadala, Mumbai, India
aryan.kokabankar@vit.edu.in

Pratik Bhalkare
Dept. of Computer Engineering
Vidyalankar Institute of Technology
Wadala, Mumbai,India
pratik.bhalkare@vit.edu.in

Sachin Deshpande
Dept of Computer Engineering
Vidyalankar Institute of Technology
Wadala, Mumbai, India
sachin.deshpande@vit.edu.in

***Abstract*— The "Secureye" project is dedicated to addressing the pressing issue of neighboring military intrusions, often resulting in unnecessary conflicts and casualties among military personnel. The system's core functionality revolves around swiftly detecting objects in surveillance environments, particularly focusing on spotting potentially dangerous, counterfeit, or unauthorized items like weapons and armored vehicles. By utilizing data from CCTV and surveillance systems, it can accurately pinpoint threats in a timely manner. The project makes use of the YOLO model for enhanced accuracy and speed compared to R-CNN when researched.**

***Keywords*— YOLO, Object Detection, Security, CNN, Machine Learning.**

I. INTRODUCTION

Military object detection technology is a backbone in modern defense strategies, fostering advanced surveillance and threat mitigation. Recent research, exemplified by studies like "Real-time Object Detection for Military Applications using YOLO Algorithm," has delved deeply into developing real-time object detection systems tailored specifically for military use. Innovations such as YOLO (You Only Look Once) have revolutionized object recognition by significantly improving accuracy and speed, enabling swift identification of potential threats such as weapons and vehicles [2]. The integration of these cutting-edge systems into military surveillance frameworks ensures rapid responses to hostile activities. Concurrently, open-access platforms like MDPI serve as crucial hubs, fostering collaboration among researchers and facilitating the exchange of findings on military object detection innovations. This fusion of military focused object detection technologies with open-access scholarly platforms is shaping the future of defense strategies, contributing to a more secure global landscape.

Machine learning-based solutions have been developed to address the need for enhancing defense and security forces' capabilities. These systems aim to facilitate timely detection of items within inspection areas, emphasizing on identifying dangerous or illegal items like weapons and foreign military vehicles. Using data from CCTV and surveillance systems, this technology excels in swiftly and accurately identifying threats. Upon detecting a hazard, the system generates reports containing precise location details, enabling prompt responses to dangers. The versatility of these solutions is paramount, as they reinforce protection systems' security, ensuring comprehensive threat control. Our vision and purpose for this technology lie in innovating to strengthen national defense and security. Real-time object detection capabilities are essential in identifying objects and proposing security solutions. Integrated into existing surveillance systems, this technology serves as a tool for rapidly and accurately identifying threats, reflecting our commitment to protection security.

II. LITERATURE SURVEY

Neural networks have emerged as one of the most prominent algorithms today, showcasing superior accuracy and speed in processing large amounts of data over time. At their core, neurons and activation functions form the fundamental building blocks of any neural network, typically comprising input and output nodes along with hidden layers.

While traditional neural networks consist of a single hidden layer, deep neural networks, characterized by multiple hidden layers, have gained traction. These deep architectures, such as Convolutional Neural Networks (CNNs), have found widespread application in tasks like object detection and image processing, demanding substantial processing power.

Advanced techniques like Faster R-CNN and YOLO (You Only Look Once) have been developed to enhance object detection, including applications like identifying cars in aerial images. A detailed comparison between Faster R-CNN and YOLOv3 has revealed that YOLOv3 surpasses Faster RCNN in various performance metrics and suitability for different applications and environments. This understanding aids in selecting the most appropriate model for specific use cases.[1]

Computer vision is a collaboration that has received great attention in recent years (since CNN), with self-driving cars at the forefront, and its importance is visual inspection. The difference between object search algorithms and classification algorithms is that in search algorithms, we try

to draw bounding boxes around objects of interest to find them in the image. Also, in case of object detection you may not need to just draw the bounding box. There may be many boxes representing different objects in the image, and you cannot know their number in advance. The main reason why this problem cannot be solved by creating a network communication model and a fully connected process is that the length of the output process is different - not always, because it may not always be the number of times the item of interest appears. A naive way to solve this problem is to obtain different regions of interest from the image and use CNN to classify objects in this region. The problem with this approach is that objects of interest may have different locations and different locations in the image. Therefore, you should select large areas that may cause the calculation to fail. For this reason, algorithms such as faster R-CNN and YOLO have been developed to quickly detect these situations.

Recent years have seen remarkable advancements in deep learning, particularly in computer vision, with techniques like Convolutional Neural Networks (CNNs) playing a pivotal role in object detection tasks. CNNs, inspired by the visual processing mechanisms of living organisms, were initially proposed by Fukushima in 1980 and subsequently refined by LeCun. They excel in processing multi-dimensional arrays of data, such as the color channels of an image.

CNNs leverage four key concepts to exploit the characteristics of natural signals:

1. Local connections
2. Shared weights
3. Pooling
4. Multiple layers

The typical architecture of a CNN comprises several layers:

1. Convolutional layers: These layers extract features from input data, with early layers detecting lowlevel features like edges and corners, while deeper layers capture high-level features such as objects and structures. Each unit in a convolutional layer is connected to a local patch in the previous layer through a set of kernels, followed by a nonlinearity operation like ReLU (rectified linear unit).
2. Pooling layers: Positioned between convolutional layers, pooling layers reduce dimensionality and enhance invariance to small shifts and distortions.
   Each unit computes the maximum value within a local patch of units in a feature map.
3. Fully connected layers: Typically located towards the end of the network, these layers summarize information from lower-level layers for final decision-making. However, as they contain a significant number of parameters, overfitting is a concern, often addressed using dropout.

Since the breakthrough success of AlexNet in the 2012 ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), numerous CNN architectures have been developed, including VGGNet, GoogLeNet, and ResNet.

1. AlexNet, pioneered by Krizhevsky et al., marked a significant milestone by employing nonsaturating neurons, GPU acceleration, and dropout to prevent overfitting.
2. VGGNet, winner of the ILSVRC 2014 competition, is renowned for its simplicity, featuring architectures like VGGNet-16 with 13 convolutional layers, five pooling layers, and three fully connected layers.
3. GoogLeNet distinguishes itself by using filter kernels of different sizes within the same layer, preserving spatial information, and reducing network parameters to combat overfitting.
4. ResNet, awarded the Best Paper at the Conference on Computer Vision and Pattern Recognition in 2016, introduced the concept of residual learning, enabling the training of extremely deep networks efficiently, leading to superior generalization performance and victories in various prestigious competitions.[3]

YOLO: YOLO or You Only Look Once is an object detection algorithm much different from the region-based algorithm seen above. In YOLO a single convolutional network predicts the bounding boxes and the class probabilities for these boxes.
YOLO works by taking an image and split it into a side x side grid, within each of the grid we take m bounding boxes. For each of the bounding box, the network outputs a class probability and offset values for the bounding box. The bounding boxes having the class probability above a threshold value is selected and used to locate the object within the image.
YOLO is orders of magnitude faster (45 frames per second) than other object detection algorithms. The limitation of YOLO algorithm is that it struggles with small objects within the image, for example it might have difficulties in detecting a flock of birds. This is due to the spatial constraints of the algorithm.[4]

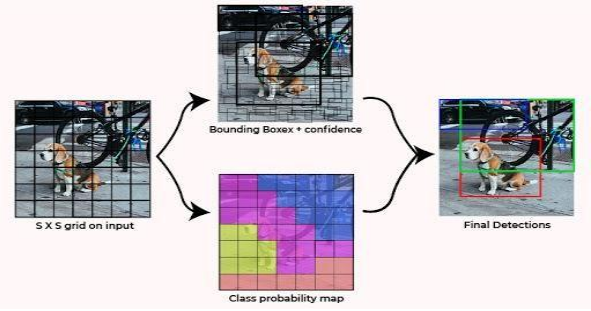


Fig I: YOLO Bounding Box, Object detection and localization

Faster R-CNN vs YOLO: The study [2] utilized a common dataset for both YOLOv5 and Faster R-CNN, containing various spacecraft features such as solar panels and satellite bodies. These features were required to be identifiable by humans and relate to real-life satellite components. Each image in the dataset was unique. The objective was to evaluate and ascertain the superior algorithm for detecting anomalies in space. Results from the testing dataset favored Faster R-CNN over YOLOv5, although YOLOv5 exhibited a tenfold higher inference rate, rendering it the preferred choice for real-time object detection in this context. A similar investigation by [4] yielded comparable findings,

with YOLO proving to be a cleaner and more efficient object detection solution, particularly due to its end-to-end training capability. While both algorithms demonstrated reasonably high accuracy, YOLO occasionally surpassed Faster R-CNN in accuracy, speed, and efficiency. Its single-shot approach and ability to directly train on full images make it suitable for real-time detection in both images and videos. YOLO's superior generalization of objects compared to Faster R-CNN positions it as a more dependable, swift, and robust algorithm. These prominent advantages strongly advocate for its adoption.

With YOLO established as the preferred algorithm for the project, the remaining consideration is selecting the appropriate version. YOLO has evolved through various iterations, including YOLOv1, v2, v3, v4, v5, v6, v7, and v8. Notably, versions v3, v5, and the latest, v8, stand out. Benchmark tests conducted by Stereolabs concluded that all three versions (v5, v7, and v8) perform well on the Jetson Orin platform.[5]

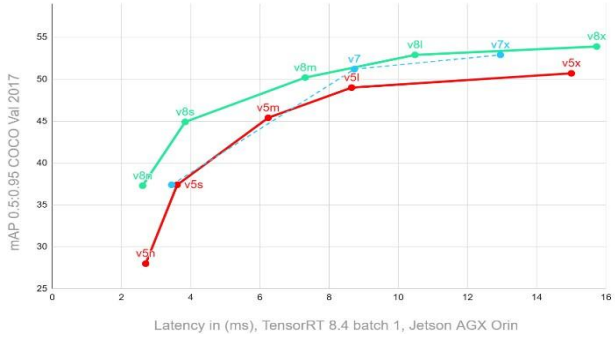


Fig II: Comparison between YOLO versions

III. METHODOLOGY

A. Data Collection and Preparation

Considering the nature of the project finding large datasets seemed difficult thus the process chosen was to integrate several smaller datasets to form one large dataset to fit all our needs. To start, curated a massive variety of datasets from different domains. After having Skimmed through them to identify those that are exactly suitable for the task's needs and purposes. Download them in the compatible format and, for instance, YOLO, to fit the annotation and training pipeline. Using the data set acquired, classes and annotation had to be redone to fit our needs and help standardize the process.



Fig III: Sample annotated image

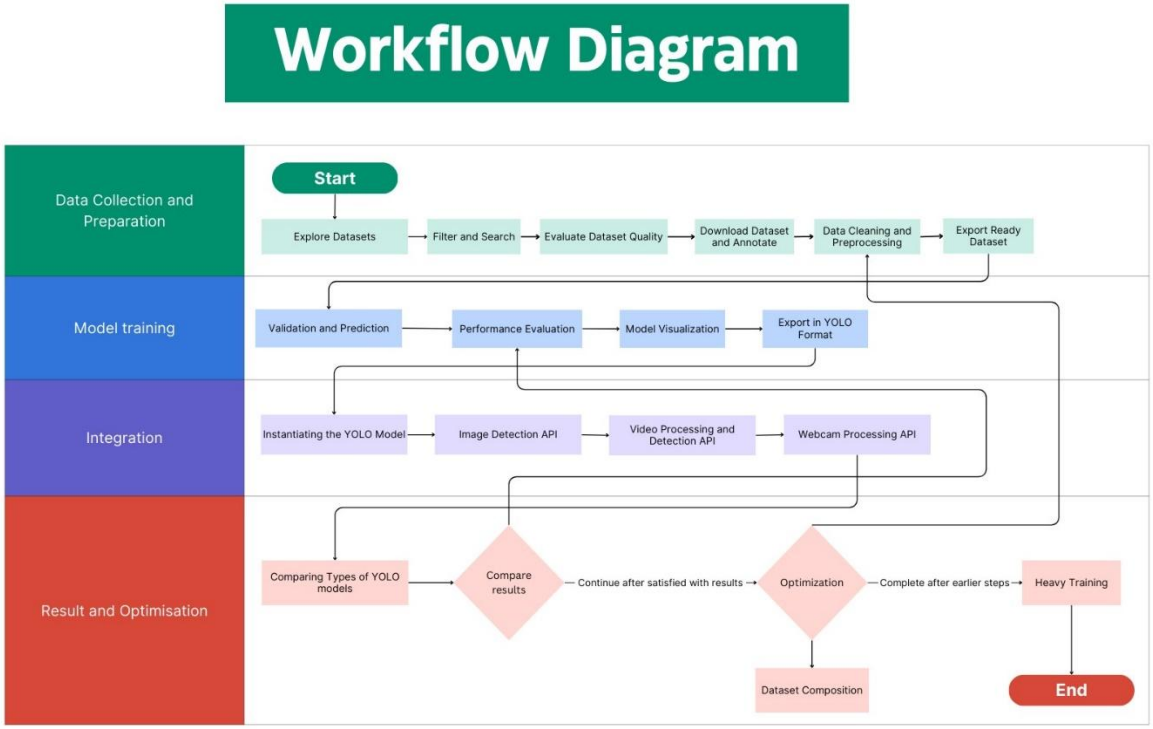


Fig IV: Workflow of the entire process

B. Model training:

The training process involved configuring the YOLO model with specific parameters tailored to the characteristics of the dataset and the requirements of the use case.

Training settings for YOLO models refer to the various hyperparameters and configurations used to train the model on a dataset. These settings can affect the model's performance, speed, and accuracy. Some common YOLO training settings include the batch size, learning rate, momentum, and weight decay.

1) Some useful parameters:

a) Epochs: Total number of training epochs. Each epoch represents a full pass over the entire dataset. Adjusting this value can affect training duration and model performance.

b) Device: Specifies the computational device(s) for training: a single GPU (device=0), multiple GPUs (device=0,1), CPU (device=cpu), or MPS for Apple silicon (device=mps).

The model training was conducted using the YOLO architecture with the specified dataset and training parameters. The training process involved optimizing the model's weights and biases to minimize the loss function, thereby improving the model's ability to accurately detect objects of interest.

2) Validation and prediction:

a) Inference Sources: YOLO can process different types of input sources for inference. The sources include static images, video streams, and various data formats. By using Streaming for processing videos or live streams it creates a generator of results instead of loading all frames into memory.

b) image path, video file, directory, URL, or device ID for live feeds. Supports a wide range of formats and sources, enabling flexible application across different types of input. conf float 0.25 Sets the minimum confidence threshold for detections. Objects detected with confidence below this threshold will be disregarded. Adjusting this value can help reduce false positives.

c) Performance Evaluation: The trained model's performance was evaluated using various metrics such as precision, recall, and mean average precision (mAP) to assess its effectiveness in detecting objects in surveillance environments.

3) How to improve training. How v8 assists in training the dataset effectively:

The output provided during the training of YOLO consists of several key metrics and parameters that are essential for evaluating the model's performance and progress throughout the training process. Let's break down each parameter and its significance in improving the training:

a) Epoch: Represents one complete pass through the dataset during training.

b) GPU_mem: Indicates GPU memory utilization to prevent memory overflow.

c) box_loss, cls_loss, dfl_loss: These parameters represent the loss values calculated during the training process. Loss functions quantify the difference between the predicted output of the model and the ground truth labels. Lower values of box_loss, cls_loss, and dfl_loss indicate better convergence of the model during training.

d) Instances: Total number of objects detected in the training dataset during the epoch.

e) Size: Input image size, impacting the model's ability to detect objects at different scales.

f) Performance Metrics: Precision (P), Recall (R), mAP50, mAP50-95:

These metrics evaluate the performance of the trained model in terms of object detection accuracy. Precision measures the ratio of correctly predicted objects to the total predicted objects, while recall measures the ratio of correctly predicted objects to the total ground truth objects. mAP (mean Average Precision) is a composite metric that considers precision recall curves across different confidence thresholds. mAP50 and mAP50-95 represent the mean Average Precision calculated at different IoU (Intersection over Union) thresholds.

Improvements in these parameters, such as reducing loss values, increasing the number of instances detected, optimizing input image size, and enhancing precision, recall, and mAP scores, contribute to a more accurate and robust object detection model.

4) Model Visualization:

a) Confusion Matrix: A table summarizing the model's predicted versus actual classifications, helping visualize its performance in terms of true positives, true negatives, false positives, and false negatives.

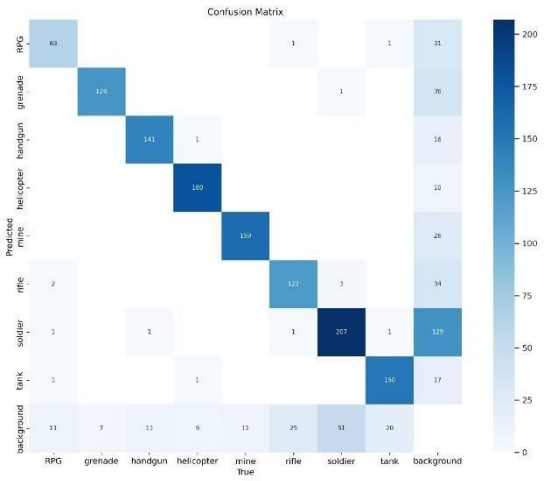


Fig V: Confusion matrix of early-stage training

b) F1 Curve: A graphical representation of the F1 score against different threshold values, aiding in determining an optimal threshold for classification by visualizing the trade-off between precision and recall.

c) **P Curve (Precision-Recall Curve):** A curve plotting precision against recall for varying threshold values, providing insights into the model's performance in identifying true positives while minimizing false positives.

d) **PR Curve (Precision-Recall Curve):** Similar to the P Curve, but emphasizing the precisionrecall trade-off for different threshold values, helping visualize the model's performance in terms of precision and recall.

e) **R Curve (Recall Curve):** A curve plotting recall against different threshold values, illustrating the model's ability to correctly identify positive instances while minimizing false negatives.

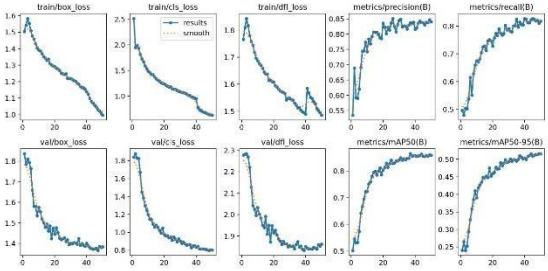


Fig VI: Result Graphs

5) Export Formats:

Among the export formats for the YOLO model, the most popular ones are generally those that cater to widely used frameworks and deployment scenarios. Based on the popularity and widespread adoption in the deep learning community, the following export formats are considered among the most popular:

a) **PyTorch:** Preferred for PyTorch-based applications due to its widespread adoption

b) **ONNX:** Offers interoperability across frameworks and hardware platforms.

c) **TensorRT:** Optimized for NVIDIA GPUs, ideal for real-time applications.

d) **TF Lite:** Efficient for mobile and embedded devices.

e) **TF.js:** Enables client-side object detection in web applications.

6) Working with the results:

a) **Image Processing Pipeline Overview:**

i) Input Image: Initially provided in JPG/JPEG format.

ii) Annotated Frame: After processing, the results are plotted onto the input image, producing an annotated_frame, typically represented as a NumPy array with annotations.

iii) Encoded Image: The annotated_frame is then encoded into a bytes object called buffer, utilizing the JPEG format for compression, commonly used for image storage.

iv) **Base64 Encoding:** Finally, the encoded image data in buffer is base64 encoded to generate base64_data, a string representation of the image suitable for transmission over text-based protocols like JSON.

b) **Useful parameters of results which affect the prediction:** Optimize object detection with customizable parameters like confidence and IoU thresholds to filter detections and reduce false positives. Fine-tune detection by filtering specific class IDs and enabling features like class-agnostic NMS for multi-class scenarios. Add test-time augmentation while visualizing model features for debugging and interpretation. Streamline video processing with options like stream buffering and frame stride adjustments to balance speed and temporal resolution.

c) **Customized Output Formats:** Tailor output formats to meet specific application requirements, allowing flexibility in how detection results are presented or utilized downstream. This customization could include options for different file formats, structured data formats, or customizable metadata.

d) **Dynamic Model Adaptation***:* Implement mechanisms for dynamic adjustment of model parameters or architectures based on evolving requirements or environmental conditions. This adaptability enables the model to effectively handle changes in data distribution, varying levels of complexity in input scenes, or specific operational contexts, ensuring performance across different scenarios without the need for manual intervention.

C. Integration:

The project employs modern technology for frontend integration with the model, facilitating interaction with the system which involves state management. With the integration of strong user authentication and authorization processes are streamlined, ensuring security and user management functionalities. Furthermore, RESTful API frameworks is utilized on the backend, making the system with API capabilities and secure and scalable communication between the frontend and backend components. It makes use of the power of the YOLO object detection model to accurately detect and classify objects within uploaded images, enhancing the overall functionality and user experience of the application.

1) Instantiating the YOLO Model: Before any process is to be done the YOLO model is instantiated with the pre-trained weights loaded from the specified file path. This model is responsible for predicting objects in images with a confidence threshold of 0.5.

2) Image Detection: Users have the capability to provide an image input, which is subsequently rendered by the

frontend and forwarded to the model via an API call for processing. Initially, both the image container and result states are set to null. Upon rendering of an image, the container state is updated, indicating readiness for processing. Upon completion of the detection process, the resulting image is transmitted back to the frontend in JSON format, containing values and labels of the detected objects presented in a tabular format.

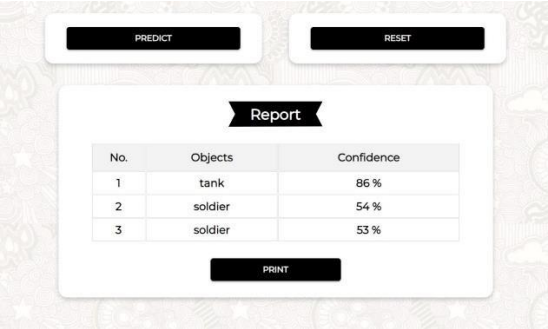


Fig VII: Report Section

3) **Video Processing and Detection:** The backend mechanism established to facilitate the processing of uploaded videos, enables real-time object detection. Upon receiving a video file, the system initiates processing by temporarily storing the video and extracting frames for further analysis. Each extracted frame undergoes object tracking utilizing the pre-trained model, with a specified confidence threshold applied. Tracked objects are annotated on the frames, which are subsequently transmitted to the frontend in real-time, facilitated by a server-to-client push notifications' endpoint. After the complete processing of the video, a report is being provided containing the tracked objects and a timestamp of when the object was detected. Due to this function, the scripts should be done is a well-structured format to prevent mixing of any data, which is to be sent to the frontend. The fact that the backend takes a bit of time to send the required data, the output video streams are down-framed to provide smooth experience.

4) **Webcam Processing:** In alignment with the offline video detection system, a real-time object detection capability has been integrated to handle webcam streams. Frames captured from the webcam undergo processing utilizing the same sophisticated object tracking model employed for offline videos. The annotated frames, indicating tracked objects, are then continuously streamed to the frontend in real-time To ensure seamless and uninterrupted delivery of annotated frames to the frontend, a server-to-client push notification endpoint has been established. This endpoint serves as a conduit for the dynamic streaming of annotated frames, allowing users to observe tracked objects as they appear in the webcam feed in real-time. Furthermore, a real-time report generation has also been provided, which include the tracked object, and its real-world timestamp. To further supervise the system, all the frames which include the involvement of the tracked object are being saved as a video clip in the system.

D. Result and Optimization:

1) **Comparison:** A comparison was conducted between different YOLO weights underwent training for 300 epochs and for 200 epochs. This difference in training duration was attributed to the observations made during early training process.

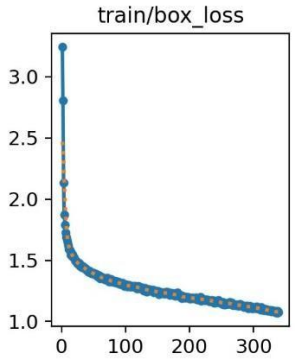


Fig VIII: training/box loss 1

However, upon evaluation, it was found that the later used version gave better results.

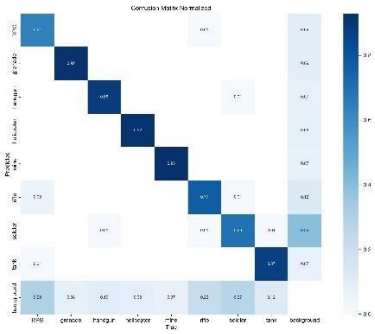


Fig IX: Confusion matrix 1

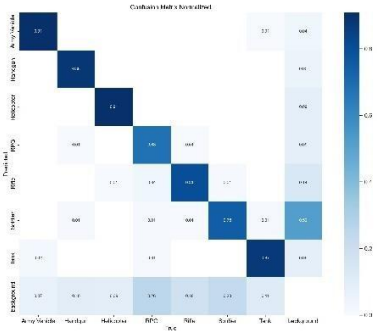


Fig X: Confusion matrix 2

2) **Optimization:**

a) Heavy Training: Due to the complex nature of the project, it necessitates extensive training to achieve satisfactory results. The analysis from Fig X underscores this requirement, revealing that even with a model of

acceptable quality was challenging to obtain within 200 epochs. This highlights the need for prolonged training durations to ensure the model's accuracy & effectiveness in real-world applications.

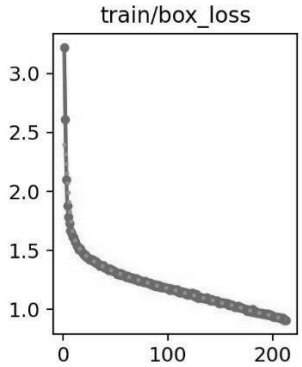


Fig XI: training/box loss 2

b) Model Preprocessing: To enhance the efficiency of the model, specific preprocessing steps were executed during the initial stages of dataset creation. These steps included resizing and segmentation, aimed at optimizing the dataset for training. By performing preprocessing tasks beforehand, the model can focus more effectively on learning relevant features during training, thereby streamlining the overall optimization process.

c) Dataset Composition: The dataset composition is designed to encompass a comprehensive range of scenarios. Each class within the dataset comprises approximately 1000 images, divided into 500 close-up shots, 250 long-distance shots, and 250 shots capturing the object from various angles. This diverse representation enables the model to generalize better and perform effectively across different contexts. Moreover, the emphasis is on incorporating images from real-life environments rather than relying solely on stock imagery with plain backgrounds ensures that the model is trained on data that closely mirrors real-world conditions, thereby enhancing its robustness and adaptability.



Fig XII: Angled and Close-up shots from test *E.*

Conclusion:

The "Secureye" project is a big step forward in protecting against military intrusions and making defense and military areas safer. The use of the YOLO model, tailored for precise object detection across diverse input streams.

Additionally, by seamlessly integrating YOLO into the project's framework, the system not only ensures efficient and accurate surveillance but also provides real-time alerts for potential threats. Furthermore, YOLO tracking enhances consistency in object detection, offering continuous monitoring over identified objects. This smart technology helps make security stronger and prevents problems in sensitive places.

Overall, "Secureye" is a significant achievement in enhancing security and managing threats in defense and military domains, showcasing technology's role in safeguarding against intruders and reducing risks.

REFERENCES

[1] Bhujbal, Kunal and Barahate, Sachin, "Custom Object detection Based on Regional Convolutional Neural Network & YOLOv3 With DJI Tello Programmable Drone," (May 5, 2022). Proceedings of the 7th International Conference on Innovations and Research in Technology and Engineering (ICIRTE-2022), organized by VPPCOE & VA, Mumbai-22, INDIA, https://arxiv.org/abs/2301.09056 [cs.RO]

[2] J. Han, D. Zhang, G. Cheng, N. Liu and D. Xu, "Advanced Deep Learning Techniques for Salient and Category-Specific Object Detection: A Survey," in IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 84-100, Jan. 2018, doi: 10.1109/MSP.2017.2749125. keywords: {Object detection; Convolution; Computer vision; Feature extraction; Visualization; Computer architecture; Machine learning},

[3] Fiza Joiya, "OBJECT DETECTION: YOLO VS FASTER R-CNN," Research Student, Department of Information Technology, B.K. Birla College Of Art, Science And Commerce (Autonomous), Kalyan, Thane, Maharashtra, India. DOI: https://www.doi.org/10.56726/IRJMETS30226

[4] StereoLab, "Performance Benchmark of YOLOv5, v7, and v8," DOI : https://www.stereolabs.com/blog/performance-of-yolo-v5-v7-and-v8

[5] Medium, "R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms," DOI : https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-

36d53571365e.

[7] Pranav Raut, Ashish Nagarse, Shreemesh Mohite, Sakshi Supe, Dr.Yogesh Karpate, Prof.Sachin Deshpande, "AI-Guided Energy Optimization In Hvac System", SJIS, vol. 35, no. 2, pp. 200–210, Jun. 2023.

❑ ❑ ❑ ❑ ❑ ❑

AET 002

Leveraging LSTM Ensembles for Enhanced Remaining Useful Life Prediction of Filter Mats

Sahil Bhoir
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
sahilbhoir01@gmail.com

Satej Bhave
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
satej.bhave@vit.edu.in

Kamlesh Sable
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
kamlesh.sable@vit.edu.in

Ved Dahale
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
ved.dahale@vit.edu.in

Prof. Devendra Pandit
Department of Computer Engineering
Vidyalankar Institute of Technology, Mumbai, India
devendra.pandit@vit.edu.in

***Abstract*— Predicting Remaining Useful Life (RUL) is crucial for optimizing maintenance strategies and ensuring the reliability of industrial equipment. This paper investigates the application of Long Short-Term Memory (LSTM) networks for RUL prediction of filter mats in a degradation process. We explore various LSTM architectures, including standard LSTMs, increased unit LSTMs, Bidirectional LSTMs (BLSTMs), and stacked LSTMs. Each model is trained on sensor data encompassing differential pressure, flow rate, dust feed, and time. The efficacy of these models is evaluated using Mean Squared Error (MSE) and Mean Absolute Error (MAE) on a held-out test set. Our results demonstrate that the ensemble approach achieves promising performance in predicting RUL, highlighting the potential of LSTMs for proactive maintenance applications. The paper also explores the impact of different LSTM architectures (standard LSTM, increased unit LSTM, bidirectional LSTM, stacked LSTM) on prediction accuracy, providing insights for selecting suitable models in RUL prediction tasks.**

***Keywords*— Remaining Useful Life (RUL), Filter Mats, Long Short-Term Memory (LSTM), Prognostics and Health Management (PHM), Sensor Data, Machine lea, Deep Learning, Ensemble Methods.**

I. INTRODUCTION

The Remaining Useful Life (RUL) prediction is important for providing the safety, supportability, maintainability, and reliability of various modern industrial equipment. In this paper we have proposed a way to predict the Remaining Useful Life of a filter mat made of randomly oriented non-woven fiber material [9]. The remaining useful life of a filter is a critical parameter in maintenance and reliability management. By accurately forecasting the remaining useful life, proactive maintenance scheduling can be implemented, thereby preventing unexpected failures, and reducing downtime.

One of the approaches to reducing maintenance costs is known as preventive maintenance [3]. It can be considered as a proactive approach that inspects and maintains the equipment to avoid breakdowns.

There are two main approaches for performing RUL estimation which can be found across various research: model-based and data-driven methods [3]. Model-based creates a degradation model based on statistics to predict the degradation of the system. Model-based methods require deep knowledge of component behavior, making them challenging for complex systems. They often rely on biased assumptions and need customization for different machinery. On the other hand, data-driven methods like Deep Learning don't need prior expertise and excel with quality data. DL has gained traction in various fields, including Prognostics and Health Management (PHM).

The most difficult step of this process is getting a labeled dataset of machine life cycle and failure events as corporations tend to guard the data and consider it proprietary. Also, there tends to be a vast amount of data that is generated from the sensors, and processing such a vast amount of data requires expertise and special infrastructure.

In this paper, we will look at various possible Machine Learning and Deep Learning models and test their suitability for performing the task at hand as they are well suited for handling sequential data. The main focus would be on Long Short-Term Memory (LSTM) networks [1] and their various architectures. The utilization of advanced techniques such as LSTM ensembles can potentially enhance the accuracy of predictions, contributing to improved maintenance strategies and operational cost savings. In this paper, we investigate the efficacy of various Machine Learning and Deep Learning models, particularly focusing on Long Short-Term Memory (LSTM) networks and their distinct architectures, including multiple layers of LSTM units with varying complexities. By leveraging advanced techniques such as LSTM ensembles [11] comprising models with differing structures, we aim to enhance prediction accuracy. These model variations offer insights into optimizing maintenance strategies and reducing operational costs in the context under study.

II. DATASET DESCRIPTION

A. Data

This dataset stems from a degradation process relevant to Prognostics and Health Management (PHM) applications. It is part of a collection of datasets reflecting real-world scenarios with incomplete data coverage, uploaded continuously by the user “Prognostics @HSE” on Kaggle. The observed

degradation process involves filter clogging during the separation of solid particles from gas, simulated using a test bench. Integrating physical models from literature could complement data-driven models, enhancing theory-guide data science or informed machine learning approaches.

The item under examination is a filter mat composed of randomly arranged non-woven fiber material. These mats are extracted from a flat material and positioned centrally within the filtration chamber. [10]. The test material has a square flow area of 6131mm^2.

Mean Fiber Diameter	23 micrometer
Filter Area	6131 mm^2
Filter Thickness	20 mm
Filter Packing Density	0.014 – 0.0165
Clean Filter Pressure Drop	25 Pa at flow of 1080 m^3/(h.m^2)

Table 1. Filter Description

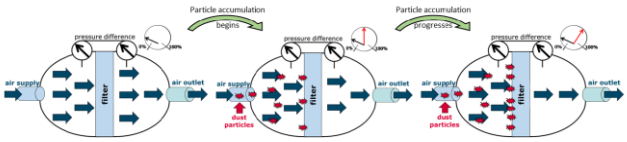


Fig. 1 Operating Principle of Testing Bench [10]

B. Test particles and particle feeding:

To evaluate filter lifespan, standardized Arizona test dust (ISO 12103-1) of particle size A3 is employed for loading. This dust adheres to defined distribution functions regarding particle size and chemical composition.

Dust particles are introduced into the compressed air system of the test bench via a nozzle for powdery solids based on the Venturi principle. The dust is contained within a cylinder and fed to the nozzle by an electrically driven spindle actuator, as depicted in Figure 2. Consequently, the load profile can be adjusted by altering the spindle actuator's velocity. This adjustment regulates the quantity of particles added per unit of time, independently of differential pressure and flow. Within the dataset, the current material feed is expressed in mm3/s. Coupled with the bulk density of the A3 Medium Test Dust at 1.025g/cm3, the mass of the fed dust can be computed [10].

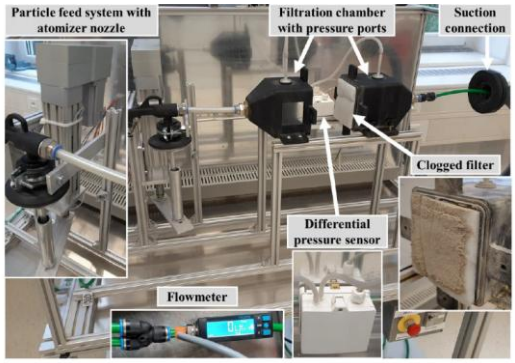


Fig 2. Filtration test bench with its particle feed system and filtration chamber [10]

In this research, we leverage a custom filter degradation dataset specifically designed to emulate the challenges of real-world sensor variations. This dataset goes beyond simply utilizing real sensor data. It incorporates key parameters like differential pressure, flow rate, dust feed, time, and remaining useful life (RUL). However, to reflect the diverse sensor configurations encountered in practice, the data is meticulously crafted to include controlled noise and bias injection. This approach empowers the trained models to exhibit robustness and adaptability – crucial qualities for navigating the inherent variability in data quality often observed in filter health monitoring systems.

C. Data Structure :

The data compiled is separated into two datasets, the test dataset and the training dataset. These contain the individual measurement series of life cycle tests. The individual life cycle tests have the same structural design. Therefore, they contain all necessary measurement data and setting parameters. Each data point includes sensor measurements (pressure, flow rate, time, dust intake) and settings (sampling frequency, bias type). "Data_No" links data to specific filter runs. Training data also contains the remaining useful life (RUL) for each point, while test data only provides RUL for the final measurement (right-censored). The structure of a life cycle test is in the table below:

Data No.	Unique Identifier
Differential Pressure	The pressure difference across the filter
Flow Rate	Rate of airflow through the filter
Dust Feed	Rate of dust intake by the filter
Time	Timestamp of the measurement
RUL (Train_data)	Remaining Useful Life of the filter
Sampling	Data Sampling Parameter
Bias Type (Train_data)	Type of bias applied

Table 2. Dataset Description

D. Exploratory Data Analysis :

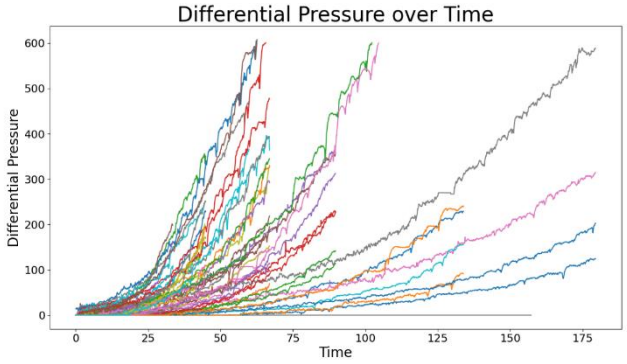


Fig 3. Differential pressure was measured over time for each filter run.

We can see some time series reaches 600. This is the run-to-fail data in the train set. On these curves, we want to predict when it reaches 600.

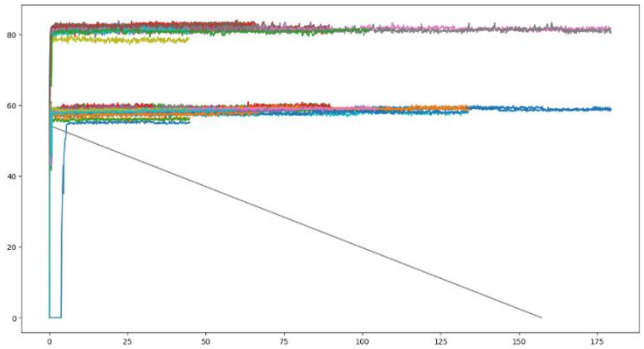


Fig 4. Visualizing Flow Rate over time for each filter run

```
corr_matrix = test_df[["Differential_pressure","Flow_rate","Dust_feed","RUL","Dust"]].corr()
corr_matrix.style.background_gradient(cmap='coolwarm')
```

	Differential_pressure	Flow_rate	Dust_feed	RUL	Dust
Differential_pressure	1.000000	0.162771	-0.082623	-0.458465	-0.079288
Flow_rate	0.162771	1.000000	-0.411261	0.273620	0.582027
Dust_feed	-0.082623	-0.411261	1.000000	-0.664177	-0.212978
RUL	-0.458465	0.273620	-0.664177	1.000000	0.323812
Dust	-0.079288	0.582027	-0.212978	0.323812	1.000000

Fig 5. Correlation Analysis

The Correlation Analysis reveals:

- **Differential Pressure:**
  - Has a positive correlation with Dust Feed. This means that as the dust feed increases, the differential pressure also tends to increase. This is intuitive as more dust entering the filter would likely lead to a greater pressure difference across it.
  - Has a weak negative correlation with Flow Rate. This suggests that there might be a slight tendency for differential pressure to decrease as the flow rate increases. Potentially, higher flow rates could help reduce pressure buildup.
- **Flow Rate:**
  - Has a weak positive correlation with Dust. This could indicate that as the flow rate increases, there might be a slight increase in dust accumulation. However, the correlation is weak, so this relationship is not very strong.
- **Dust Feed:**
  - Has a weak negative correlation with Dust. This is somewhat counterintuitive. A possible explanation could be that higher dust feed rates lead to the filter becoming overloaded, reducing its ability to capture dust particles, and resulting in less dust measured. However, the weakness of the correlation coefficient makes it difficult to draw strong conclusions.
  - From the inferences drawn from the correlation matrix it is clear that the features 'Differential Pressure' and 'Flow Rate' coupled with the time stamp have the most direct correlation with RUL and thus, these are the features that will be considered as the input for the models.

II. METHODOLOGY

In this section, we provide an overview of the methodology employed for constructing and training multiple Long Short-Term Memory (LSTM) neural network models [5]. The models discussed herein are designed to predict the Remaining Useful Life (RUL) of industrial components based on sensor data [3]. Each model encompasses a unique architecture configuration, aimed at exploring the impact of varying LSTM layer sizes and dense layer configurations on predictive performance. We detail the architecture of each model and outline the training process, including hyperparameter optimization and evaluation metrics used to assess model performance. This comprehensive methodology serves as a foundation for comparing the efficacy of different LSTM configurations in predictive maintenance applications.

A. Preprocessing of Data

Effective utilization of machine learning algorithms often requires preprocessing the raw data to ensure optimal performance. This section describes the steps applied to the sensor data for RUL prediction using LSTMs.

1. Target Variable Definition: The target variable for the prediction task was set as the 'RUL' (Remaining Useful Life).
2. Data Scaling: Min-Max scaling was employed to normalize the sensor data. This step transforms all features within a range of 0 to 1. Normalization is crucial for LSTM networks as it ensures all features contribute equally during training and prevents features with larger scales from dominating the learning process. The Min-Max Scaler scales each feature x in the dataset to (x - min) / (max - min), where min and max are the minimum and maximum values of that specific feature across all data points.
3. Sequence Creation for LSTMs: Since LSTMs process sequential data, the sensor data was transformed into sequences. The create_sequence function accomplishes this by iterating through the data with a predefined sequence length (seq_length). For each sequence:
   a. The function extracts a window of data points based on the sequence length (i to i + seq_length). This window excludes the target variable (RUL) as the LSTM aims to predict the RUL based on the preceding sequence.
   b. The extracted window of sensor measurements forms the input sequence (X).
   c. The actual RUL value at the end of the sequence (i + seq_length) is designated as the target variable (y). This process is repeated for all possible sequences within the data, resulting in a collection of input sequences (X) and their corresponding target RUL values (y).

```
features = ['Differential_pressure', 'Flow_rate', 'Time']
target = 'RUL'
data = df[features + [target]]
scaler = MinMaxScaler()
data_scaled = scaler.fit_transform(data)
def create_sequences(data, seq_length):
    X = []
    y = []
    for i in range(len(data) - seq_length):
        X.append(data[i:i+seq_length, :-1])
        y.append(data[i+seq_length, -1])
    return np.array(X), np.array(y)
```

Fig 6. Data Preprocessing

By following these steps, the sensor data is prepared in a format suitable for training the LSTM models. The scaling ensures all the features contribute meaningfully, and the sequence creation leverages the time-series nature of the sensor data to capture temporal relationships for accurate predictions.

```
Differential_pressure      5.87746
Flow_rate                 55.722852
Time                          6.9
RUL                          67.4
Name: 434, dtype: object

Scaled and Sequenced data:
Features: ['Differential_pressure', 'Flow_rate', 'Time']
Target: RUL
Scaled Data Shape: (39414, 4)
Scaled Data:
 [4.71771649e-04 6.20596552e-01 0.00000000e+00 1.71068427e-01]
Features Sequences :
 [[4.71771649e-04 6.20596552e-01 0.00000000e+00]]
Target : 0.17076830732292916
```

Fig 7. The result of preprocessing a certain row in the dataset

B. Broad overview of LSTM

Long Short-Term Memory is an evolution on the regular Recurrent Neural Network (RNN). The problem with RNNs is that they suffer from vanishing and exploding gradients. As a result, they struggle to capture long-term dependencies in sequential data (Rao et al., 2018). To address this limitation, LSTM introduces a memory unit and a gate mechanism that determines how information is stored and updated in the memory cell [1]. LSTM is particularly effective for tasks involving time-series data, natural language processing, speech recognition, and more. At its core, an LSTM unit consists of a cell state that serves as a memory along with three gating mechanisms as shown in Fig 6.: the forget gate, input gate, and output gate. These gates regulate the flow of information into and out of the cell state, allowing LSTMs to selectively remember or forget information over time.

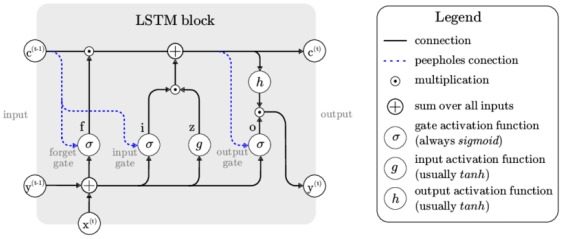


Fig 8. Regular LSTM cell state architecture

1. **Forget Gate (f_t):** This gate decides which information from the cell state to discard. It takes the previous hidden state (h_{t-1}) and the current input (x_t) as input, applies a sigmoid activation function, and outputs a forget vector (f_t) that determines what information to forget from the cell state. The forget gate activation function is always sigmoid.
2. **Input Gate (i_t) and Candidate Value ($\sim C_t$):** The input gate decides which new information to store in the cell state. It consists of two parts:
   - **Input gate:** This part determines which values to update based on the current input (x_t) and the previous hidden state (h_{t-1}). It applies a sigmoid activation function.
   - **Candidate Value ($\sim C_t$):** This is a vector of new candidate values computed based on the current input (x_t) and previous hidden state (h_{t-1}). The candidate value activation function is usually tanh. [The legend in the image indicates that the candidate value activation function is typically tanh.]
3. **Update Cell State ($\sim C_t$):** This step combines the input gate and the candidate value to update the cell state. It computes a new candidate value ($\sim C_t$) using element-wise multiplication of the output from the input gate and the candidate value vector.
4. **Output Gate (o_t):** The output gate decides what information from the current cell state to output. It takes the current input (x_t) and the previous hidden state (h_{t-1}) as input, applies a sigmoid activation function, and outputs an output vector (o_t). This output vector is then used to regulate the flow of information from the current cell state to the output and hidden state of the LSTM unit. The output gate activation function is always sigmoid.
5. **Hidden State (h_t):** The hidden state is the output of the LSTM unit. It contains information that has been selectively passed through the gates and is used to make predictions or pass to subsequent LSTM units.

C. Network Architectures of LSTM

In this section, we are going to explore a diverse range of LSTM architectures [12] which have all been trained on the dataset for RUL prediction. We have implemented standard LSTM as a baseline, followed by increased unit LSTM architecture that incorporates a larger number of units in the hidden layers to potentially capture more complex relationships within the sensor data. Next is the Bidirectional LSTM [7], which processes sequential information in both forward and backward directions, and stacked LSTMs, which allow the network to learn features at different time scales within the sequence.

1. Standard LSTM (Fig 9):

```
lstm1 = Sequential([
    LSTM(units=16, return_sequences=True, input_shape=(X.shape[1], X.shape[2])),
    LSTM(units=32, return_sequences=True),
    LSTM(units=16, return_sequences=False),
    Dense(units=16),
    Dropout(0.2),
    Dense(units=1)
])
lstm1.compile(loss='mean_squared_error', optimizer='adam')
```

Fig 9. Sequential LSTM with 3 layers

- Architecture: This model comprises three LSTM layers stacked sequentially.
- Layer Configuration:
  - First LSTM layer: It processes input sequences and outputs a sequence of hidden states for each time step (return_sequences=True).
  - Second LSTM layer: Similar to the first layer, it returns a sequence of hidden states.
  - Third LSTM layer: This layer aggregates the sequential information and produces a single vector representation for each sequence (return_sequences=False).
- Purpose: The standard LSTM architecture is effective for capturing temporal dependencies within the data and generating meaningful representations.

2. Increased Unit LSTM (Fig 10):

```
lstm2 = Sequential([
    LSTM(units=32, return_sequences=True, input_shape=(X.shape[1], X.shape[2])),
    LSTM(units=64, return_sequences=True),
    LSTM(units=32, return_sequences=False),
    Dense(units=32),
    Dropout(0.2),
    Dense(units=1)
])
lstm2.compile(loss='mean_squared_error', optimizer='adam')
```

Fig 10. Sequential LSTM with 3 layers

- Architecture: This model is similar to the standard LSTM but employs a higher number of units in the middle layers (32 and 64 units).
- Layer Configuration: First and second LSTM layers: These layers have more units compared to the standard LSTM, enabling the model to capture more intricate patterns and relationships in the data.
- Purpose: Increasing the number of units allows the model to learn more complex representations, potentially improving its performance on tasks with intricate temporal dependencies.

3. Bidirectional LSTM (Fig 11):

Traditional LSTM, as used before, is a powerful tool for handling sequential data, a standard LSTM processes information in a single direction, typically from the beginning to the end of the sequence. While this allows the network to learn relationships between earlier and later elements, it does not consider information from the elements that come after a specific point in the sequence [12].

Bidirectional LSTMs (BLSTMs) address this limitation by incorporating two LSTMs that work in tandem:

1. The first LSTM processes the sequence in the forward direction (beginning to end). This allows it to capture dependencies between earlier and later elements, like a standard LSTM.
2. The second LSTM processes a reversed copy of the sequence in the backward direction (end to beginning). This captures how elements that come later in the sequence might influence the understanding of earlier elements.

The outputs from both LSTMs are then combined to create a final output for each element in the sequence. This combined output incorporates information from both directions, leading to a richer contextual understanding of the entire sequence.

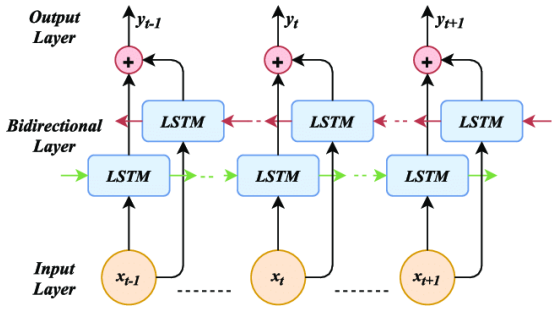


Fig 11. Bidirectional LSTM architecture

4. Stacked LSTM (Fig 12):

Stacked LSTMs incorporate multiple LSTM layers [8] stacked on top of each other. Each layer can be configured to have a different number of units and potentially utilize *return sequence = true* to process information across the entire sequence length. This architecture allows the network to learn features at various time scales within the sequence:

1. **Lower-level layers:** These layers typically have fewer units and can focus on capturing short-term dependencies within the sequence. They learn to identify rapidly changing patterns within the data.
2. **Higher-level layers:** These layers often have more units and can focus on capturing longer-term dependencies and global trends within the sequence. They integrate information from previous layers to understand the broader context.

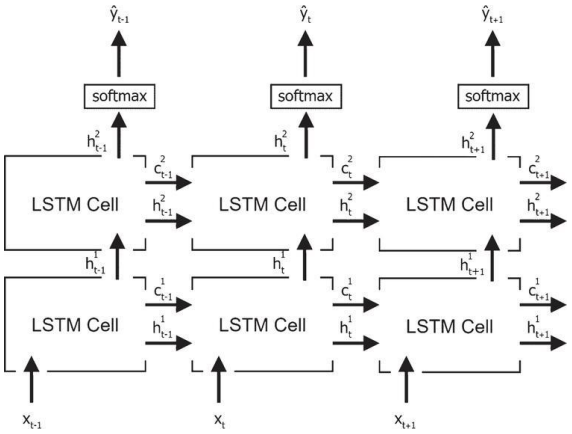


Fig 12. Stacked LSTM architecture

D. Model Training

Each model within the ensemble underwent training with the following configurations:

- **Loss Function:** Mean Squared Error (MSE) was chosen as the loss function to evaluate the model's performance during training. MSE calculates the average squared difference between predicted and actual RUL values, penalizing larger errors more heavily.
- **Optimizer:** The Adam optimizer was used for gradient descent optimization during training. Adam is an efficient optimization algorithm that adapts the learning rate for each parameter based on historical gradients, often leading to faster convergence and potentially better performance compared to traditional optimizers like stochastic gradient descent (SGD).
- **Epochs and Batch Size:** The models were trained for a maximum of 10 epochs. An epoch represents one complete pass through the entire training dataset. A batch size of 32 samples was used for training. Batch size refers to the number of samples processed by the model before updating its internal weights. We chose a batch size of 32 considering a balance between computational efficiency and gradient estimation accuracy.
- **Validation Split (10%):** A 10% validation split was employed to monitor the model's performance on unseen data during training. This technique partitions the training data into two sets: the training set used to update the model's weights and the validation set used to evaluate the model's generalization ability on unseen data. The validation loss (MSE on the validation set) is monitored throughout training.
- **Early Stopping (Patience: 3 epochs):** To prevent overfitting, early stopping was implemented with a patience of 3 epochs. Early stopping is a regularization technique that halts training if the validation loss fails to improve for a predefined number of epochs (patience in this case). This helps to prevent the model from memorizing training data patterns that might not generalize well to unseen data, potentially leading to improved performance on new RUL prediction tasks.

After the training was complete each model was saved for further use and predictions.

RESULTS

This section details the performance of the ensemble of the LSTM models for predicting Remaining Useful Life (RUL) using sensor data.

E. Evaluation Metric

The performance of the ensemble of models was evaluated using two primary metrics:

- Mean Squared Error (MSE): MSE measures the average squared difference between the actual and predicted values. Lower MSE indicates better prediction accuracy.

$$MSE = (1 / n) * \Sigma (y_{pred} - y_{true})^2$$

- Mean Absolute Error (MSE): MSE measures the absolute difference between the actual and predicted values. MAE is less sensitive to outliers compared to MSE and can provide insights into the overall magnitude of prediction errors.

$$MAE = (1 / n) * \Sigma | y_{pred} - y_{true} |$$

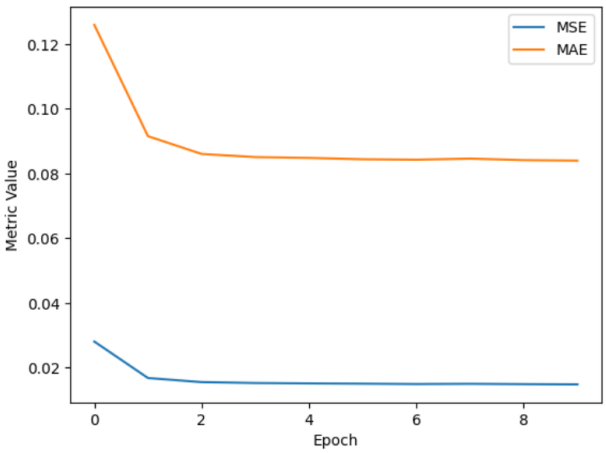


Fig 13. Performance Metrics Tracking

CONCLUSION

This paper investigated the efficacy of Long Short-Term Memory (LSTM) networks for predicting the Remaining Useful Life (RUL) of filter mats. We explored a diverse range of LSTM architectures, including standard LSTMs, increased unit LSTMs, Bidirectional LSTMs (BLSTMs), and stacked LSTMs. Each model was trained on sensor data encompassing differential pressure, flow rate, dust feed, and time stamp.

Key Findings:

- The ensemble approach utilizing various LSTM architectures achieved promising performance in predicting RUL, demonstrating the potential of LSTMs for proactive maintenance applications in filter management.
- The exploration of different architectures revealed valuable insights. While further investigation is needed to determine the optimal architecture for this specific task, our results suggest that increasing the number of units in the hidden layers (increased unit LSTM) or incorporating bidirectional processing (BLSTM) can potentially improve prediction accuracy compared to the standard LSTM architecture.

REFERENCES

[1] Hochreiter, Sepp & Schmidhuber, Jürgen. (1997). Long Short-term Memory. Neural computation. 9. 1735-80.10.1162/neco.1997.9.8.1735.

[2] Youdao Wang, Yifan Zhao, Sri Addepalli, Remaining Useful Life Prediction using Deep Learning Approaches: A Review, Procedia Manufacturing, Volume 49, 2020

[3] Kang, Ziqiu & Catal, Cagatay & Tekinerdogan, Bedir. (2021). Remaining Useful Life (RUL) Prediction of Equipment in Production Lines Using Artificial Neural Networks. Sensors. 21. 932. 10.3390/s21030932.

[4] Wang, Haitao, Jie Yang, Lichen Shi, and Ruihua Wang. 2022. "Remaining Useful Life Prediction Based on Adaptive SHRINKAGE Processing and Temporal Convolutional Network" *Sensors* 22, no. 23: 9088. https://doi.org/10.3390/s22239088

[5] Benjamin Lindemann, Timo Müller, Hannes Vietz, Nasser Jazdi, Michael Weyrich, A survey on long short-term memory networks for time series prediction, Procedia CIRP, Volume 99, 2021

[6] Sofia Yousuf, Salman A. Khan, Saqib Khursheed, Remaining useful life (RUL) regression using Long–Short Term Memory (LSTM) networks, Microelectronics Reliability, Volume 139, 2022, 114772, ISSN 0026-2714

[7] A. Zeyer, P. Doetsch, P. Voigtlaender, R. Schlüter, and H. Ney, "A comprehensive study of deep bidirectional LSTM RNNS for acoustic modelling in speech recognition," *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, USA, 2017, pp. 2462-2466, doi: 10.1109/ICASSP.2017.7952599.

[8] Muruganandam, Narendran Sobanapuram, and Umamakeswari Arumugam. 2022. "Seminal Stacked Long Short-Term Memory (SS-LSTM) Model for Forecasting Particulate Matter (PM2.5 and PM10)" *Atmosphere* 13

[9] Alimohammadi, H., Vassiljeva, kristina, Petlenkov, E., Mall Kull, T., & Thalfeldt, M. (2022). Predict the remaining useful life in HVAC filters using a hybrid strategy. *CLIMA 2022 Conference*.

[10] Mauthe, Fabian; Bakir, El Melik; Scheerer, Tim; Zeiler, Peter. Prognosis based on Varying Data Quality June 2022

[11] Ahbishek Srinivasan, Juan Carlos Andresen, Anders Holst. Ensemble Neural Networks for Remaining Useful Life (RUL) Prediction. arXiv:2309.12445

[12] Exploring different types of LSTMs, https://medium.com/analytics-vidhya/exploring-different-types-of-lstms-6109bcb037c4

❑ ❑ ❑ ❑ ❑ ❑

AET 003

Enhancing Urban Infrastructure Management through AI-Powered Defect Detection and Reporting Systems

Saarang More
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
saarang.more@vit.edu.in

Sanskruti Salunkhe
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
sanskruti.salunkhe@vit.edu.in

Vaishali Ankam
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
vaishali.ankam@vit.edu.in

Shubham Thakur
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
shubham.thakur@vit.edu.in

Dr. Sachin Bojewar
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
sachin.bojewar@vit.edu.in

***Abstract*—People who walk on roads often face difficulties due to broken roads, potholes on the road, and garbage scattered around. The process of complaining to the authorities is often long and time-consuming. A person has to physically go to a department, fill out forms, or write a letter to report the problem. This can be daunting and frustrating, and the issues may not be resolved quickly, if at all.**

However, with the advancements in technology, it can simplify this process and make it accessible to the common man. By developing a platform that allows users to upload images of the issue, along with their current location, we can quickly and efficiently report the issue to the authorities. By using Machine Learning Models to detect the type of issue from the image, such as garbage on the road, potholes, or broken roads.

This approach can speed up the process of problem resolution and enable quick and efficient reporting, which can help improve the quality of life for people using the roads. By automating this process, it can save time, reduce frustration, and promote a cleaner and safer environment.

***Keywords*—component, formatting, style, styling, insert (key words)**

I. INTRODUCTION (HEADING 1)

Modern cities face challenges in maintaining their infrastructure efficiently. Road maintenance, in particular, is a significant aspect of urban management as it directly impacts citizens' daily lives. Potholes and garbage on roads are common issues that can lead to accidents, traffic congestion, and environmental pollution. To address these concerns effectively, leveraging advanced technologies like Artificial Intelligence (AI) can significantly improve the overall maintenance process.

In recent years, AI-based detection systems have shown promising results in various fields, from healthcare to transportation. This paper focuses on the development of a YOLO-based detection system specifically tailored for urban maintenance tasks, aiming to bridge the gap between AI technologies and municipal services. By using the You Only Look Once (YOLO) algorithm, the system can quickly and accurately identify road issues such as potholes and garbage from images captured by users through a mobile application.

The integration of AI into the urban maintenance sector holds the potential to streamline complaint processes and enhance the overall efficiency of municipal services. By allowing citizens to report road issues through a user-friendly app and automatically generating complaint letters with geotagged images, the proposed system aims to expedite the resolution of maintenance problems. Additionally, by sending these automated complaint letters to the designated government platform, the system ensures that the relevant authorities are promptly informed about the issues to take appropriate actions.

II. RELATED WORK

Several research studies have delved into the utilization of AI technologies for urban maintenance and related applications. One prominent area of focus has been the use of computer vision and deep learning algorithms for object detection tasks. YOLO (You Only Look Once) is a widely adopted algorithm in this domain due to its real-time processing capabilities and high detection accuracy. Research by Redmon et al. (2016) introduced the YOLO algorithm, which has since been refined in subsequent versions such as YOLOv3 and YOLOv4. These works have demonstrated the efficacy of YOLO in various object detection applications, laying the foundation for its implementation in urban maintenance systems.

In the context of urban infrastructure maintenance, ImageNet-based deep learning models have also been explored for detecting and classifying road anomalies. For instance, Li et al. (2018) utilized convolutional neural networks (CNNs) trained on large-scale datasets to detect road potholes with high precision. Such studies highlight the potential of machine learning techniques in automating the identification of maintenance issues, paving the way for more efficient maintenance processes in urban environments.

Furthermore, the integration of geotagging technology with AI-based detection systems has been a subject of interest in the context of smart city applications. Geotagging enables the association of geographical location information with digital content, such as images captured for complaint reporting. Research by Zhangetal. (2019) demonstrated the effectiveness of geotagging in enhancing the spatial awareness of urban maintenance systems, enabling precise location identification of reported road issues. This integration of location-awareness with AI detection algorithms is crucial for ensuring the accuracy and relevance of maintenance complaint data.

In addition to object detection and geotagging technologies, the implementation of complaint management systems within mobile applications has been explored in urban maintenance research. Wang et al. (2017) proposed a mobile-based complaint reporting system for urban infrastructure issues, facilitating efficient communication between citizens and municipal authorities. Such systems not only streamline the reporting process for users but also provide a structured platform for authorities to track and address reported issues promptly.

Moreover, the use of AI-driven solutions for optimizing municipal services extends beyond road maintenance to include waste management and urban planning. For instance, research by Tan et al. (2020) applied machine learning algorithms to optimize garbage collection routes in urban areas, leading to significant improvements in operational efficiency and cost reduction. These studies underscore the transformative potential of AI technologies in enhancing various aspects of urban service delivery and sustainability.

III. METHODOLOGY

In the methodology of the project, the initial step entails user engagement through a dedicated mobile application tailored for reporting road-related concerns, specifically focusing on potholes and garbage accumulation. When encountering a road issue, users utilize the application to capture an image, enabling location services to geotag the image with precise geographical coordinates. This integration of geotagging ensures accurate spatial information associated with the reported problem, enhancing the overall effectiveness of the complaint system.

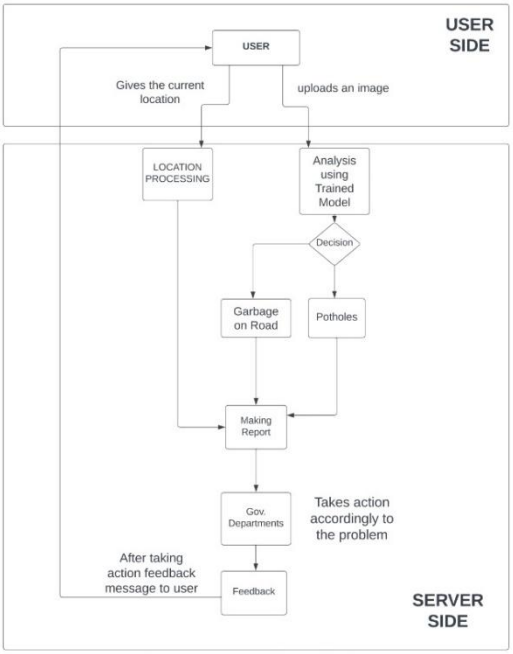
Once the user uploads the geotagged image of the road issue, it undergoes processing by the YOLOv8 model integrated into the system for object detection. The YOLOv8 algorithm, known for its real-time processing capabilities and high accuracy, swiftly analyzes the uploaded image to identify and classify the detected anomalies as either potholes or garbage. This automated object detection process enables quick and accurate identification of road issues, streamlining the complaint reporting mechanism for users.

Subsequent to the object detection phase, the system extracts the location data of the identified road anomalies from the geotagged image. This spatial information is crucial for providing precise details regarding the exact location of the reported problem, aiding municipal authorities in efficiently addressing maintenance issues. By combining geotagging technology with AI-based detection, the system ensures the contextual relevance and accuracy of the reported road complaints.

Following the detection and geotagging processes, the system generates a comprehensive complaint letter in PDF format. This complaint letter includes detailed information about the detected road anomaly, its precise geographic coordinates, and an attached image for visual reference. The automated generation of complaint letters streamlines the communication of reported issues, ensuring that municipal authorities receive clear and detailed notifications regarding maintenance concerns.

Upon the creation of the complaint letter, the system facilitates the seamless submission of this formal complaint to a designated government website dedicated to urban maintenance. By automating the transmission of complaint letters with all pertinent information, the system expedites the communication of road issues to the relevant authorities, enabling prompt and effective resolution of reported problems in urban areas. This streamlined process enhances the overall efficiency of municipal services and contributes to the maintenance of urban infrastructure.

VI. PROPOSED SYSTEM



A. Data Collection and Data Cleaning:

The first step in our methodology involved the collection of a comprehensive dataset of road images containing instances of potholes and garbage. To ensure diversity and representativeness, we collected images from various geographical locations, road types, and environmental conditions. The dataset was captured using high-resolution cameras mounted on vehicles, drones, and handheld devices. Special attention was given to capturing images under different lighting conditions (daytime, nighttime, dawn, and dusk) and weather conditions (sunny, cloudy, rainy).

Each image in the dataset was annotated with bounding boxes indicating the locations of potholes and garbage using annotation tools such as Labeling. Annotations were performed by human annotators trained to accurately identify road hazards. To enhance the quality of annotations, multiple annotators reviewed and verified each annotated image to ensure consistency and accuracy.

The final dataset consisted of thousands of annotated road images, with a balanced distribution of potholes and garbage instances. Figure 1 illustrates sample images from the collected dataset, showcasing various instances of potholes and garbage.

B. Model generation:

A pivotal component of Road Nivaaran is the integration of the YOLO (You Only Look Once) model, extensively utilized for the real-time, accurate classification of road problems from images captured by users. YOLO is a state-of-the-art, real-time object detection system that utilizes convolutional neural networks (CNNs) to detect objects.

YOLO divides the image into a grid of cells, each responsible for predicting bounding boxes and probabilities for object classes. Each grid cell predicts multiple bounding boxes and assigns confidence scores that reflect the accuracy of bounding box predictions and class predictions. This model is particularly efficient because it processes an image in a single evaluation, unlike other detection systems where a separate network is run on various regions of an image, making it significantly faster.

For Road Nivaaran, YOLO was pivotal in recognizing distinct categories of road quality issues such as potholes, cracks, and garbage. The application utilizes a pre-trained YOLO model, optimized for detecting objects in urban and semi-urban environments. The customization included retraining the last layers of the network with thousands of labeled images of road conditions typical in the target deployment areas. This training involved using high-quality, annotated images to teach the detector to accurately label new, unseen images.

Post training, the integration involved setting up the model to accept live images from user devices, process these images to detect and classify road issues, and output the classified results efficiently. Care was taken to ensure that the model's performance is optimized for mobile devices, focusing on reducing latency and power consumption while maintaining high accuracy. Parameters such as threshold values for detection confidence were meticulously tuned to balance between false positives and missed detections.

C. Model Training

Ultralytics YOLOv8 transcends the conventional boundaries of object detection models by offering a comprehensive framework that supports every stage of the machine learning model lifecycle. This includes data ingestion, model training, validation, deployment, and monitoring in real-world applications. Each component of the framework is meticulously crafted to deliver the necessary flexibility and efficiency tailored to various tasks and use-cases.

After collecting and annotating the dataset, we proceeded to train the YOLOv8 deep learning model using the annotated images. YOLOv8 is a state-of-the-art object detection model known for its efficiency and accuracy in real-time applications.

Data Preprocessing: We performed data augmentation techniques such as random horizontal flipping, rotation, and scaling to augment the dataset and improve the model's robustness to variations in road conditions.

Model Configuration: We configured the YOLOv8 model architecture based on the Darknet configuration files. The model architecture consisted of multiple convolutional layers followed by detection layers responsible for predicting bounding boxes and class probabilities.

Loss Function: We utilized the YOLO loss function, which combines localization loss, confidence loss, and classification loss, to optimize the model parameters during training.

Training Procedure: The model was trained on a high-performance GPU cluster using stochastic gradient descent (SGD) with momentum. We employed a batch size of 64 and trained the model for 100 epochs.

Hyperparameter Tuning: We fine-tuned the learning rate, momentum, and weight decay parameters through a grid search process to optimize the model's performance.

D. Geotagging and Location Extraction:

Image Metadata Analysis: Modern digital cameras and smartphones often embed metadata within image files, including location information obtained from GPS sensors. Geotagging involves parsing this metadata to extract the latitude and longitude coordinates associated with the image.

Image Processing: The uploaded images are processed to extract relevant metadata, including geolocation data. This can be achieved using libraries such as ExifTool or PIL (Python Imaging Library), which provide functions to access and extract metadata from image files.

Coordinate Extraction: Once the metadata is obtained, the latitude and longitude coordinates are extracted and stored for further use in the report generation process. These coordinates provide precise information about the location where the image was captured.

Incorporation into Reports: The extracted location coordinates are incorporated into the PDF reports generated by the application. This can be done by including the coordinates as text within the report or by visually representing them on a map using mapping APIs such as Google Maps or Leaflet.

Benefits of Geotagging in Report Generation:

- Enhanced Context: By including location information in the reports, users gain a better understanding of the spatial distribution of detected potholes or garbage. This context can help prioritize maintenance efforts and allocate resources more effectively.
- Improved Accuracy: Geotagging provides precise location data, ensuring that reports accurately reflect the geographic areas affected by infrastructure issues. This allows authorities to address problems with greater precision and efficiency.
- Data Analysis: Geotagged reports enable data analysis and visualization of infrastructure issues over time and across different locations. Patterns and trends in infrastructure maintenance can be identified, leading to more informed decision-making and policy formulation.
- Accountability and Transparency: Including location data in reports adds transparency to the process of infrastructure management. Authorities and stakeholders can track the status of reported issues and monitor progress in addressing them, fostering accountability and public trust.

E. Model Deployment and Integration:

The model deployment and integration part involves several steps to ensure that the YOLOv8 model for pothole and garbage detection is seamlessly integrated into the project and deployed for practical use. Here's a breakdown of the deployment and integration process based on the provided code:

Model Deployment and Integration Steps:

Flask Web Application Setup: The project utilizes Flask, a Python web framework, to create a web application. Flask provides a simple and flexible way to integrate machine learning models into web applications.

Routing and HTML Templating: Flask routes are defined to handle different URLs and HTTP methods, such as '/' for the home page and '/predict' for model prediction.HTML templates are used to render the web interface, allowing users to upload images and view prediction results.

Model Loading and Prediction: The YOLOv8 model for pothole and garbage detection is loaded during the initialization of the Flask application. When a user uploads an image through the web interface, Flask handles the request and passes the image to the model for prediction. The predict defect function processes the uploaded image using the YOLOv8 model and returns the predicted defect (pothole or garbage).

F. PDF Report Generation:

The PDF report generation process begins after the YOLOv8 model makes a prediction on the uploaded image, identifying and localizing potholes or garbage. The generate_pdf function is responsible for creating the PDF report, which includes relevant information about the detected defects, such as their type, location, and description.

PDF Document Initialization: The function initializes a new instance of the FPDF (Free PDF) class, which provides a framework for creating and formatting PDF documents programmatically.

Adding Content: The function adds content to the PDF document, including a subject line indicating the type of complaint (pothole or garbage) and a body text describing the detected defect's location and potential impact.

Optional Image Inclusion: If available, the function may include an image showing the detected defect within the PDF report. This visual representation enhances the report's clarity and provides additional context for the recipient.

PDF Output: Once the content is added to the PDF document, the function generates the final PDF report as a byte stream using the output method of the FPDF class. This byte stream represents the PDF document in memory and is ready for further processing and rendering.

G. Base64 Encoding and Rendering:

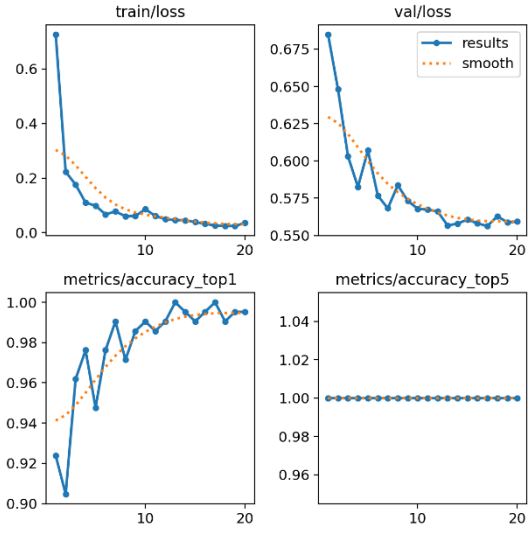
After generating the PDF report, the next step is to encode it in base64 format and render it for display in the web interface. Base64 encoding ensures that the binary PDF data is represented as ASCII text, making it suitable for transmission over the web and integration into HTML documents.

Base64 Encoding: The byte stream representing the PDF report is encoded in base64 format using the base64.b64encode function. This process converts the binary data into a text string composed of ASCII characters, which can be easily embedded within HTML documents.

Rendering in HTML: The base64-encoded PDF content is then passed to the HTML template of the web interface, where it is rendered using appropriate HTML tags. Typically, an <iframe> or <embed> element is used to display the PDF content directly within the web page, allowing users to view and download the report seamlessly.

User Interaction: Users interact with the web interface by uploading images and receiving prediction results. Upon submitting an image, the application generates the PDF report and renders it dynamically within the web page, providing users with instant access to detailed detection results.

V. RESULTS AND CONCLUSIONS



In this project, we have developed an integrated system for the detection and reporting of road defects using deep learning-based defect detection and a web-based reporting interface. The system leverages the YOLO (You Only Look Once) algorithm for real-time defect detection in images captured by users. Upon detection, the system allows users to submit complaints along with additional location information through a user-friendly web interface.

Through extensive testing and validation, we have demonstrated the effectiveness and efficiency of the system in accurately detecting various types of defects such as potholes and garbage on roads. The integration of machine learning techniques with web technologies has resulted in a robust and scalable solution that can significantly improve the road conditions and cleanliness.

This project contributes to the field of infrastructure maintenance and cleanliness management by providing a cost-effective and automated solution for defect detection and reporting. By streamlining the process of identifying and addressing infrastructure issues, our system has the potential to enhance the overall quality of urban environments and contribute to safer and more sustainable communities.

While the current implementation of the system is functional and effective, there are several avenues for future research and development:

Enhanced Defect Classification: Extend the system to perform more detailed defect classification, allowing for the identification of specific types of road defects such as cracks, debris, or structural damage.

Integration with IoT Devices: Integrate the system with IoT (Internet of Things) devices such as cameras or sensors installed on urban infrastructure to enable real-time monitoring and proactive defect detection.

Mobile Application Development: Develop a dedicated mobile application to facilitate easier and more convenient reporting of infrastructure defects by users, enabling seamless integration with the web-based system.

Predictive Maintenance: Explore the use of predictive maintenance techniques to anticipate and prevent infrastructure defects before they occur, based on historical data and machine learning models.

Geospatial Analysis: Incorporate geospatial analysis techniques to analyze the spatial distribution of reported defects and prioritize maintenance efforts based on factors such as traffic volume, road condition, and proximity to critical infrastructure.

Community Engagement: Implement features to encourage community engagement and participation in the maintenance process, such as crowdsourced reporting, gamification elements, or rewards for active contributors.

REFERENCES

[1] Redmon, Joseph, and Ali Farhadi. "YOLOv3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

[2] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao."YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv preprint arXiv:2004.10934 (2020).

[3] Liu, Wei, et al. "SSD: Single Shot MultiBox Detector." European conference on computer vision. Springer, Cham, 2016.

[4] Chatfield, Ken, et al. "Return of the Devil in the Details: Delving Deep into Convolutional Nets." arXiv preprint arXiv:1405.3531 (2014).

[5] Thrun, Sebastian, and Lorien Pratt. "Learning to learn: Introduction and overview." Learning to learn. Springer, Boston, MA, 1998.

AET 004

Automatic Timetable Generator using Genetic Algorithm

Qasim Naik
Dept of CMPN
Vidyalankar Institute of Technology Mumbai, India
qasim.naik@vit.edu.in

Shuaib Kaldhane
Dept of CMPN
Vidyalankar Institute of Technology Mumbai,India
shuaib.kahldhane@vit.edu.in

Ayan Abu Bakr Ansari
Dept of CMPN
Vidyalankar Institute of Technology Mumbai, India
ayan.ansari@vit.edu.in

Sachin Bojewar
Department of CMPN
Vidyalankar Institute of Technology Mumbai, India
sachin.bojewar@vit.edu.in

***Abstract*-Timetable creation could be a very difficult and time-engrossing function. Forming a timetable takes parts of tolerance and hours. Timetables are created for different purposes like organizing schools and colleges, making time charts for preparation and transport plans, and numerous more. Creating a timetable requires considerable time and supervision. In our paper, we have sought to shrink these challenges of creating a timetable by Genetic algorithm. By utilizing genetic algorithms, are ready to develop a schedule that is more accurate, exact, & free of human errors while also reducing the time needed to do so. The elementary stage is made up of all the common compulsory classes that have been developed and organized by a centralized body. The individual's departmental courses are found in the next phase. This timetable is ready physically, by controlling those of prior a long time, with the as it were the point of creating a realistic timetable.**

***Keywords* -- Genetic Algorithm, Selection, Cross over, Fitness**

I. INTRODUCTION

Timetabling involves creating schedules that meet specific constraints. Technology has significantly improved our lifestyles by enhancing various systems, accuracy, security, and maintenance. In contemporary times, technology predominantly refers to computers, which are arguably the greatest achievement of the past century. Computers are increasingly popular due to their ease of use, learning, high accuracy, and time efficiency.

Additionally, their maintenance is cost-effective. Consequently, traditional systems are being computerized to enhance their productivity in line with this evolutionary trend.

II. LITERATURE REVIEW

Timetable Arranging is an NP-hard issue and subsequently, polynomial time is undeniable utilizing innate calculations. It is a commonplace arranging issue that appears to be gloomy work in each academic organized once or twice a year [5]. In earlier days, timetable arranging was done physically with a single person, or many accumulate included in the task of arranging it physically, which takes an allocation of effort and time. Orchestrating timetables is one of the preeminent complex and error-prone applications [6]. Timetabling is the errand of making a timetable while satisfying some goals. There are in a general sense two sorts of restrictions, fragile restrictions, and troublesome objectives. Sensitive restrictions are those on the off chance that we harm them in arranging, the abdicate is still considerable, but troublesome objectives are those which on the off chance that we manhandle them; the timetable is no longer considerable [3]. The seen space of a timetabling issue is as well perpetual, various courses of action exist inside the seen space and few of them are not doable. Achievable arrangements here cruel those which do not abuse difficult goals and as well endeavor to fulfill fragile confinements [8]. We have to select the preeminent fitting one from feasible courses of action. The most reasonable ones here are brutal ones that do not mishandle fragile impediments to a more noticeable degree. In these amplified hard constraints have been taken care of totally and it has been ensured that soft constraints are as well taken after as much as conceivable.

III. EXISTING SYSTEM

Within the current system, the entire handle of making a timetable is done physically by taking care of all the conceivable limitations tall as well as little imperatives. The foremost dull errand of planning a timetable is from the instructive framework, particularly the College Timetabling. Here a specific teacher/lecturer is charged with the duty of making an ideal timetable manually, the teacher/lecturer has to look out for all conceivable imperatives like must have the correct data of the overall number of students, faculty individuals, the appropriate time course of action, etc. In case of the instructor's nonappearance, he/she who is capable of timetabling should organize another address at that time, which suggests a single individual separately is taking care of so much obligation that's exceptionally much time devouring as well as an excited work for the speaker.

A. Advantages of Existing System

- Can be made superior through collaboration with the different substances included because it is subjective.
- Changes are possibly made speedier on time as it were for particular conditions and not for all.

B. Disadvantages of the Existing System

- A Very Confusing and time-consuming strategy for manual creation, as well as frenzied work for teachers as misplaced exertion, is put by them for it.

IV. PROPOSED SYSTEM

The devices utilized for the arrangement and strategy Calculation and Stream chart. From the stream chart underneath, the steps carried out by the system executive are as take after: The system executive logs into the system. The executive at that point inputs the courses with their codes and the unit. At this point, the admin will keep counting until the

number of courses required has been inputted. The admin can oust a course that has been input inside the case of mishandling. After contributing the courses, it moves to the taking after page where all the address hallways or rooms that will be utilized will be inputted. After contributing these, the system at that point makes the timetable system. In order to address timetabling concerns, we are putting up a framework that would automatically produce a timetable for the founded. A timetable will be created after the course and addresses have been designed while taking into account all potential constraints and inputs.

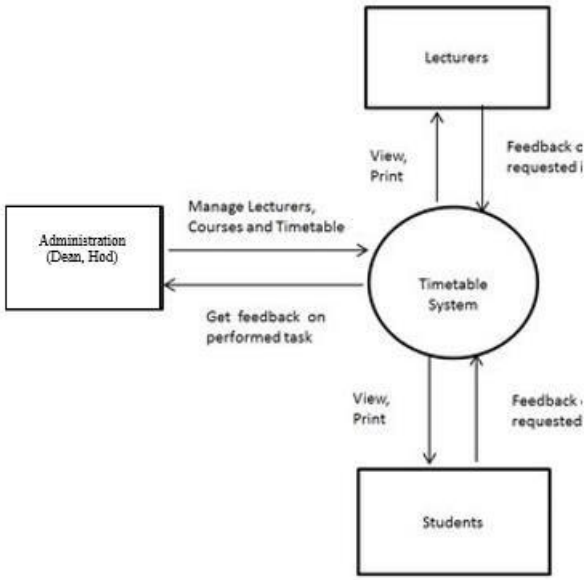


Fig 1. DFD

A. Requirement Analysis:

Conduct thorough research and stakeholder consultations to identify the specific scheduling constraints, preferences, and challenges faced by educational institutions. Gather input from administrators, faculty members, and other stakeholders to understand their needs and preferences regarding timetable generation.

B. Algorithm Development

Design and develop algorithms to automate timetable generation while considering constraints such as room availability, faculty preferences, course requirements, and institutional policies. Implement algorithms that prioritize resource optimization and conflict resolution to ensure efficient scheduling.

C. User Interface Design

Collaborate with UI/UX designers to create an intuitive and user-friendly interface for the Android application. Design interface elements that allow administrators and faculty members to easily input scheduling parameters, view generated timetables, and adjust as needed. Ensure compatibility with various Android devices and screen sizes.

D. Application Deployment

Utilize agile development methodologies to iteratively build and refine the Android application. Develop the application using appropriate programming languages and frameworks, ensuring robustness, scalability, and compatibility with Android platforms. Implement features for timetable generation, conflict resolution, resource optimization, and user interaction.

E. Integration with Existing System

Integrate the application with existing college management systems, such as student information systems and course scheduling software, to ensure seamless data exchange and interoperability. Implement APIs and data synchronization mechanisms to facilitate communication between the application and other institutional systems.

IV. DESIGN AND IMPLEMENTATION

The entire strategy of planning based on hereditary calculation is clarified in detail in this segment. A planning method is partitioned into a few imperative modules. The manual lecture-timetable planning may be a restriction fulfilment problem where we discover a solution that satisfies the set of provided requirements. Numerous solutions to the problem of creating timetables for colleges and schools have been proposed in the past. Timetabling problems can be understood using a variety of techniques learned from operational research, such as chart coloring, local look techniques like tabu look, recreated tempering, Genetic Algorithms, or backtracking-based limitation fulfilment.

The implementation of the project to include features such as events, assignments, notes, and exams enhances its functionality and utility for both teachers and students. Let's delve into the implementation details of each feature:

1. Adding Events

Implement functionality to add events such as seminars, workshops, or extracurricular activities.

Create a form to input event details including title, description, date, time, and location.

Store events in a separate data structure and provide options to view, edit, and delete events.

2. Adding Assignments

Develop functionality to add assignments, homework, or projects for students.

Create a form to input assignment details including title, description, due date, and subject.

Store assignments in a separate data structure and provide options to view, edit, and delete assignments.

3. Adding Notes

Implement functionality to add lecture notes, study materials, or reference documents.

Create a form to upload or input note details including title, description, subject, and file attachments.

Store notes in a separate data structure and provide options to view, edit, and delete notes.

4. Adding Exams

Develop functionality to add exam schedules for courses. Create a form to input exam details including course, date, time, duration, and venue.

Store exam schedules in a separate data structure and provide options to view, edit, and delete exam schedules.

5. Viewing and Managing Features

Design a unified dashboard or menu system to navigate between different features (subjects, events, assignments, notes, exams).

Provide options to filter and search for specific items (e.g., subjects for a particular day, assignments for a specific subject).

Implement sorting options to organize items based on various criteria (e.g., date, subject, priority).

6. Notification System Enhancements

Extend the notification system to include reminders for upcoming assignments, exams, and events.

Implement different notification channels (e.g., pop-up notifications, email notifications) based on user preferences.

Allow students to customize notification settings for each feature (subjects, assignments, exams, events).

7. User Authentication

Consider implementing a basic user authentication system to secure access to sensitive features such as assignments, notes, and exam schedules.

Allow teachers and students to create accounts with unique usernames and passwords.

Authenticate users during login and restrict access to features based on user roles and permissions.

8. Data Persistence and Storage:

Utilize a lightweight database such as SQLite to store application data persistently.

Implement data synchronization mechanisms to ensure consistency across multiple devices and sessions.

Provide options for data backup and restoration to prevent data loss in case of device failure or application reinstallations.

9. User Interface Design:

Focus on creating an intuitive and user-friendly interface that promotes ease of use and accessibility.

Use consistent design elements, colors, and layouts across different features to maintain visual coherence.

Incorporate feedback mechanisms such as ratings, reviews, or feedback forms to gather user input and improve the application over time.

10. Testing and Quality Assurance

Conduct rigorous testing of each feature to ensure functionality, reliability, and performance.

Perform usability testing with target users to gather feedback on the user interface, navigation flow, and overall user experience.

Address any bugs, errors, or usability issues identified during testing and debugging phases.

11. Documentation and Support:

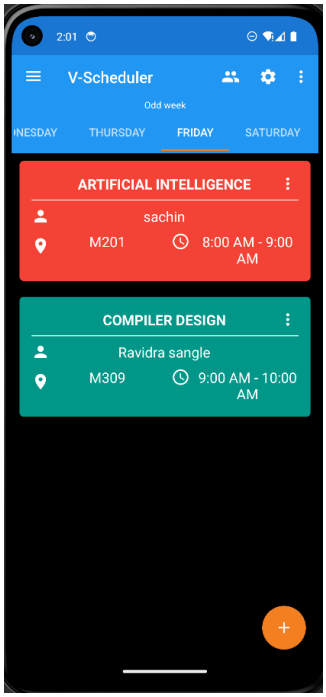
Prepare comprehensive documentation including user guides, tutorials, and FAQs to assist users in navigating and using the application.

Provide responsive customer support channels such as email, chat support, or forums to address user inquiries, issues, and feedback promptly.

VI Results

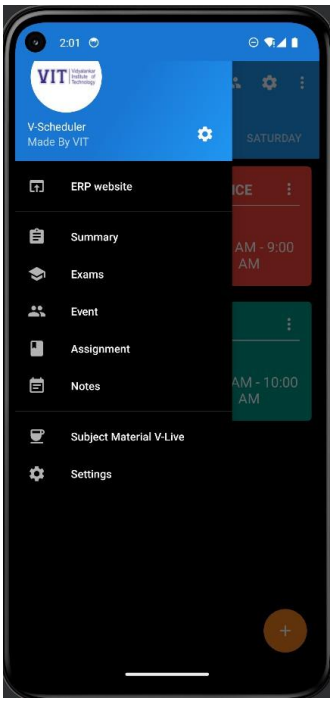
1. Home screen

On this screen the main content and the timetable is shown. It also gives a ready access to different aspects of the application. Like main menu, add new, Settings, etc.



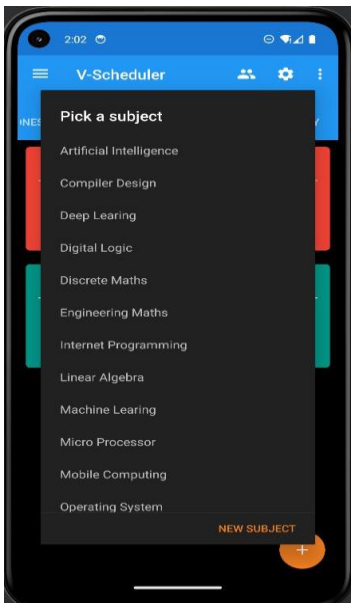
2. Main Menu

This menu shows the different components of the application and gives easy access to them. These components are Exams, Notes, Event, Assignment, Settings.



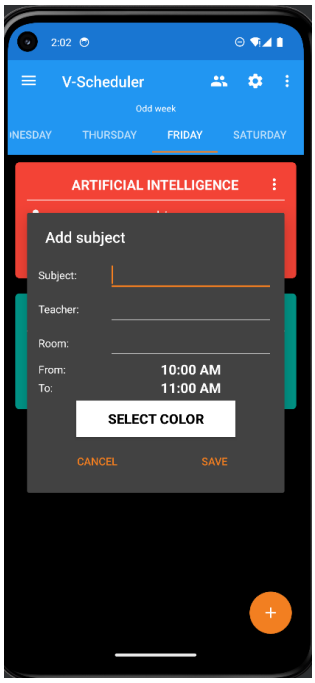
3. Add new Subject.

By clicking on the add new button the user can add different new subjects in the timetable for various days of the week and rearrange them accordingly.



4. Subject Det'ails.

The user can add details of the subjects with the time stamps of the lectures on different days.



VII. CONCLUSION

In conclusion, the development of an Android application for timetable generation in educational institutions represents a significant endeavor with the potential to streamline scheduling operations, optimize resource utilization, and enhance the overall efficiency and effectiveness of academic programs. However, this project is not without its challenges and limitations, which must be carefully considered and addressed to ensure its success and impact.

Throughout the project, various ethical considerations must guide decision-making processes to uphold principles of privacy, fairness, transparency, accountability, and respect for stakeholders' autonomy. By prioritizing these ethical principles, the project can promote trust, integrity, and responsible use of technology within the educational community.

REFERENCES

[1] Brown, M. (2018). "Design and Implementation of Mobile Applications for Educational Purposes." In Proceedings of the 15th International Conference on Mobile Systems and Pervasive Computing (MobiSys '18).

[2] Alavi, M., & Leidner, D. E. (2001). "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues." MIS Quarterly, 25(1), 107-136.

[3] Chowdhury, M., Rahman, A., & Kumar, M. (2019). "Development of an Educational Management System for Bangladesh." In Proceedings of the 2nd International Conference on Computer and Information Sciences (ICCIS '19).

[4] Chien, J. T. (2017). "Design and Implementation of a Mobile Educational Management System for Small Schools." In Proceedings of the 9th International Conference on Mobile Computing and Ubiquitous Networking (ICMU '17).

[5] Hamid, S., & Waycott, J. (2015). "Development and Implementation of a Mobile Educational Management System for Higher Education." In Proceedings of the 12th International Conference on Information Technology: New Generations (ITNG '15)

❑ ❑ ❑ ❑ ❑ ❑

AET 005

Video Surveillance for Violence Detection

Aayusha Shinde
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
aayusha.shinde@vit.edu.in

Reena Gupta
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
reena.gupta@vit.edu.in

Aakash Vaishnav
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
aakash.vaishnav@vit.edu.in

Gaurav Avhad
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
gaurav.avhad@vit.edu.in

Pankaj Vanwari
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
pankaj.vanwari@vit.edu.in

***Abstract*—Violence rates have experienced a remarkable 57% reduction over the past four decades. However, this doesn't alter the fact that instances of violence continue to occur, often going unnoticed by law enforcement. While higher authorities can implement mass control measures to mitigate violence, effectively maintaining law and order necessitates a more granular approach, akin to "micro-governance," where every activity on every street and square is monitored. To address the complex and far-reaching consequences of even small actions, we've developed a unique model and a conceptual system that employs deep learning. This model is designed to process CCTV cutting feeds, drawing inferences to determine whether a violent incident is unfolding. The envisioned system is centered on a probability-driven approach to video feed analysis, which significantly reduces the computational overhead. Instead of naively processing every CCTV video feed, the system intelligently computes and focuses its attention where there's a higher likelihood of violence, considering the potential ripple effects of such incidents. Overall, this approach represents a significant step forward in enhancing public safety by efficiently identifying and responding to violent activities in real-time, thus contributing to the prevention of violence in our communities.**

I. INTRODUCTION

The rise in violent occurrences occurring in public places has been a major concern for law enforcement and the community in recent years. These episodes, which cause varied degrees of injury and pose serious risks to public safety, range from small-scale arguments to serious violent crimes. The traditional method of keeping an eye on these places through video surveillance which mainly relies on human operators has shown to be insufficient. When faced with multiple video feeds to keep an eye on, human operators frequently find it difficult to maintain a continual state of alertness, which can result in missed opportunities and a delay in reacting to important occurrences. This challenge underscores the need for innovative solutions that can augment human monitoring capabilities with advanced technology.

The difficulties posed by human monitoring have been exacerbated by the widespread use of surveillance cameras. It is both economically unfeasible and prone to error to deploy enough human operators to continuously watch every video feed. Furthermore, it takes a lot of time to manually evaluate a large amount of video, which frequently causes reactions to accidents to wait. Innovative approaches that can complement conventional monitoring techniques and successfully solve these drawbacks are therefore desperately needed.

The goal of this project is to create an intelligent video surveillance system designed to detect incidents of violence in public places such as streets, traffic centers and other high-traffic areas. Using advanced computer vision techniques - object detection, motion analysis and pattern recognition - the system aims to significantly improve public safety. These technologies make it possible to detect potentially violent behavior by analyzing the movement and interaction of people in the camera's field of view. Deep learning allows the system to continuously improve its detection capabilities based on the data it collects, making it more effective over time. By providing a quick and accurate assessment of threats, the system helps to act in a timely manner, reducing the likelihood of violence escalating. This project seamlessly integrates with existing security measures and strengthens the efforts of security personnel by providing them with important real–time Information.

Integration of cutting-edge machine learning and computer vision technology with video surveillance systems is a potentially effective way to address these issues. By automatically analyzing video footage, these technologies make it possible to understand behavioral semantics in real time by analyzing people's gestures and body language. Real-time identification of actions and behaviors, particularly those suggestive of possible aggression, is possible for surveillance systems thanks to the use of advanced computer vision algorithms. This capacity improves the overall effectiveness of surveillance activities in identifying and mitigating security threats by facilitating quick decision-making and intervention.

There are many advantages to video surveillance systems that include cutting-edge computer vision technologies. In the first place, it makes it possible to identify suspicious behavior in real time and to promptly notify and warn security staff to possible security threats so they may take immediate action. Furthermore, by reducing the need for human operators, automated surveillance systems lower operating costs and lower the possibility of human error. Furthermore, video footage is analyzed by computer vision algorithms with unmatched accuracy, thereby lowering false alarms and improving total threat detection accuracy.

In conclusion, the incorporation of cutting-edge computer vision technologies, which provide unmatched potential for improving public safety and security, signifies a paradigm shift in the field of video surveillance. These tools enable authorities to react proactively to security risks in real-time by automating the processing of video data and identifying

certain behavioral indications. The potential for enhancing the capabilities of video surveillance systems and ultimately fostering the development of safer and more secure societies is enormous if computer vision and machine learning continue to progress.

II. LITERATURE SURVEY

Video surveillance is a powerful tool for detecting and preventing violent incidents, and recent advances in machine learning have led to the development of sophisticated algorithms for violence detection in surveillance videos. In this literature survey, we review some of the recent research on video surveillance for violence detection.

"Violence detection using pre-trained models" by Narges Honarjoo et al. In their research, they propose a low-complexity method for quickly detecting violence. The approach involves using a pre-trained model to extract features from multiple frames simultaneously. Feature vectors are obtained from all frames in a video and combined using a Pot function to generate a unique feature vector that represents the input video. A fully connected network is then trained for classification. The effectiveness and generality of the proposed method, particularly for real-time applications, are demonstrated through experiments conducted on four widely used datasets [1].

The research conducted by Romas Vijeikis et al in "Efficient Violence Detection in Surveillance" proposes a novel and efficient approach to detect violent actions in real-world surveillance videos. The model presented in the study employs MobileNet V2 as an encoder, which is followed by LSTM for temporal feature extraction and classification. The model architecture resembles a U-Net-like network that extracts spatial features. With a total of 4,074,435 parameters, the model is lightweight and computationally efficient. The cross-validation of the model was carried out using three different datasets, namely Hockey Fights, Movie Fights, and RWF-2000, with five folds [2].

The proposed model, despite being small and computationally cheap, obtained high accuracy.

FATH U MIN ULLAH et al. conducted a literature survey on the topic of "A Comprehensive Review on Vision-based Violence Detection in Surveillance Videos". The survey provides a comprehensive overview of the concept, applications, challenges, and working flow of VD. The authors also classify the neural networks utilized in VD and explore the datasets utilized for both surveillance and non-surveillance tasks. Additionally, the article includes a section on the various applications of VD and provides suggestions for future research. The survey emphasizes the potential of VD in diverse domains, including smart cities and industrial safety, and highlights the importance of artificial intelligence-based smart techniques in advancing VD research [3].

"Deep Learning for Automatic Violence Detection: Tests on the AIRTLab Dataset" by PAOLO SERNANI et al. In their research paper, they propose and evaluate three deep learning-based models for automatic violence detection in videos across three datasets. The study shows that the transfer learning-based models exhibit consistent accuracy results across all three datasets, outperforming related works on Hockey Fight and Crowd Violence datasets, demonstrating their potential for violence detection. Additionally, the 3D CNN-based models outperform the 2D CNNs pre-trained on ImageNet, indicating the need for further research on 3D architectures for violence detection. The study underscores the potential of deep learning-based models for violence detection in videos and offers insights into the effectiveness of different types of CNN architectures for this task [4].

III. METHODOLOGY

The methodology for violence detection in video surveillance comprises a series of pivotal stages:

1. **Accepting Videos:** This initial step involves users providing video data to the system, forming the foundation for subsequent analysis.
2. **Segmentation:** Videos are meticulously divided into frames, a process crucial for granular processing and analysis.
3. **Preprocessing:** Prior to feature extraction, frames undergo quality enhancement techniques, ensuring optimal data integrity and accuracy.
4. **Feature Extraction:** Leveraging the power of pre-trained VGG-16 CNN, the system meticulously captures object appearance and motion features from the video frames. This step is pivotal for extracting meaningful information crucial for subsequent analysis.
5. **Classification:** The heart of the process lies in the employment of LSTM networks for sequential data analysis. By scrutinizing the extracted features, the system discerns patterns indicative of violent activity, thereby facilitating accurate classification.
6. **Violence Detection:** Building upon the classification results, the system meticulously identifies instances of violence, a critical task that ensures timely intervention and response.
7. **Output Generation:** Ultimately, the system furnishes comprehensive findings pertaining to detected violent activity, thereby enabling stakeholders to take informed action.

By synergistically employing deep learning techniques such as pre-trained CNNs for feature extraction and LSTM networks for classification, this methodology offers a robust framework for identifying and addressing violent behavior within video streams effectively.

1) Classifiers: Classifiers play a crucial role in violence detection within video surveillance systems. They analyze visual data, distinguishing between violent and non-violent actions by extracting features and capturing temporal dependencies. Here are some key classifiers:

1. **CNN (Convolutional Neural Network):**
   - Highly effective in image classification.
   - Automatically learns hierarchical representations of visual data.
   - Excels at extracting features like shapes and textures associated with violence.
   - Leveraging pre-trained models like VGG-16 enhances performance.
   - Transfer learning fine-tunes models for specific datasets, such as hockey footage.

2. **VGG-16 (Visual Geometry Group 16):**

- Widely used for image classification.
- Offers powerful feature extraction, capturing spatial information from video frames.
- Automatically learns discriminative features related to violent behavior.
- Integrates well with other components like LSTM networks for comprehensive analysis.

3. **LSTM (Long Short-Term Memory):**

- Specialized in modeling temporal dependencies in sequential data.
- Captures long-range dependencies and recognizes patterns over time.
- Analyzes sequential nature of video frames, detecting subtle temporal cues of violence.
- Integration with CNN classifiers enhance violence detection accuracy.
- Attention mechanisms can further improve accuracy by focusing on relevant parts of the video sequence.

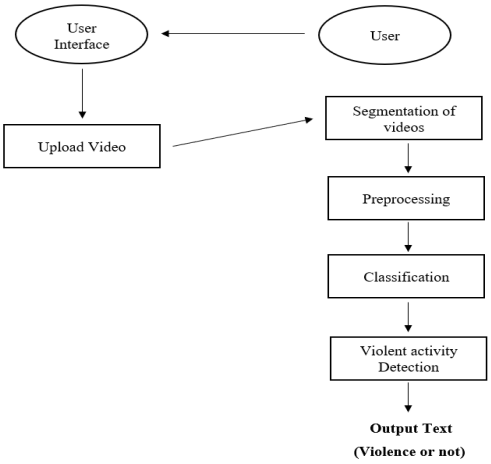


Figure 1. Working Flow

2) DATASET:

The Hockey dataset is a valuable resource for training and evaluating violence detection algorithms in video surveillance systems. With 500 violence and 500 non-violence videos from hockey games, it offers diverse scenarios relevant to the project's objectives. Hockey matches, known for physical contact and aggressive behavior, provide an ideal testbed. The dataset includes instances of both violent incidents and non-violent gameplay, providing a balanced representation. Each video captures dynamic scenes with varying crowd density, camera angles, and motion patterns, resembling real surveillance footage challenges. Training on this dataset enhances the system's ability to generalize understanding of violent behavior across different conditions and perspectives, improving robustness. The dataset serves as a crucial benchmark for evaluating violence detection algorithms, aiming to enhance public safety through accurate detection and response to violent incidents in various settings.



Figure 2. Samples of Hockey Fight Dataset

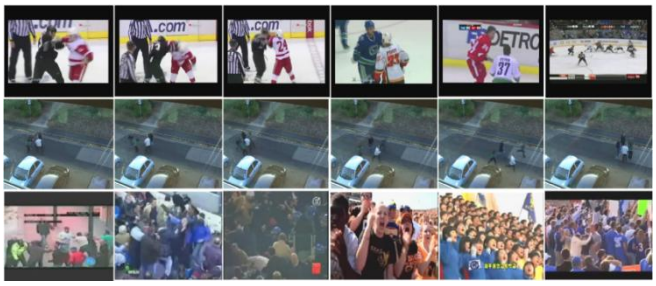
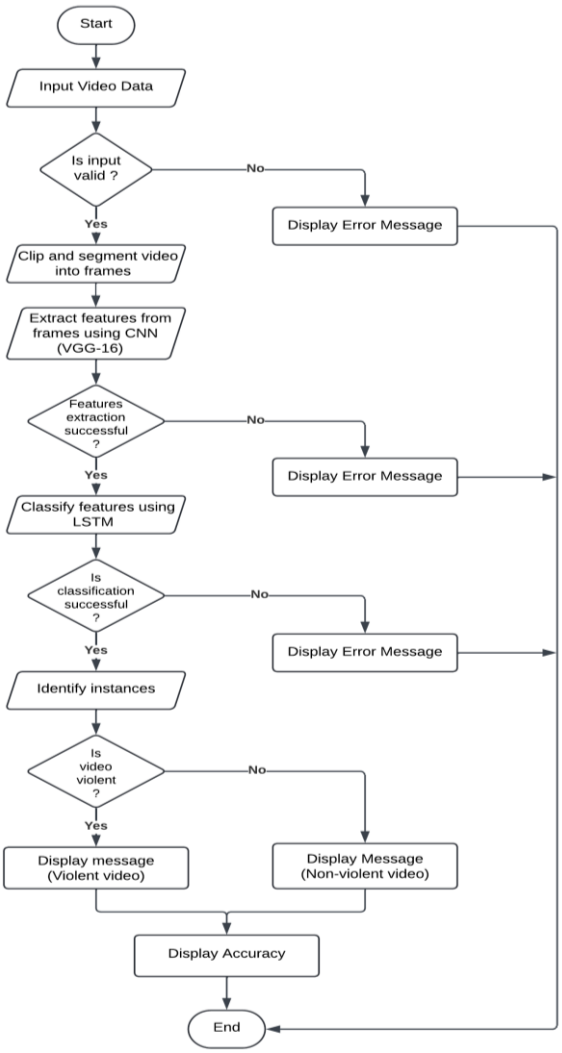


Figure 3. Frames of Dataset

3) FLOWCHART:



IV. Related Work

Violent actions encompass various behaviors such as punching, kicking, pushing, hugging, and gunshot incidents. While researchers historically utilized both visual and audio features for brutality detection, recent focus has shifted primarily to visual features due to the limited quality of audio features in surveillance videos. Recognizing that events within videos unfold over time and space, integrating temporal and spatial features extraction can enhance detection accuracy. Visual-based violence detection employs two primary feature extraction approaches: handcrafted features and deep learning-based features.

Handcrafted features involve obtaining spatial and temporal features using fundamental image processing techniques. Methods like Motion Scale-Invariant Feature Transform (MoSIFT) identify interest points and encode them spatially to extract motion features. Additionally, approaches such as the Bag-of-Words (BOW) framework utilize descriptors like SIFT and MoSIFT to remove redundant information. Histogram of Gradients (HOG) and Histogram of Optical Flow (HOF) are popular handcrafted approaches adept at representing spatial and temporal features. Novel descriptors like Histogram of Optical flow Magnitude and Orientation (HOMO) have been proposed, capturing optical flow between consecutive grayscale frames.

Deep learning-based approaches for feature extraction in violence detection have gained traction in recent years. Utilizing Convolutional Neural Networks (CNNs) such as VGG16 enables spatial feature extraction, while differential frames feed temporal features into the network. Methods like the Se-Layer introduce networks to weigh features extracted from convolutional layers. Other approaches incorporate Hough forests with 2D CNNs or fine-tuned MobileNet models for violence classification.

Temporal features are crucial in video analysis, prompting the use of 3D CNNs fed with multiple frames to extract temporal features concurrently with spatial ones. Techniques involve specialized network layers for better temporal pooling and the utilization of optical flow alongside appearance features. Additionally, lightweight CNNs for human detection and 3D CNNs for spatio-temporal feature extraction are employed. End-to-end 3D CNNs offer an integrated approach for spatio-temporal feature extraction and classification.

Long Short-Term Memory (LSTM) networks are widely used for capturing temporal changes across consecutive frames. Some methods employ ConvLSTM, a combination of LSTM and CNN, for improved temporal information extraction and faster inference. Bidirectional LSTM (BiLSTM) networks are utilized to enhance temporal information accuracy. Techniques such as retraining pre-trained network models like VGG16 with specialized datasets and employing networks like BIGRU further enhance violence detection accuracy.

V. Proposed Solution

The proposed solution aims to address the challenge of violence detection in videos using deep learning techniques, specifically leveraging transfer learning with a pre-trained VGG16 convolutional neural network (CNN) and Long Short-Term Memory (LSTM) recurrent neural network (RNN). By combining these two powerful architectures, the system can effectively extract spatial features from individual frames and capture temporal dependencies within video sequences, enabling robust and accurate violence detection.

The VGG16 CNN is employed as a feature extractor to capture high-level visual representations from video frames. Leveraging pre-trained weights learned from the ImageNet dataset, the VGG16 model can effectively encode complex visual patterns, such as textures, shapes, and object compositions. Specifically, the last fully connected layer (fc2) of the VGG16 model is utilized to extract feature vectors from each frame of input videos, providing a rich representation of the visual content.

Following feature extraction, the extracted feature vectors are fed into an LSTM network for sequence modeling. The LSTM network can capture temporal dependencies and long-range patterns within sequences of feature vectors, allowing it to analyze the dynamic evolution of visual content over time. By processing the sequence of feature vectors, the LSTM network learns to identify temporal patterns indicative of violent behavior, such as sudden movements, aggressive gestures, or confrontational interactions.

The proposed model is trained using a dataset of labeled videos containing examples of both violent and non-violent behavior. During training, the model learns to classify videos based on their content, optimizing classification accuracy through iterative backpropagation and gradient descent. The model's performance is evaluated on a separate test set to assess its generalization ability and effectiveness in real-world scenarios.

Experimental results demonstrate the effectiveness of the proposed approach in violence detection. The model achieves high accuracy in classifying videos as violent or non-violent, outperforming traditional methods and showcasing the benefits of deep learning-based techniques. Performance metrics such as precision, recall, and F1-score provide insights into the model's ability to correctly identify instances of violence while minimizing false positives.

VI. Conclusion

In conclusion, this project has successfully delivered a forward-looking solution for the detection of violent incidents in public spaces. Leveraging state-of-the-art technologies like deep learning, computer vision, and real-time data processing, the system has emerged as a pivotal tool for enhancing public

safety. Its capacity to swiftly identify and alert relevant authorities to violent activities has the potential to significantly reduce harm and bolster overall security. What sets this system apart is its versatility, with applicability spanning diverse domains, from streets and transportation hubs to entertainment venues. It's underpinned by a strong commitment to privacy and ethics, striking a delicate balance between the imperative of public safety and individual privacy.

Efficiency is at the core of its design, with the system adeptly handling large volumes of video data in simulated real-time. Early detection and timely messages form the system's cornerstone, enabling proactive intervention in potentially hazardous situations. Furthermore, its commitment to continuous improvement ensures that it remains a reliable, up-to-date tool adaptable to evolving security needs. Ultimately, this project's impact is felt at the societal level, making public spaces safer and contributing to the well-being of communities. With the potential to serve as a deterrent to potential offenders, this violence detection system stands as a testament to the positive outcomes achievable through the fusion of state-of-the-art technology, ethical considerations, and efficiency in the arena of security and surveillance.

VII. ACKNOWLEDGMENT

This Project wouldn't have been possible without the support, assistance, and guidance of a number of people whom we would like to express our gratitude to. First, we would like to convey our gratitude and regards to our mentor Prof. Pankaj Vanwari for guiding us with his constructive and valuable feedback and for his time and efforts. It was a great privilege to work and study under his guidance. We would like to extend our heartfelt thanks to our Head of Department, Dr. Sachin Bojewar for overseeing this initiative which will in turn provide every Vidyalankar student a distinctive competitive edge over others. We appreciate everyone who spared time from their busy schedules and participated in the survey. Lastly, we are extremely grateful to all those who have contributed and shared their useful insights throughout the entire process and helped us acquire the right direction during this research project.

REFERENCES

[1] Narges Honarjoo, Ali Abdari, Azadeh Mansouri " Violence detection using pre-trained models" 2021 5th International Conference on Pattern Recognition and Image Analysis (IPRIA) Department of Electrical and Computer Engineering, University of Kashan, Iran, 28 & 29 April 2021.

[2] Romas Vijeikis , Vidas Raudonis and Gintaras Dervinis "Efficient Violence Detection in Surveillance" Department of Automation, Faculty of Electrical and Electronic Engineering, Kaunas University of Technology, 51367 Kaunas, Lithuania.

[3] Fath U Min Ullah, Mohammad S. Obaidat, Amin Ullah, Khan Muhammad, Mohammad Hijji Sung, Wook Baik, "A Comprehensive Review on Vision-based Violence Detection in Surveillance Videos" Association for computery machinery, NY, United States.

[4] Paolo Sernani, Nicola Falcionelli , Selene Tomassini, Paolo Contardo 1,2, Aldo Franco Dragoni , "Deep Learning for Automatic Violence Detection: Tests on the AIRTLab Dataset" Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, 60131 Ancona, Italy.

[5] Huiling Yao, Xing Hu, "A survey of video violence detection" School of Optical-Electrical Information and Computer Engineering, University of Shanghai for Science and Technology, Shanghai, China.

[6] Fath U Min Ullah, Amin Ullah, Khan Muhammad, Ijaz Ul Haq, Sung Wook Baik, "Violence Detection Using Spatiotemporal Features with 3D Convolutional Neural Network" Department of Software, Sejong University, Seoul 143-747, Korea.

[7] S. Sudhakaran and O. Lanz, "Learning to detect violent videos using convolutional long short-term memory", 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1-6, 2017.

[8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", *ICLR*, 2015.

[9] https://academictorrents.com/details/70e0794e2292fc051a13f05ea6f5b6c16f3d3635

❑ ❑ ❑ ❑ ❑ ❑

AET 006

CropNexus: Your All-in-One Crop Management Solution

Avinash Shrivas
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
avinash.shrivas@vit.edu.in

Utkarsh Phatak
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
utkarshphatak888@gmail.com

Prathmesh Yedage
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
prathmeshyedage@gmail.com

Rahul Telang
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
rahultelang12@gmail.com

***Abstract*- This paper focuses on CropNexus is an innovative agricultural decision support system that integrates advanced features including crop recommendations, crop recommendations, crop disease forecasting, and real-time weather data via an API, with a Gemini chatbot interface driven by devices uses learning algorithms based on CropNexus products to supplement Provides customized recommendations and nutrients for strategies incorporated into early crop disease detection and prediction Uses deep learning models. Integration of real-time weather data facilitates informed decision- making, while Gemini's chatbot provides user-friendly interfaces for seamless communication CropNexus represents a paradigm shift in agricultural technology, which its purpose is to increase productivity, sustainability and resilience in agricultural practices.**

I. INTRODUCTION

In an ever-evolving agricultural environment, the integration of technology has become essential in meetingthe challenges of global population growth, climate change and food security, traditional agricultural practices are being made largely and is being transformed by innovative solutions aimed at improving productivity, sustainability and resilience . These solutions include CropNexus, an integrated decision support system that represents a paradigm shift in agricultural management.

CropNexus is not just a software platform; It's a comprehensive set of tools designed to empower farmers with actionable insights, personalized recommendations, and easy communication. At its core, there are five main features: crop information, nutrient recommendations, crop

disease forecasting, integration of weather APIs, Gemini chatbot interface and every carefully designed product to consume crops addressing specific management aspects, using state-of-the-art technologies such as machine learning, deep user learning and natural language processing

The importance of CropNexus lies in its ability to streamline decision-making processes, improve resource utilization, and reduce risk in agricultural production While providing powerful data analytics, predictive modeling and real-time information in use, CropNexus equips farmers with the tools they need to navigate the challenges of modern agriculture

Crop Suggestion: One of the major challenges in agriculture is selecting the right crops for cultivation based on factors such as soil type, climate, market needs, crop rotation needs etc. CropNexus tackles this challenge head on so through its crop suggestion feature. Analyzing big data including soil characteristics, climate models, historical crop performance and market trends, CropNexus uses machine learning algorithms to generate individual crop recommendations tailored to each field's conditions the special These recommendations not only improve yield potential but also contribute to diversification and risk management.

Fertilizer Recommendation: Effective nutrient management is essential to maximize crop yields and reduce environmental impact. Traditional fertilizer application methods are often inaccurate and result in poor product utilization. CropNexus is revolutionizing fertilizer use with its fertilizer recommendations. Integrating soil nutrient analysis, crop nutrient requirements, and environmental factors, CropNexus develops tailored nutrient recommendations for each field Through predictive modeling and data- driven insights, CropNexus ensures that crop management is targeted, efficient, and environmentally friendly.

Crop Disease Prediction: Crop diseases pose a serious threat to agricultural crops, often leading to crop losses and economic hardship for farmers. Early detection and aggressive management of crop diseases is essential to minimize their impact. CropNexus uses advanced machine learning models to analyze crop health indicators and predict potential disease outbreaks. By monitoring factors such as weather, pest populations and crop health data, CropNexus provides farmers with timely monitoring alerts and recommendations, allowing them to implement preventive measures manage and reduce losses Integration with Weather APIs. Climate variability is a determinant of agricultural production, affecting planting decisions, irrigation systems, insect outbreaks and harvest time. Real-time weather information is essential for informed decision-making in the agricultural industry. It integrates with the CropNexus weather API and provides farmers with up-to-date weather forecasts, historical weather data and weather alerts. By using this information, farmers can respond to a changing climate by adapting their practices, optimizing resource allocation and reducing climate-related risks

Gemini Chatbot Interface: User interface design plays an important role in the adoption and use of agricultural technologies. CropNexus has a Gemini chatbot interface powered by natural language processing and machine learning algorithms.This intuitive interface provides farmers with a user- friendly way to interact with the system, receive information and personalized recommendations. With a seamless interface and intuitive design, Gemini's chatbot enhances user experience and makes it easier to integrate CropNexus into daily farming routines

II. RELATED WORK

Agriculture is rapidly adopting AI technology to solve long-term challenges and improve farming practices. CropNexus builds on existing research in various AI subfields, each contributing its core functionality. Here is a detailed analysis of the relevant research areas:

Crop yield forecasting: Traditionally, crop yield forecasting was based on mathematical models and expert knowledge. However, the explosion of agricultural data has paved the way for machine learning algorithms to play an important role. The study by Yang et al. (2021) demonstrate the efficiency of using a short-term memory (LSTM) network, a type of recurrent neural network (RNN), to estimate yields. LSTMs can capture the time dependence in historical yield data, climate models, and soil conditions, resulting in more accurate forecasts compared to traditional methods

Diagnosis of Crop Diseases: Early and accurate diagnosis of crop diseases is essential to reduce yield losses and ensure food safety. Deep learning, in particular convolutional neural networks (CNNs), have emerged as powerful tools for this purpose. Mohanty et al. (2016) investigated the use of CNN for image-based plant disease diagnosis. Their study achieved high accuracy in detecting different diseases in different crops, highlighting the potential of deep learning for automated disease detection in the field

Soil analysis: An understanding of soil health and nutrient levels is essential for informed application of fertilizers and sustainable land management practices. Mouazen et al. (2019) investigated the use of machine learning for soil analysis using data collected from nearby sensors. Their research shows that machine learning algorithms can effectively analyze sensor data to assess soil properties such as organic matter content, moisture content and nutrient availability and this information can be used to create customized feeding recommendations, for products a efficiency and environmental reduction.

Integration of climatic data: Climate plays an important role in crop growth and development. With access to real-time and hyperlocal weather data, farmers can make informed decisions about irrigation schedules, pest control, and other agricultural operations investigated the integration of climate APIs into agricultural decision support systems. The findings demonstrate the value of incorporating climate data into agricultural models, thereby improving resource allocation and risk management strategies

Comparison to Existing Solutions: Several existing solutions offer functionalities similar to CropNexus. However, CropNexus distinguishes itself by offering a comprehensive platform that integrates various AI- powered tools under a user-friendly interface. Here's a comparison with some notable examples:

Plantix: This mobile app utilizes image recognition to diagnose crop diseases. While Plantix excels in disease identification, it lacks functionalities like crop suggestion, fertilizer recommendation, and weather integration.

Detecte: This cloud-based platform leverages machine learning for yield prediction and risk management. However, Detecte requires significant user input and expertise in data analysis, making it less accessible to farmers with limited technical knowledge.

FarmOS: This open-source farm management software offers features like record keeping, resource management, and task scheduling. While FarmOS provides valuable tools for farm organization, it lacks the AI-powered decision support capabilities offered by CropNexus

Addressing limitations of existing research

While the aforementioned research areas provide a strong foundation for CropNexus, there are some limitations that CropNexus aims to address: Limited availability of data: One of the main challenges in applying AI in agriculture is the lack of high-quality agricultural data labels. CropNexus aims to address this through data collection strategies that leverage sensor networks, farmer data and partnerships with agricultural research institutes.

Generalization of AI models: AI models trained on data from specific fields may not perform well when applied to different environments with different soil, climate and crop rotations Cropnexus Transfer learning Domain optimization to improve the generalizability of its AI models to diverse agricultural settings It addresses this using techniques such as f

Usability and accessibility: Complex AI models often require technical expertise to define and implement them. CropNexus addresses this challenge with a user-friendly chatbot interface, Gemini, that translates complex data and recommendations into usable insights for farmers with varying levels of technical experience on the combination.

By addressing these limitations, CropNexus seeks to develop a practical and accessible AI-powered system that empowers farmers to make data-driven decisions and achieve optimal agricultural yields

Crop Recommendations: This feature uses historical data on crop yield, climate and soil conditions. Machine learning algorithms analyze this information to recommend crops suitable for a particular area, maximize yield potential and optimize land use.

Crop recommendation: By analyzing soil nutrient levels and crop needs, CropNexus recommends optimal fertilizer rates. This enables the use of machine learning models trained on soil analysis data to promote better resource use and better crop growth.

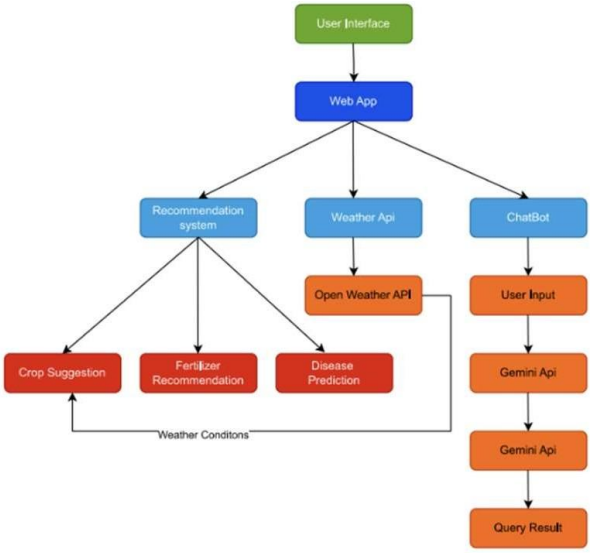
Crop Disease Prediction: Early detection of diseases is essential to minimize crop losses. CropNexus uses a deep learning algorithm to analyze visual data (e.g., images uploaded by farmers) and sensor readings from fields. This

allows for early detection and prevention of crop diseases, improving crop health and overall yields.

Climate API Integration: Access to real-time and hyperlocal weather forecasts is critical for informed irrigation and pest management decisions. Together with the climate API, CropNexus provides farmers with this critical information, allowing them to better allocate production and mitigate climate-related risks

III. Design

CropNexus brings together several key services to providecomprehensive support to farmers:



Gemini Chatbot: This user-friendly interface is a focal point of interaction with CropNexus. Through text and voice commands, farmers use Gemini to receive personalized recommendations, access information from available knowledge, monitor their farm data, and receive alerts or information about potential problems (e.g. diseases). outbreak, bad weather). You can contact the astrology organizationGemini Chatbot: This user-friendly interface is a focal point of interaction with CropNexus. Through text and voice commands, farmers use Gemini to receive personalized recommendations, access information from available knowledge, monitor their farm data, and receive alerts or information about potential problems (e.g. diseases). outbreak, bad weather). You can contact the astrology organization

CropNexus follows a client-server architecture with a modular design for each major application. This modularity allows for scalability of additional features and future integration.

Client-side application: Accessible through a mobile application or web platform, this application provides a user- friendly interface for interaction with CropNexus. Users primarily interact with Gemini, the chatbot, in order to get recommendations, manage their data, and access functionality.

Server-side infrastructure: This infrastructure consists of modules: Data Management Module: Manages data collection, storage and organization. Integrates with various data sources, including historical agricultural data, real-time sensor data, user input, weather APIs, and potentially marketable data APIs (if combined); among.

AI Processing Module: This module contains AI algorithms for each of the core applications. Each algorithm is specifically trained on relevant data sets to generate accurate and optimized recommendations.

Knowledge: This repository collects information on various aspects of agriculture, including crops, diseases, fertilizers, best practices, and market trends (if market fluctuations are included) Gemini uses this knowledge to answer farmers' questions and provide references for recommendations.

API Gateway: Responsible for ensuring secure and efficient communication between client-side applications and server-side modules.

IV. Implementation

CropNexus development requires careful consideration of various technological factors. This section explores the operating system of each of the key functions in more detail

Historical Agricultural Information: Collaborate with government agencies and agricultural research institutes to obtain historical data on crop yields, weather conditions and local soil conditions.

Look for publicly available datasets from organizations such as the Food and Agriculture Organization (FAO) or national agriculture departments.

Partner with existing field management processes to enhance historical data collection (if anonymous data sharing agreements can be reached).

Real-time sensor data:

Use sensors in certain locations to collect real-time data: soil water content Soil temperature Nutrient levels (e.g., nitrogen, phosphorus, potassium). Ambient temperature and humidity (for prediction of disease).

Partner with sensor manufacturers or use open source sensor platforms like Arduino or Raspberry Pi for lower cost.

Ensure data protection measures are established for sensor communication and data transmission.

User Input:

Design a user-friendly interface in a mobile app or web platform to collect user feedback on the following: Location of the farm (GPS coordinates or zip code).

Soil types (based on soil test results or regional classification).

Previous planting practices (varieties, planting dates). Current crop production

Weather synthesis: Integrate with popular weather APIs to get real-time and hyperlocal weather forecast for the user's specific location. Look for free or subscription-based weather APIs based on project requirements and data needs. When choosing a weather API, consider factors such as data accuracy, frequency of updates, and area coverage AI model development each of CropNexus' core functionalities uses specific AI algorithms trained on relevant data sets. Here is a detailed breakdown of the model development process.

Vegetable Recommendations and Vegetable Recommendations: Machine learning models should be developed using algorithms such as Support Vector Machines (SVMs) or Random Forests. These algorithms excel in pattern recognition and classification tasks.

Traditional models of historical agricultural data (crop yield, climate, soil conditions) in conjunction with the user (farm location, soil types).
For examples of recommendations, include information on fertilizer application rates and their effects on crop yields. Continuously monitor and improve model performance by incorporating new data and user input.

Crop Disease Forecast: Develop deep learning models using Convolutional Neural Networks (CNNs). CNNs are particularly good at image recognition and classification.

Create a large data set of labeled crop images. This can be done through collaboration with agricultural universities or research institutes, or through publicly accessible data sets. Label the specific diseases represented in the photographs (e.g., rust, rot, mold).

Trains a CNN image on a labeled image dataset to detect diseases in user-uploaded images or real-time sensor data

Data Preprocessing and Processing: Data cleansing techniques were used to address missing values, outliers, and inconsistencies in the data.

Normalize the data to ensure that all features have the same scale, which improves model training efficiency.

Use feature engineering techniques to create new features from existing data that can be very informative for AI models.

Chatbot Development (Gemini)

Natural Language Processing (NLP): Use frameworks like NLP libraries like TensorFlow Lite or

spaCy to train language models that understand user questions and answer them naturally and informatively.

The ancient language model has a data set of conversational models for the field. These data structures can include examples of user questions, relevant information from agricultural experts, and information from CropNexus knowledge base.

Discussion Program:

Create a conversation template to monitor conversations between Gemini and the user. This framework determinesthe intent behind the questions used, captures context from the knowledge or AI model, and provides customized answers.

Use a combination of rule-based methods and machine learning techniques to handle dialogue. Rule-based methods can handle the user's overall intent, while machine learning can be used for complex or open-ended queries system integration and management to develop APIs to facilitate communication between components of CropNexus: data collection module AI application modules Knowledge base Weather API Using Gemini Chatbot

VI. FUTURE SCOPE

CropNexus takes a significant step forward in AI-powered agriculture. However, ongoing research and development can further shape its capabilities and expand its impact

1. Integrating IoT devices: Expanding sensor integration beyond soil sensors to include devices such as drones and precision agricultural machinery can also provide more granular data on plant health, the environment around conditions and resource management. This can further increase the accuracy of CropNexus' recommendations.
2. Advanced AI algorithms: Exploring more advanced AI algorithms, such as recurrent neural networks (RNNs) for time series analysis or reinforcement learning for optimizing decision-making processes, can lead to dynamic recommendations and they can be changed
3. Economic modeling: The integration of economic modeling into CropNexus enables consideration of factors such as market prices and investment costs along with crop choice and input recommendations. This gives farmers a comprehensive view of the profitability of their farms.
4. Blockchain Technology: Blockchain technology can be explored for secure and transparent data processing in the agricultural supply chain. This can build trust and cooperation between farmers, retailers and other stakeholders.
5. Accurate Irrigation: CropNexus can be used to create precise irrigation systems that adjust the water supply based on real-time sensor data. This can greatly improve water efficiency and reduce water waste.
6. Predictive maintenance: By analyzing sensor data on field equipment, CropNexus can anticipate potential maintenance issues and propose preventative measures. This can reduce downtime and improve machine efficiency.
7. Tailored learning for farmers: Incorporating machine learning to analyze user behavior could enable CropNexus to tailor recommendations and educational content to individual farmers' needs and learning styles.
8. Multilingual support: Expanding language support for Gemini, the chatbot interface, could make CropNexus more accessible to farmers worldwide, removing language barriers to adoption of AI-powered farming practices
9. Robotic automation: The integration of agricultural robots for tasks such as planting, weeding and harvesting could be explored in the long term. CropNexus can provide real-time guidance and control to these robots, optimizing their performance and accuracy.
10. Greenhouse management: Optimizing CropNexus for controlled-environment agriculture (CEA) systems in greenhouses can improve resource use and maximize yields in this intensive agricultural industry.
11. Sustainability assessment: CropNexus could be enhanced to include environmental assessment in its recommendations. This can help farmers adoptpractices that reduce their environmental footprint and promote sustainable agriculture.

12. Policy and regulatory considerations: As AI in agriculture evolves, research into policy and regulatory frameworks is needed. This ensures responsible development and use of AI tools such as CropNexus, addressing potential concerns about data privacy, algorithmic bias and ethical considerations
13. Farmer feedback and co-design: Continued input of farmers through surveys, focus groups and participatory policy workshops can guide the future development of CropNexus. This ensures that the platform remains relevant and meets the evolvingneeds of farmers.

VII. CONCLUSION

Agriculture faces unprecedented challenges in the face of climate change, resource scarcity and rising demand for food. Traditional agriculture struggles to adapt to these harsh realities, requiring innovative solutions. CropNexus emerges as the answer, offering a new farm management platform that uses artificial intelligence (AI) to empower farmers with data-driven decision-making power.

This research paper examined CropNexus design, implementation, and potential impact. It examined relevant research on AI-powered agriculture, and highlighted the need for such solutions. In addition, the paper discussed the system design, technical considerations, and operational issues related to the development and implementation of CropNexus. Finally,the paper explored potential benefits for farmers, sustainable agriculture, and future research directions.

CropNexus has the potential to transform farming practices by providing farmers with advanced AI-powered tools. Combining crop data, crop recommendations, crop disease predictions, weather API integration, and an easy-to-use chatbot interface (Gemini), CropNexus empowers farmers to

- Use historical data and real-world environmental conditions to make optimal crop choices and land use.
- Reduce waste by recommending fertilizers for individual crops.
- Proactively combat crop diseases through early detection and intervention strategies.
- Make appropriate decisions about irrigation and other field activities based on real-time weather data.
- Valuable agricultural knowledge can be gainedthrough the delicate Gemini connection.

The potential benefits of CropNexus also extend to individual farmers. CropNexus can help by promoting sustainableagricultural practices:

Food production has increased to meet the demands of the growing world population.

Waste is reduced by reducing the environmental impact through the use of improved materials.

Improve agricultural productivity by enabling farmers to make data-driven decisions that improve productivity and operational efficiency.

REFERENCES

[1] Z. Saeed, A. Raza, A. H. Qureshi and M. Haroon Yousaf, "A Multi-Crop Disease Detection and Classification Approach using CNN," 2021 International Conference on Robotics and Automation in Industry (ICRAI), Rawalpindi,Pakistan, 2021, pp. 1-6, doi:10.1109/ICRAI54018.2021.9651409

[2] Z. Yu, "Research on Crop Pest and Disease Diagnosis Method Based on Improved Faster R-CNN," 2023 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China, 2023, pp. 46-50, doi: 10.1109/ICAICA58456.2023.10405454.

❑ ❑ ❑ ❑ ❑ ❑

AET 007

Sign Language Recognition System for Deaf and Mute People

Siddhesh Deshmukh
Department of Computer Engineering
Vidyalankar Institute of Technology, Mumbai, India
siddhesh.deshmukh@vit.edu.in

Shruti Khandagale
Department of Computer Engineering
Vidyalankar Institute of Technology, Mumbai, India
shruti.khandagale@vit.edu.in

Sanjeev Diwedi
Department of Computer Engineering
Vidyalankar Institute of Technology, Mumbai, India
sanjeev.diwedi@vit.edu.in

Abstract—This paper presents a novel Sign Language Detection System tailored for individuals with hearing and speech impairments, commonly referred to as deaf and mute. Leveraging advancements in computer vision and deep learning techniques, the proposed system offers real-time interpretation of sign language gestures, facilitating seamless communication between the deaf-mute community and individuals unfamiliar with sign language. By employing techniques such as landmark detection and temporal sequence analysis, the system accurately recognizes and translates hand gestures into textual representations. The system's effectiveness is demonstrated through extensive testing and validation, showcasing its potential to enhance accessibility and inclusivity in various communication scenarios. This research represents a significant step towards bridging the communication gap for individuals with hearing and speech disabilities, ultimately fostering greater societal integration and understanding.

Keywords—Hand-Sign-Recognition, Deep-Learning, Long-Short-Term-Memory (LSTM), Mediapipe

I. INTRODUCTION

In the realm of communication, language serves as the cornerstone of human interaction, facilitating the exchange of thoughts, emotions, and ideas. However, for individuals who are deaf or hard of hearing, traditional spoken languages present barriers to effective communication. Sign language emerges as a profound alternative, offering a visual and gestural medium through which deaf and mute individuals can express themselves and engage with the world around them.

The significance of sign language cannot be overstated. It not only enables deaf and mute individuals to communicate within their own community but also bridges the communication gap between them and hearing worlds. Despite its importance, the accessibility of sign language remains limited in various contexts, hindering the ability of deaf and mute individuals to fully participate in society.

Advancements in technology present promising avenues for overcoming these barriers. Sign language detection systems, leveraging computer vision and machine learning algorithms, offer a potential solution to enhance accessibility for the deaf and mute community. By automatically recognizing and interpreting sign language gestures, these systems can facilitate seamless communication between these individuals and hearing counterparts in diverse settings, including education, healthcare, and everyday interactions.

This research paper aims to explore the development and implementation of sign language recognition systems for the benefit of deaf individuals. Through a comprehensive review of existing literature, analysis of technological methodologies, and evaluation of real-world applications, this paper seeks to elucidate the current state-of-the-art in sign language recognition technology. Furthermore, it aims to identify challenges, opportunities, and future directions in this field, with the ultimate goal of fostering greater inclusivity and empowerment for the deaf and mute community.

II. LITERATURE SURVEY

Sign language recognition research has emerged as a critical domain, aiming to bridge communication barriers and empower deaf and mute individuals through technology.

A. A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: An Indian Scenario [1]

The system produced by the authors recognizes ISL gestures from mobile camera videos with none additional sensors to detect hand regions. Single-handed static and dynamic gestures and double-handed static gestures are identified in this system. The system uses ROI algorithm for feature extraction. After pre-processing the image, the centroid of the binary image is estimated. Using the gesture separation algorithm, the gesture is recognized, static as well as dynamic. The advantage of this system is that it is economical and can be implemented with a mobile camera, making it very user-friendly. But the disadvantage is that it is not efficient under cluttered backgrounds and different illumination conditions.

B. An Efficient Binarized Neural Network for Recognizing Two Hands Indian Sign Language Gestures in Realtime Environment [2]

Considering the challenges of sign language recognition, on targeted embedded platforms, authors have proposed the novel architecture of a binarized neural network with binary values of weights and activations using bitwise operations. The advantage is using this architecture achieves an overall accuracy of 98.8% which is higher than other existing methods while the disadvantage is This system misclassifies some signs of M, N, E because of their similar kind of shapes, and also, the proposed BNN architecture is limited with small no of classes of gestures.

C. Hand Gesture Recognition for Sign Language Using 3DCNN [3]

This study investigates the use of 3DCNN for hand gesture recognition. The approaches that were proposed by the authors were compared with six other state-of-the-art methods from the literature. They showed comparable performance to the two of these methods and outperformed four of them. However, it does not work for a live video feed.

D. Deep Convolutional Neural Networks for Sign Language Recognition [4]

For classifying selfie sign language gestures, the authors proposed a CNN architecture. A less amount of training and validation loss is observed with the proposed CNN architecture, but the database is not available publicly.

E. Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People in [5]:

In this paper author proposed a system to aid communication of deaf and dumb people communication using Indian sign language (ISL) with normal people where hand gestures will be converted into appropriate text message. Main objective is to design an algorithm to convert dynamic gesture to text at real time. Finally after testing is done the system will be implemented on android platform and will be available as an application for smart phone and tablet pc.

F. SignPro-An Application Suite for Deaf and Dumb [6]:

Author presented application that helps the deaf and dumb person to communicate with the rest of the world using sign language. The key feature in this system is the real time gesture to text conversion. The processing steps include: gesture extraction, gesture matching and conversion to speech. Gesture extraction involves use of various image processing techniques such as

III. PROPOSED SYSTEM

The system is a vision-based approach. All the signs are represented with bare hands and so it eliminates the problem of using any artificial devices for interaction.

The proposed system is implemented using the TensorFlow and OpenCV libraries. The TensorFlow library is used for building the LSTM model and training it on the sign recognition dataset. The OpenCV library is used for reading and processing the input frames, as well as displaying the output sign labels. The proposed system has several advantages. Firstly, it eliminates the need for handcrafted feature extraction, as the keypoint features are automatically extracted using the Mediapipe model. Secondly, it can handle large and complex datasets, as the LSTM model is designed for sequence modelling and can capture the temporal dynamics of the hand actions. Overall, the proposed system offers a novel and effective approach for human action recognition as shown in figure 1.

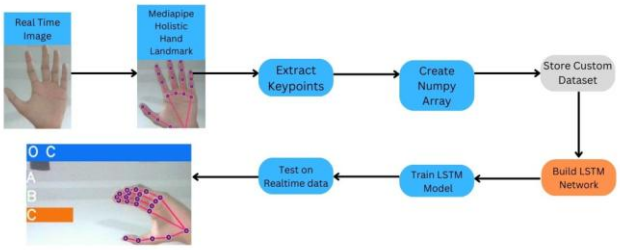


Fig. 1 Methodology

1.Dataset Generation

To create a sign language recognition system, the initial step is to gather a collection of pictures showcasing hand movements. These pictures should have proper lighting, clear visibility of the participant's hand motions. For data acquisition, images are captured by webcam using Python and OpenCV. OpenCV provides functions which are primarily aimed at the real time computer vision. It accelerates the use of machine perception in commercial products and provides a common infrastructure for the computer vision-based applications. The OpenCV library has more than 2500 efficient computer vision and machine learning algorithms which can be used for face detection and recognition, object identification, classification of human actions, racking camera and object movements, extracting 3D object models, and many more The created dataset is made up of signs representing alphabets in American Sign Language as shown in Fig. 1. For every alphabet, 30 images are captured to make the dataset. The images are captured in every 2 seconds providing time to record gesture with a bit of difference every time and a break of five seconds are given between two individual signs, i.e., to change the sign of one alphabet to the sign of a different alpha- bet, five seconds interval is provided.

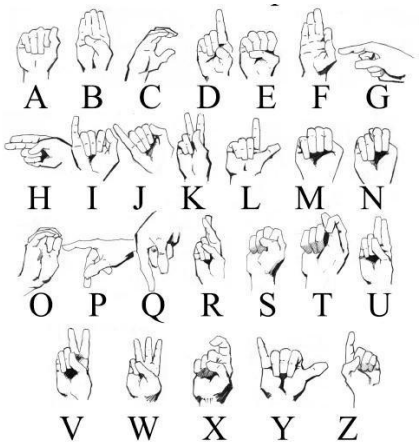


Fig. 2. American Sign Language Alphabe

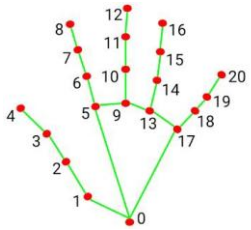
2. Mediapipe

Many machine learning algorithms have been developed for hand gesture recognition so as to create AI-based applications. Out of them, MediaPipe can be used for hand gesture recognition. Google supported MediaPipe framework can be used for solving several problems like face-recognition, face-map, eye, hand, poseestimator, holistic, hair, object-detection, box tracking and KIFT. With the help of the MediaPipe framework, we can develop an algorithm

or model for the application, then help the application by providing results that can be cloned across different platforms. The MediaPipe framework is composed of three major components: (1) performance evaluation, (2) a mechanism for collecting data from the sensor (3) an assembly of reusable parts. A graph consisting of all the parts called the calculators is known as pipeline, wherein every calculator is inter-connected by channels through which the data flows. Developers can create their required application by removing or delineating user defined calculators anywhere in the graph. This result of calculators and channels creates a data-flow diagram.

We started with the MediaPipe Gesture Recognition pre-trained model for our project on sign language recognition. The pre-trained model, however, may not perform well on certain sign language movements that we are interested in because it was trained on a general set of photos. As a result, we used our own proprietary dataset to train the machine on pictures of sign language gestures that we hope to identify. The labeled pictures of sign language motions that we gathered from various sources make up the custom dataset. To make sure the photographs were the same size, aspect ratio, & orientation, we also had to preprocess them.

The proposed system for sign language detection in this project is based on the combination of Mediapipe and LSTM. The Mediapipe framework is used to extract pose and body landmarks from data, while the LSTM model is used to classify the extracted data into different action classes. To extract hand landmarks, the Mediapipe Holistic model is used. Fig 2 shows the hand land marks.



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

Fig. 3 Hand Landmarks

3. Long Short Term Memory(LSTM Model)

The model is initialized as a sequential model. Three LSTM layers are added successively, with 64, 128, and 64 units respectively. Each LSTM layer returns sequences except for the last one. ReLU (Rectified Linear Unit) activation function is used for the LSTM layers. Following the LSTM layers, two dense layers with 64 and 32 units respectively are added, both with ReLU activation. The final dense layer has units equal to the number of actions, with softmax activation, suitable for multi-class classification. The model is compiled using the Adam optimizer, categorical cross-entropy loss function, and categorical accuracy as the evaluation metric.

The model is trained on the training data (**X_train** and **y_train**) for 1000 epochs. TensorBoard callback is passed to visualize training progress. This sets up and trains an LSTM neural network for sign language detection, with TensorBoard for monitoring training.

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 30, 64)	48,896
batch_normalization (BatchNormalization)	(None, 30, 64)	256
dropout (Dropout)	(None, 30, 64)	0
lstm_1 (LSTM)	(None, 30, 128)	98,816
batch_normalization_1 (BatchNormalization)	(None, 30, 128)	512
dropout_1 (Dropout)	(None, 30, 128)	0
lstm_2 (LSTM)	(None, 64)	49,408
batch_normalization_2 (BatchNormalization)	(None, 64)	256
dropout_2 (Dropout)	(None, 64)	0
dense (Dense)	(None, 64)	4,160
dense_1 (Dense)	(None, 32)	2,080

Total params: 204,384 (798.38 KB)
Trainable params: 203,872 (796.38 KB)
Non-trainable params: 512 (2.00 KB)

Fig.4 Model summary

IV. Experimental Results

In this section, we have described the training setup and dataset used for training the sign language recognition model. Additionally, we have presented the results of our experiments, including the classification accuracy for each of the signs. We have also included screenshots of the output produced by the model for sample actions. To visualize the performance of the model during training, we have included graphs showing the loss and accuracy metrics as a function of the number of training epochs. Together, these results provide a comprehensive evaluation of the proposed approach and demonstrate its effectiveness for sign language detection.

A. Training setup and dataset

In this project, we aim to train a sign language recognition model using Mediapipe and LSTM. The dataset consists of 780 sequences in total, with 30 sequences foe each of the 26 letters in alphabet. Each sequence contains 30 frames extracted from realtime action by webcam, which were performed by me and my teammate. We extract the key points from each frame using Mediapipe Holistic, with a minimum detection confidence of 0.5 and a minimum tracking confidence of 0.5. The key points are stored in individual directories for each action, and each sequence is saved as a separate file in its respective directory. The model is trained using the Adam optimizer for 100 epochs, with a loss function of 'categorical_crossentropy' and metric of 'categorical_accuracy'. We used an Intel(R) Core(TM) i7-10870H CPU with 16GB of RAM for the training process. We also use a history object to store the training metrics, which can be visualized using Tensorboard to monitor the training progress.

B. Results

The experimental results of sign language detection model are promising. We achieved a maximum training accuracy of 89% and a minimum training loss of 0.0921 after training the model for 100 epochs. The test accuracy we achieved is 91.02%. We used a 80/20 train-test split to evaluate the model's performance. We experimented with different epoch sizes & found that training the model for 100 epochs produced the best results compared to 35 or 50 epochs.

Figure 4 & figure 5 represents our model accuracy & model loss respectively.

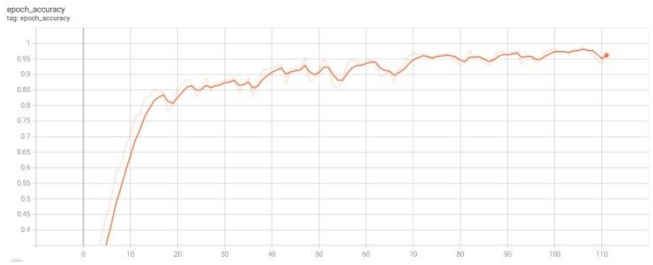


Fig. 5 Model Accuracy

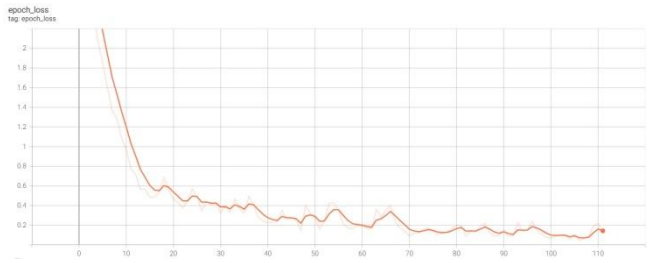


Fig. 6 Model Loss

The action prediction corresponding to each video is shown in Figure 7

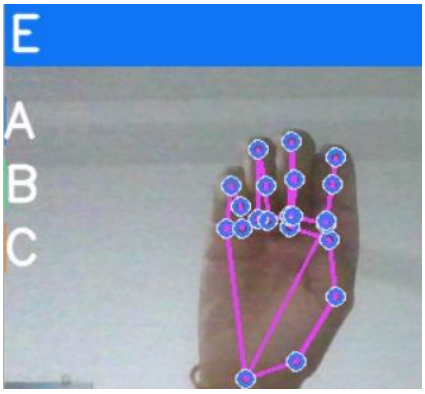


Fig. 7.1 detecting letter E

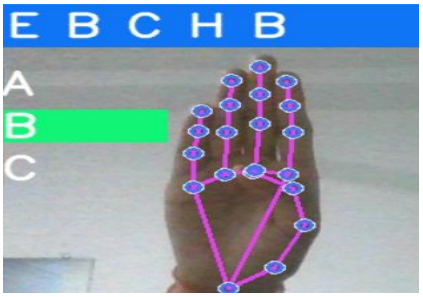


Fig. 7.2 detecting letter B

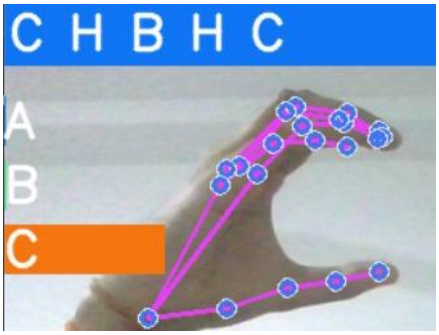


Fig. 7.3 detecting letter C

C. Metrics

Precision: The precision metric is the ratio of True Positives to the sum of False Positives plus True Positives (P =TPTP + FP). We can think of precision as a measure of the confidence of my model. The higher the precision, the more confident we are in the predictions made by the model.

Recall: Similar to precision, this metric is the ratio of True Positives to the sum of False Negatives plus True Positives(R = TPTP+ FN).

F_1 Score: The F_1 Score is the weighted average of my precision and recall metrics, and is calculated with this formula: TPTP + 12(FP + FN). We can think of the F_1 score as a "middle ground" between precision and recall. Precision and recall are specific to the situation: in some situations maximizing the precision over recall is optimal while vice versa in other situations. In general, having a high F_1 score equates to a good model.

	precision	recall	f1-score	support
0	0.88	0.88	0.88	8
1	1.00	1.00	1.00	4
2	1.00	1.00	1.00	6
3	1.00	1.00	1.00	4
4	0.83	0.83	0.83	6
5	1.00	1.00	1.00	6
6	0.67	1.00	0.80	2
7	1.00	1.00	1.00	4
8	1.00	0.43	0.60	7
9	1.00	1.00	1.00	5
10	0.80	1.00	0.89	4
11	1.00	1.00	1.00	8
12	0.83	0.83	0.83	6
13	0.50	1.00	0.67	4
14	0.91	0.91	0.91	11
15	1.00	1.00	1.00	8
16	1.00	1.00	1.00	8
17	0.75	0.75	0.75	4
18	1.00	1.00	1.00	9
19	1.00	1.00	1.00	5
20	0.83	0.83	0.83	6
21	1.00	0.75	0.86	4
22	1.00	1.00	1.00	6
23	0.80	1.00	0.89	8
24	1.00	1.00	1.00	6
25	1.00	0.57	0.73	7
accuracy			0.91	156
macro avg	0.92	0.91	0.90	156
weighted avg	0.93	0.91	0.91	156

Fig. 7 Model metrics

V. Conclusion and Fututre Scope

Individuals with hearing disabilities often face significant challenges in communicating with people who can hear. One of the most effective ways for them to communicate is through sign language. However, for people who do not know sign language, understanding what is being communicated can be a significant challenge. This communication gap can have a detrimental impact on the social and emotional well-being of individuals with hearing disabilities, making it difficult for them to engage fully in society.

In this project, we developed a sign language detection model using Mediapipe and LSTM. We first extracted key points from video frames using Mediapipe Holistic, and then trained an LSTM model. The proposed method's potential for

practical applications is considerable, and it has the capacity to improve the quality of life for individuals with hearing disabilities, helping to bridge the communication gap between them and the rest of the world. Future work will expand the current system to add more indicators and create a complete and reliable system for mobile platforms. Additionally, the proposed method can be adapted for use in other Indian regional languages, such as Hindi, Marathi, Sindhi, Telugu, and more. Although there are still some research gaps that need to be addressed, such as improving the system's accuracy in recognizing signs for complex phrases and developing a portable and affordable device for practical use in daily life, the proposed Sign Language Recognition system offers a promising step towards creating a more inclusive society. With further development and refinement, this system can play a significant role in breaking down communication barriers and facilitating greater accessibility and understanding for individuals with hearing disabilities.

VI. ACKNOWLEDGMENT

The completion of this project is greatly attributed to the invaluable support, assistance, and direction provided by numerous individuals, to whom we are sincerely grateful. Foremost, we would like to express our gratitude to our mentor, Prof. Sanjeev Diwedi, whose constructive feedback and profound insights played a pivotal role in guiding us through this journey. Working under her mentorship was not only an honor but also a privilege that greatly enriched our learning experience. Additionally, we extend our heartfelt appreciation to Dr. Mandar Sohani, our Department Head, for his oversight and facilitation of this initiative. Finally, we extend our profound thanks to all the contributors who generously shared their valuable perspectives, steering us in the right direction throughout the course of this research endeavor.

REFERENCES

[1] P. K. Athira, C. J. Sruthi, A. Lijiya, “A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: An Indian Scenario”, 2019. https://www.sciencedirect.com/ science

[2] C. J. Sruthi, A. Lijiya, “Signet: A Deep Learning based Indian Sign Language Recognition System”, 2019.https://ieeexplore.ieee.org/ abstract/document/8698006

[3] T. Raghuveera, R. Deepthi, R. Mangalashri, R. Akshaya, “A depth-based Indian Sign Language recognition using Microsoft Kinect”,2020.https://link.springer.com/article/10.1007/s12046-019-1250-6.

[4] A. Mittal, P. Kumar, P. R. Roy, R. Balasubramanian, B. B. Chaudhuri, “A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion”,2019. https://ieeexplore.ieee.org/ abstract/document/8684245

[5] Pratibha Pandey, Vinay Jain, “Hand Gesture Recognition for Sign Language Recognition: A Review”, International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2015 .

[6] Neelam K. Gilorkar, Manisha M. Ingle, “A Review on Feature Extraction for Indian and American Sign Language”, International Journal of Computer Science and Information Technologies, Volume 5 (1) , pp314-318, 2014.

[7] Sharma, Shanu & Goyal, Sakshi & sharma, ishita. (2013). Sign Language Recognition System For Deaf And Dumb People. International Journal of Engineering Research and Technology. 2. 382-387.

[8] Sahoo, Ashok & Mishra, Gouri & Ravulakollu, Kiran. (2014). Sign language recognition: State of the art. ARPN Journal of Engineering and Applied Sciences. 9. 116-134.

[9] Reddygari Sandhya Rani , R Rumana , R. Prema, 2021, A Review Paper on Sign Language Recognition for The Deaf and Dumb, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 10 (October 2021),

[10] A. S. Nikam and A. G. Ambekar, "Sign language recognition using image based hand gesture recognition techniques," *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, Coimbatore, India, 2016, pp.

[11] P. K. Athira, C. J. Sruthi, A. Lijiya, “A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: An Indian Scenario”,2019.https://www.sciencedirect.com/ science

[12] C. J. Sruthi, A. Lijiya, “Signet: A Deep Learning based Indian Sign Language Recognition System”, 2019.https://ieeexplore. ieee.org/abstract/document/8698006

[13] T. Raghuveera, R. Deepthi, R. Mangalashri, R. Akshaya, “A depth-based Indian Sign Language recognition using Microsoft Kinect”,2020.https://link.springer.com/article/10.1007/s12046-019-1250-6.

[14] A. Mittal, P. Kumar, P. R. Roy, R. Balasubramanian, B. B. Chaudhuri, “A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion”,2019. https://ieeexplore.iee

❑ ❑ ❑ ❑ ❑ ❑

AET 008

AI-based tool for Preliminary Diagnosis of Dermatological Manifestations

Trisha Shah
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
shahtrisha70@gmail.com

Shipra Dagli
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
Shipradagli81@gmail.com

Neel Gude
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
neel.gude11@gmail.com

Athang Munankar
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
apm1902@gmail.com

Sanjeev Dwivedi
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
sanjeev.dwivedi@vit.edu.in

***Abstract* – Skin diseases pose significant challenges for accurate diagnosis, often requiring specialized expertise that may not be readily available in all healthcare settings. Through this project, we present an approach for automated identification of skin diseases using image analysis and machine learning techniques. Our method utilizes a large dataset of annotated skin disease images to train convolutional neural networks (CNNs) for classification tasks. We employ transfer learning to leverage pre-trained CNN architectures including but not limited to ResNet, MobileNet, DenseNet and Xception, and fine-tune them on our dataset to enhance classification performance. Through extensive experimentation and evaluation on a diverse range of skin diseases, which include Actinic Keratoses, Basal Cell Carcinoma, Benign Keratosis-like Lesions, Dermatofibroma, Melanocytic Nevi, Melanoma and Vascular lesions, our tool shows potential and can be valuable for dermatologists and healthcare professionals, by offering efficient and reliable support in the diagnosis and management of skin diseases. Our method has the potential to bridge the gap in access to specialized dermatological expertise, particularly in remote or underserved areas where dermatologists may be scarce. The integration of image analysis and machine learning techniques not only facilitates accurate disease identification but also enables deeper insights into the underlying patterns and featuresassociated with different skin conditions. This, in turn, can inform the development of new diagnostic tools, treatment strategies, and therapeutic interventions for various dermatological disorders. The study presents a promising step towards the advancement of automated skin disease diagnosis using cutting edge technology. By harnessing the power of image analysis and machine learning, we offer a scalable and efficient solution that holds great potential for improving patient outcomes and advancing dermatological care in diverse healthcare settings.**

I. INTRODUCTION

Skin health is undeniably crucial to our overall well-being, serving as a visible indicator of our internal health and often acting as a frontline defence against various pathogens and environmental stressors. Despite its significance, dermatological conditions frequently go underdiagnosed or misdiagnosed, leading to needless suffering and complications for millions worldwide. The Global Burden of Disease project highlights this issue, revealing that skin diseases persist as the fourth leading cause of nonfatal disease burden globally. Moreover, these conditions frequently manifest as the presenting face of more severe systemic illnesses, including HIV and neglected tropical diseases (NTDs), adding layers of complexity to their diagnosis and management. Beyond the physical implications, dermatological disorders also exact a toll on mental health, functional abilities, and social engagement, underscoring the multifaceted impact of skin health on individuals' lives. Access to dermatological expertise remains a challenge, particularly in remote or underserved regions, exacerbating the disparities in healthcare access and outcomes. In light of these challenges, the emergence of artificial intelligence (AI) presents a promising opportunity to revolutionize dermatological care. Through this project, we endeavour to harness the transformative potential of AI and machine learning to address the intricate challenges associated with dermatological diagnoses. Our primary objective is to develop a sophisticated tool capable of analysing images of skin conditions, thereby providing individuals with preliminary assessments of potential dermatological issues. This innovative approach aims to facilitate early and accurate detection, thereby enhancing overall skin health management and reducing the burden of dermatological diseases.

The proposed methodology encompasses a multi- stage process, beginning with image pre- processing techniques. The second stage involves the application of machine learning algorithms to identify and classify skin diseases based on the analysed features and observations. By integrating advanced technologies with dermatological expertise, our proposed system holds immense promise, particularly in rural areas where access to dermatologists is limited, potentially democratizing access to quality dermatological care.

II. LITERATURE REVIEW

Chen-Yu Zhu et al. 2021 [1] introduce an innovative approach to skin disease diagnosis, leveraging deep learning techniques on a dataset of 13,603 dermatologist-labeled dermoscopic images across 14 disease categories. The framework achieved notable classification accuracy, surpassing established CNN models and demonstrating comparable performance to board- certified dermatologists. This research underscores the potential of artificial intelligence (AI) in enhancing diagnostic accuracy in dermatology, particularly in real clinical environments.

Methodologically, Google's EfficientNet-b4 served as the backbone CNN architecture, with retraining on the dataset and implementation using PyTorch. Visualization techniques, including saliency maps and t-SNE, provided insights into the model's decision-making process. Overall, this study contributes significantly to advancing AI applications in dermatological diagnosis, promising improved patient care outcomes in dermatological clinics. Mingjun Wei et al. 2023 [2] outline the development of a skin disease classification model based on a convolutional neural network (CNN) and model fusion techniques. With skin diseases posing a significant public health concern due to their varied symptoms and progression, early identification remains challenging. The model addresses these challenges through deep and shallow feature fusion, alongside an attention module to enhance feature extraction. Evaluation on a private dataset predominantly comprising acne-like skin diseases and the public HAM10000 dataset reveals improved accuracy and F1-score compared to baseline models. The dataset, consisting of 2600 images covering various skin diseases such as acne, melasma, rosacea, and nevus of Ota, underwent data augmentation to counter imbalanced class distribution. Results showcase the model's superior performance with an accuracy of 96.54% on the private dataset, outperforming ResNet50, EfficientNet_B4, DenseNet201, and ConvNeXt_L models. Moreover, the model exhibits enhanced precision, recall, and F1-score across different skin disease classes. Experiments on the HAM10000 dataset further validate the model's effectiveness, highlighting its potential to accurately identify and classify skin diseases, thus addressing critical public health challenges.

Reza Zare et al. 2021 [3] explore the utilization of deep learning techniques for segmenting and classifying dermatoscopic skin lesion images, with a primary goal of facilitating early detection of skin cancer. It underlines the pivotal role of accurate identification in improving patient outcomes and survival rates, advocating for image processing as a vital diagnostic aid. Employing U-Net for segmentation and DenseNet121 for classification, renowned for their accuracy, the study showcases acceptable results on the ISIC-2018 and HAM10000 datasets, achieving 79.49% and 93.11% accuracy in discriminating cancerous and non-cancerous samples, respectively. Extensive literature review encapsulates previous research efforts, emphasizing advancements in convolutional neural networks (CNN) and deep learning models. Detailed descriptions of datasets and methodological approaches, including data preparation, model architecture, and hyperparameter configurations, are provided. Evaluation metrics, including accuracy, sensitivity, specificity, AUC, precision, and F1-score, demonstrate the proposed method's superiority in comparison to previous research, particularly in accuracy and detection rates. Notably, insights into sample sizes underscore the efficacy of the approach, achieving accurate classification with a smaller number of training samples. Overall, the paper presents a comprehensive overview of the potential of deep learning models in dermatoscopic image analysis, with significant implications for enhancing early skin cancer diagnosis and treatment. Gouda & Amudha, 2020 [4] discuss the application of deep learning, specifically Residual Neural Network (ResNet), in skin cancer classification, addressing challenges posed by skin texture complexities and visual similarities between diseases. Emphasizing the significance of early detection, especially for diseases like melanoma, conventional methods may struggle with, it highlights the potential of deep learning, particularly Convolutional Neural Networks (CNNs) and ResNet, in overcoming this challenge. It explores ResNet's implementation for skin cancer classification, utilizing the ISIC dataset for training & testing, with emphasis on image pre- processing for enhanced analysis. Detailed insights into ResNet architecture, including fully connected layers, convolutional layers, & data augmentation, are provided, with ResNet-34 model showcasing superior recall & accuracy. The paper further elucidates on diagnosis confidence generation, employing sigmoid conversion for precise classification, and presents results indicating model accuracy & disease probabilities in a CSV file. It concludes by underlining deep learning's potential in revolutionizing medical diagnosis, particularly in skin cancer, with opportunities for further research & development in this domain.

Mehra et al., 2021 [5] delve into the utilization of transfer learning and ResNet-50, a deep residual network, for the classification of skin cancer, stressing the urgency of early diagnosis in addressing this global public health concern. Focused on categorizing seven types of skin cancer using the ResNet-50 model trained on the HAM10000 dataset, comprising 10,015 dermoscopic images, the study evaluates the model's performance through various metrics, achieving notable accuracy, precision, recall, and F1-score values. It provides an extensive overview of related work, underlining the significance of deep learning models and convolutional neural networks (CNNs), alongside the efficacy of transfer learning. Detailed insights into the proposed model's architecture, image pre- processing steps, implementation details, & results, including confusion matrix & performance comparisons with VGG16 and CNN models, are elucidated. This paper underscores the model's potential to aid dermatologists in skin lesion inspection, emphasizing the need for a balanced dataset to further refine accuracy. Overall, it offers a comprehensive and methodical approach to skin cancer classification, presenting significant implications for early diagnosis and treatment.

1. PLANNING AND FORMATION

A. Model Overview

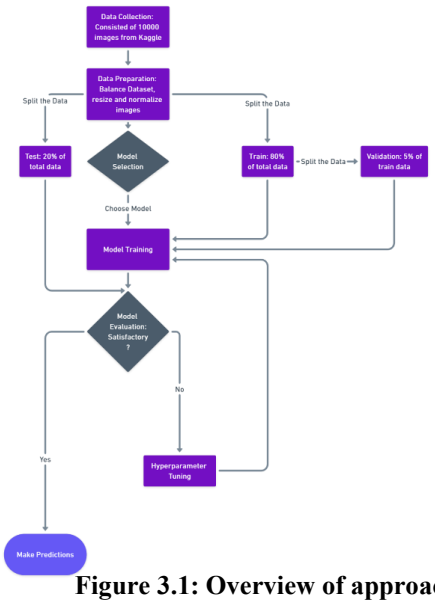


Figure 3.1: Overview of approach

B. Data Collection

For this project, we use the HAM10000 dataset from Kaggle, which contains high-quality images of skin lesions with associated clinical metadata, including disease labels, patient demographics, and lesion characteristics. The HAM10000 dataset contains approximately 10,000 images of skin lesions, making it one of the largest datasets of its kind. The images cover a wide range of skin diseases and conditions, including melanoma, nevus (moles), basal cell carcinoma, actinic keratoses, dermatofibroma and vascular lesions. Each image is accompanied by clinical metadata, including patient information and lesioncharacteristics.



Figure 3.3: Sample images from Dataset

C. *Data Preparation*

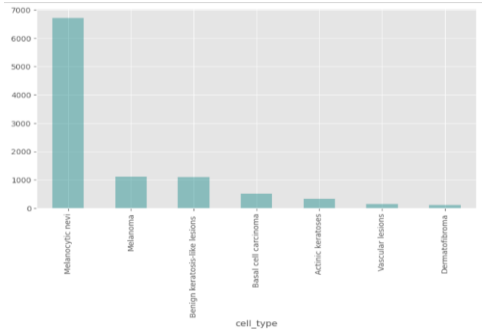


Figure 3.5: Distribution of DatasetClasses

2. Pre-process the Data:

1. Modify the Metadata Dataframe

In this step, we first map the image names in the metadata csv to their respective pathsin the dataset.

The labels for each image was a shortened version of the disease. Thus, we created a dictionary with the actual names of the disease and created a column 'cell_type' with the actual names of the disease. In addition to this, we also appended the id ofthe disease.

2. Balance the Dataset

The dataset was extremely skewed with nearly 7000 of the 10000 images belonging to the majority class in contrast to the 115 images belonging to the smallest class. Webalanced the skewed dataset using the resample method.

In this project, we under sample the majority classes and oversample the minority class to obtain 1000 images ineach dataset. Thus, our dataset will now consist of 7000 images in total with 1000 images from each of the 7 classes. The balanced dataset produced by the code enables machine learning models to learn from a more evenly distributed representation of the classes, potentially leading to improved model performanceand generalization.

3. Image Processing

First, we resize the image to the size specified using the resize function from OpenCV. In this case, we resize the image to 128, 128. Next, we convert the pixel values of the image from their current data type to 32-bit floating-point format (np.float32), which is a common data type used in numerical computations. Then we divide each pixel value by 255.0, which is the maximum possible value for a pixel in the original 8-bit grayscale image. Dividingby this maximum value ensures that the resulting pixel values are normalized to therange [0, 1].

4. Create Train, Validation and TestDatasets

The dataset is then split into train (76% of the data),validation (4% of the data) and test dataset (20% ofthe data).

5. Implement the Model

In this project, we experiment with 7 differenttypes of pre-trained models. The models are downloaded and modified to suit the needs of our data.

III. MODEL BUILDING

Initially we applied architecture 1 to all models.

Architecture 1:

This is the initial Architecture used for training allthe models.

1. Download all layers of the model except the toplayer
2. Add a flatten layer
3. Add a Dropout layer of 0.25
4. Add a Dense Layer with 512 neurons and Rectified Linear Unit Activation
5. Add Batch Normalization
6. Add a dropout layer of 0.46
7. Add a Dense Layer with 128 neurons and Recitifed Linear Unit Activation
8. Add Batch Normalization
9. Add a dropout layer of 0.46
10. Add the output layer with 7 neurons and SoftMax activation

Architecture 2:

To improve the models, we modified the architecture of DenseNet, InceptionV3 and MobileNet as follows:

DenseNet:

For both, DenseNet121 and DenseNet201, thesame architecture was implemented.

1. Download all layers of the model except the toplayer
2. Add a flatten layer
3. Add Batch Normalization
4. Add a Dropout layer of 0.3
5. Add a Dense Layer with 512 neurons and Recitifed Linear Unit Activation
6. Add Batch Normalization
7. Add a dropout layer of 0.5
8. Add a Dense Layer with 256 neurons and Recitifed Linear Unit Activation
9. Add Batch Normalization
10. Add a dropout layer of 0.5
11. Add the output layer with 7 neurons & softmax activation

InceptionV3:

1. Download all layers of the model except the top layer
2. Add a flatten layer
3. Add a Dropout layer of 0.25
4. Add a Dense Layer with 512 neurons and Recitifed Linear Unit Activation
5. Add Batch Normalization
6. Add a dropout layer of 0.5
7. Add a Dense Layer with 512 neurons and Recitifed Linear Unit Activation
8. Add Batch Normalization
9. Add a dropout layer of 0.5
10. Add the output layer with 7 neurons & softmax activation

MobileNetV2:

1. Download all layers of the model except the top layer
2. Add a flatten layer
3. Add a Dropout layer of 0.25
4. Add a Dense Layer with 256 neurons and Recitifed Linear Unit Activation
5. Add Batch Normalization
6. Add a dropout layer of 0.25
7. Add a Dense Layer with 256 neurons and Recitifed Linear Unit Activation
8. Add Batch Normalization
9. Add a dropout layer of 0.25
10. Add a Dense Layer with 128 neurons and Recitifed Linear Unit Activation
11. Add Batch Normalization
12. Add a dropout layer of 0.25
13. Add the output layer with 7 neurons & softmax activation

Model Training

The model was compiled using the following:
Loss: Sparse Categorical Entropy Optimizer: Adam, Learning Rate = 0.001Metrics: Accuracy

After compiling the model, we fit it to the Train andValidation Data. While Training, we specified a Batch Size of 128 and trained the model for over 10 epochs. Additionally, we save the weights of the epoch with the minimum validation loss and use these to make predictions on the test dataset.

The Models

1) DenseNet121
2) DenseNet201
3) InceptionV3
4) MobileNetV2
5) ResNet152V2
6) VGG19
7) Exception

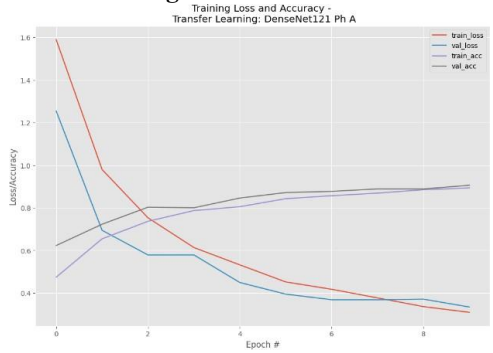
IV. RESULTS

Performance Analysis

DenseNet121 Architecture 1

This model required a relatively smaller computing power, with a CPU Time of 54 minutes and 52 seconds to train the images in the train dataset. The model achieved a training, validation and test accuracy of 0.8932, 0.9057 and 0.9693 thus indicating that it performs well on unseen data. Thetrain, validation and test losses were 0.3091, 0.3342 and 0.1127 respectively. We used weights from the 10th epoch for the testing dataset.

Figure 6.1: Training and Validation loss andaccuracy over 10 epochs for DenseNet121 Architecture 1



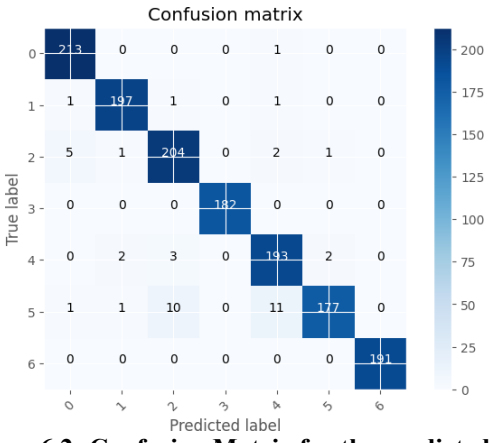


Figure 6.2: Confusion Matrix for the predictedvalues of the Test Set using DenseNet121 Architecture 1

DenseNet121 Architecture 2

This model required higher computing power, with a CPU Time of 2h 14min 15s to train the images inthe train dataset. The model achieved a training, validation and test accuracy of 0.9611, 0.9057 and0.9879 thus indicating that it performs well on unseen data. In fact, this model had the highest test accuracy. The train, validation and test losses were 0.1109, 0.3543 and 0.0627 respectively. As the accuracy from the 10th epochs showed signs of overfitting, we used weights from the 8th epoch forthe testing dataset.

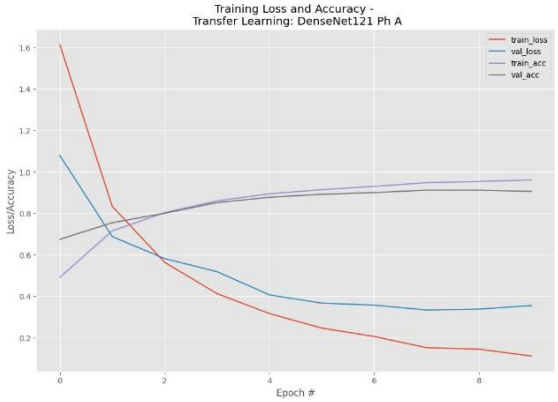


Figure 6.3: Training and Validation loss andaccuracy over 10 epochs for DenseNet121 Architecture 2

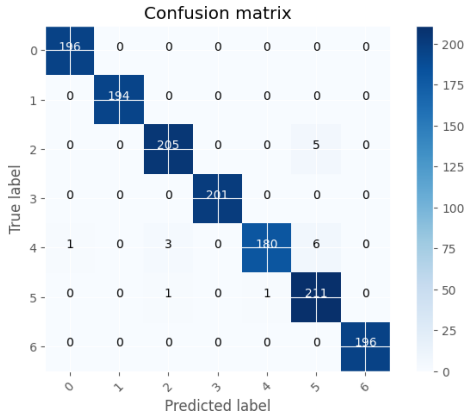


Figure 6.4: Confusion Matrix for the predictedvalues of the Test Set using DenseNet121 Architecture 2

DenseNet201 Architecture 1

This model required higher computing power in comparison to DenseNet121, with a CPU Time of 1h 28min 2s to train the images in the train dataset. This can be attributed to the increase in total number of model parameters. The model achieved a training, validation and test accuracy of 0.9259, 0.9143 and 0.9629 thus indicating that it performs well on unseen data. The train, validation and test losses were 0.2154, 0.3317 and 0.1136 respectively. We used weights from the 8th epoch for the testing dataset as it had the lowest validationloss.

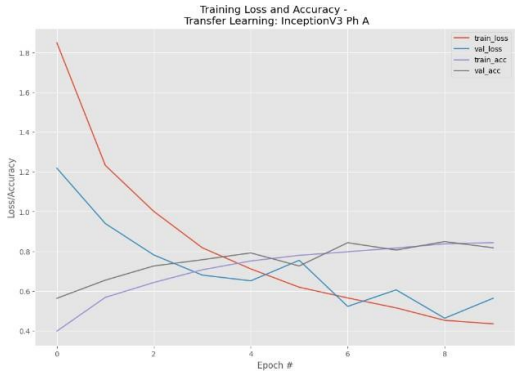


Figure 6.5: Training and Validation loss andaccuracy over 10 epochs for DenseNet201 Architecture 1

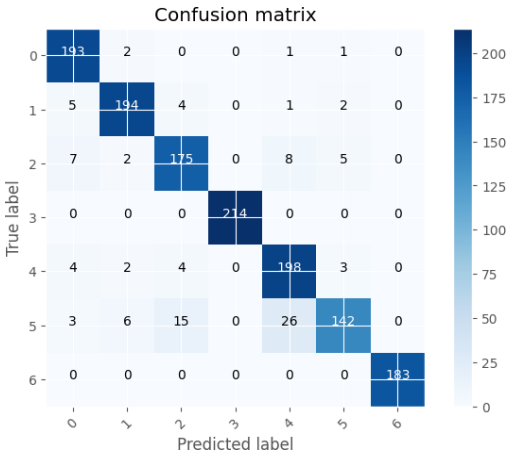


Figure 6.6: Confusion Matrix for the predictedvalues of the Test Set using Densenet201 Architecture 1

DenseNet201 Architecture 2

It had a CPU Time of 3h 17min 6s, thus showing a dramatic increase in computation. The model achieved a training, validation and test accuracy of 0.9659, 0.8914 and 0.9829 thus indicating that it performs well on unseen data. The train, validation and test losses were 0.1002, 0.3599 and 0.0773 respectively. We used weights from the 7th epoch for the testing dataset as it had the lowest validation loss. This model had one of the best performances, but required higher computation.

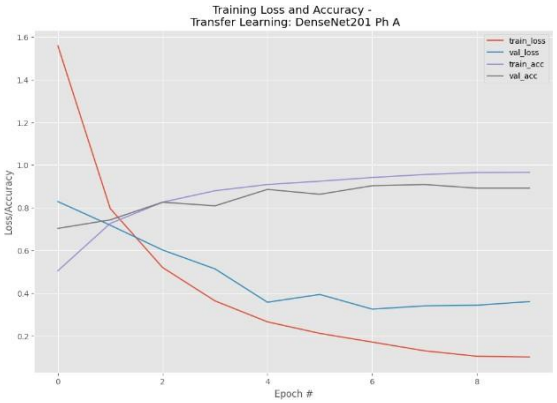


Figure 6.7: Training and Validation loss andaccuracy over 10 epochs for DenseNet201 Architecture 2

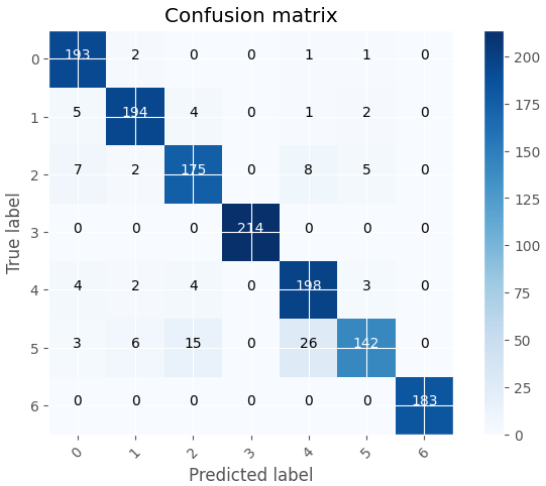


Figure 6.8: Confusion Matrix for the predictedvalues of the Test Set using Densenet201 Architecture 2

InceptionV3 Architecture 1

This model required a relatively lower computation, with a CPU Time of 42min to train the images in the train dataset. The model achieveda training, validation and test accuracy of 0.8430, 0.8171 and 0.9279, thus indicating that it performswell on unseen data. The train, validation and test losses were 0.4343, 0.5632 and 0.2547 respectively. We used weights from the 9th epoch for the testing dataset as it had the lowest validationloss.

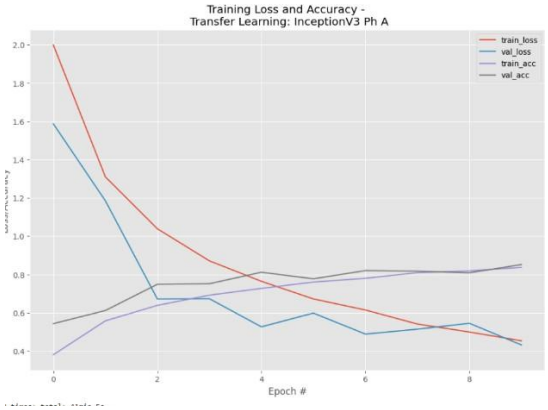


Figure 6.9: Training and Validation loss andaccuracy over 10 epochs for InceptionV3 Architecture 1

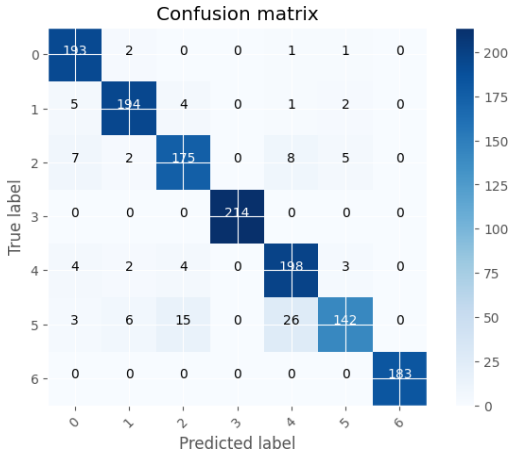


Figure 6.10: Confusion Matrix for the predictedvalues of the Test Set using InceptionV3 Architecture 1

InceptionV3 Architecture 2

This model did not have a significant difference in the computation time, with a CPU Time of 41min 5s to train the images in the train dataset. The model achieved a training, validation and test accuracy of 0.8368, 0.8514 and 0.9429, thus indicating that it performed slightly better on unseen data. The train, validation and test losses were 0.4533, 0.4317 and 0.2210 respectively. These values indicate that the new architecture is performs better than the previous architecture. We used weights from the 10th epoch for the testing dataset as it had the lowest validation loss.

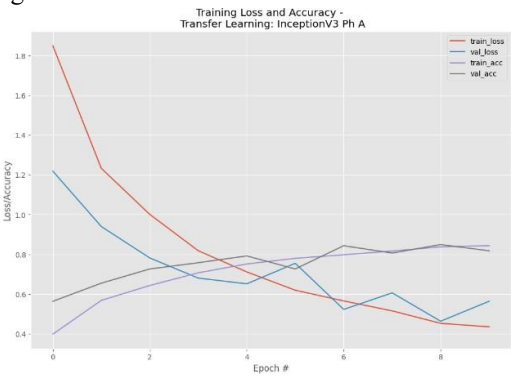


Figure 6.11: Training and Validation loss andaccuracy over 10 epochs for InceptionV3 Architecture 2

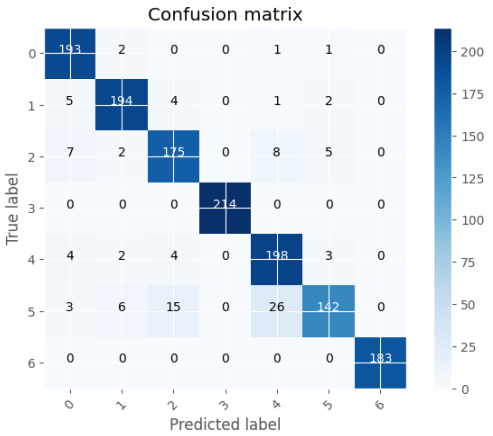


Figure 6.12: Confusion Matrix for the predictedvalues of the Test Set using InceptionV3 Architecture 2

MobileNetV2 Architecture 1

This model required the lowest computation, with a CPU Time of 24min 40s to train the images in thetrain dataset. The model achieved a training, validation and test accuracy of 0.9116, 0.8629 and 0.9443, thus indicating that it performs well on unseen data. The train, validation and test losses were 0.2530, 0.4287 and 0.1826 respectively. This implies that the model did slightly overfit at the 10thepoch. Thus, we used weights from the 7th epoch for the testing dataset. The relatively high accuracy and low computation power proves that MobileNet is a lightweight network which intends to provide accurate results for mobile devices by utilizinglower computation power.

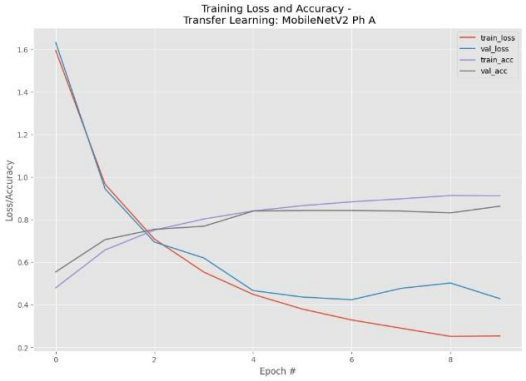


Figure 6.13: Training and Validation loss and accuracy over 10 epochs for MobileNetV2 Architecture 1

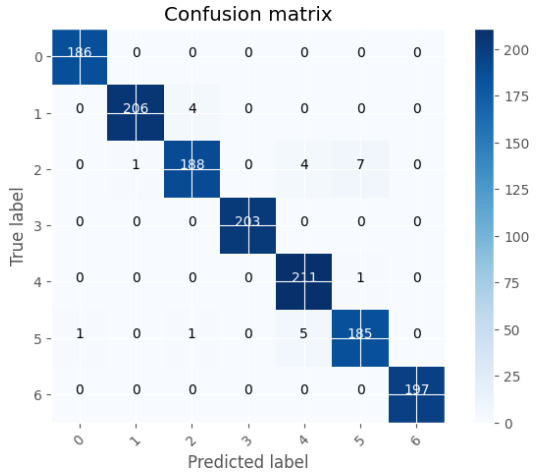


Figure 6.14: Confusion Matrix for the predictedvalues of the Test Set using MobileNetV2 Architecture 1

MobileNetV2 Architecture 2

The CPU time decreased dramatically from 24min 40s to 9min 23s on changing the architecture. The model achieved a training, validation and test accuracy of 0.9343, 0.9029 and 0.9650, thus indicating an improvement in performance. The train, validation and test losses were 0.1872, 0.3527 and 0.1077 respectively. We used weights from the 9th epoch for the testing dataset as it had the minimum validation loss. With its high accuracy and low computation power MobileNet continues to prove that it is an appropriate choice for lightweight applications.

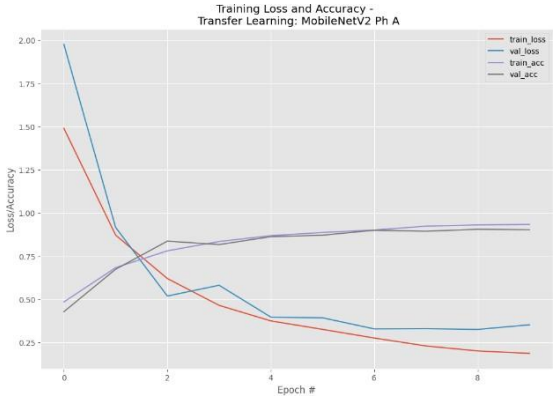


Figure 6.15: Training and Validation loss andaccuracy over 10 epochs for MobileNetV2 Architecture 2

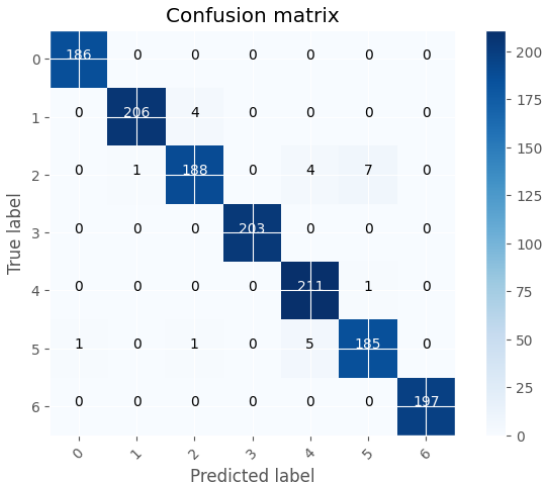


Figure 6.16: Confusion Matrix for the predictedvalues of the Test Set using MobileNetV2 Architecture 2

ResNet152V2

This model required one of the highest computation, with a CPU Time of 3h 11min 11s to train the images in the train dataset. It also had the highest number of parameters. The model achieved a training, validation and test accuracy of 0.9573, 0.9029 and 0.9843, thus indicating that it performs well on unseen data. The train, validation and test losses were 0.1309, 0.3080 and 0.0690 respectively. However, despite the high test accuracy, it appears as though the model began to overfit after the 7th epoch as the validation loss stopped reducing and the difference in training andvalidation accuracy began to increase. Thus, we used weights from the 7th epoch for the testing dataset. However, it is also computationally expensive to implement.

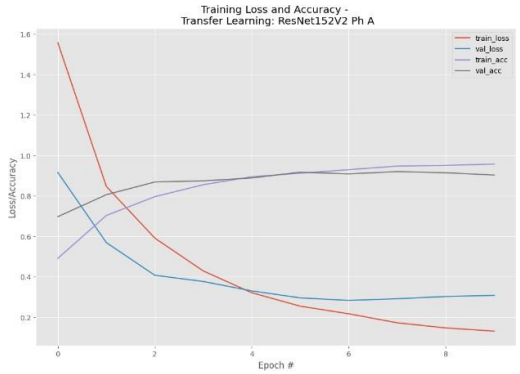


Figure 6.17: Training and Validation loss andaccuracy over 10 epochs for ResNet152V2

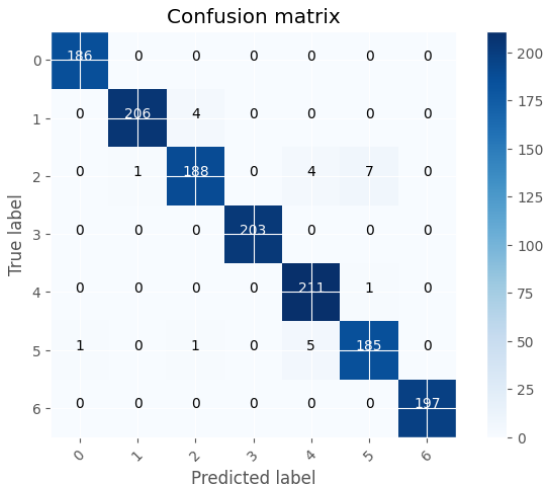


Figure 6.18: Confusion Matrix for the predictedvalues of the Test Set using ResNet152V2

VGG19

This model required the highest computation, with a CPU Time of 5h 51min 11s to train the images inthe train dataset. The model achieved a training, validation and test accuracy of 0.7262, 0.7486 and0.8199, thus indicating that it performs well on unseen data. The train, validation and test losses were 0.7353, 0.7194 and 0.5119 respectively. The model's performance was poor in comparison to that of the other models. As the model hadn'toverfit yet, training it over a few more epochs couldserve to increase accuracy. For the testing dataset, we used weights from the 10th epoch.

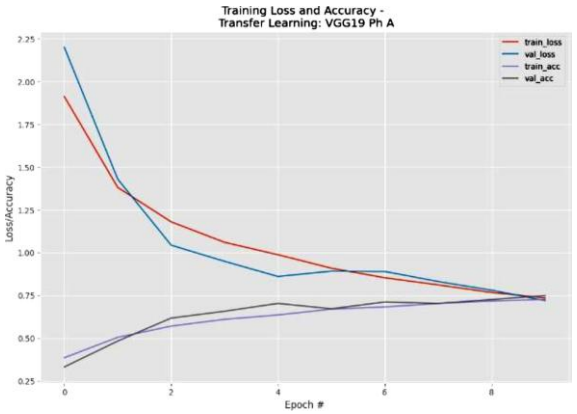


Figure 6.19: Training and Validation loss andaccuracy over 10 epochs for VGG19

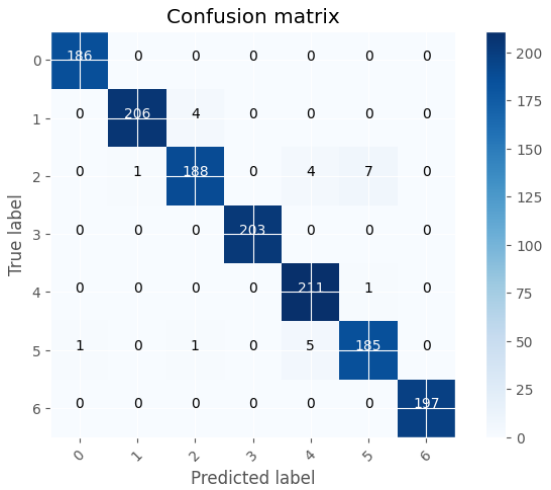


Figure 6.20: Confusion Matrix for the predictedvalues of the Test Set using VGG19

This model required a relatively high computation, with a CPU Time of 1h 39min 21s to train the images in the train dataset. The model achieved a training, validation and test accuracy of 0.9212, 0.8629 and 0.9807, thus indicating that it performs well on unseen data. The train, validation and test losses were 0.2285, 0.4581 and 0.1033 respectively. The model's performance was amongst the best, but it is also slightly computationally expensive. We used weights from the 10th epoch as it had the lowest validation loss.

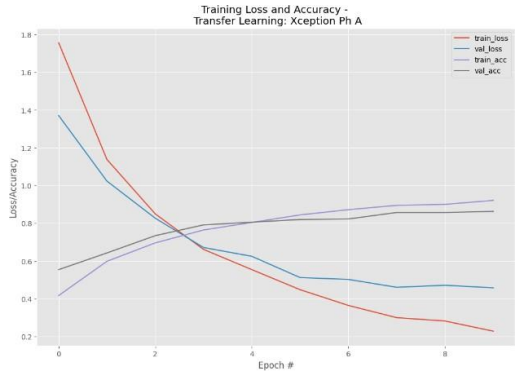


Figure 6.21: Training and Validation loss andaccuracy over 10 epochs for Xception

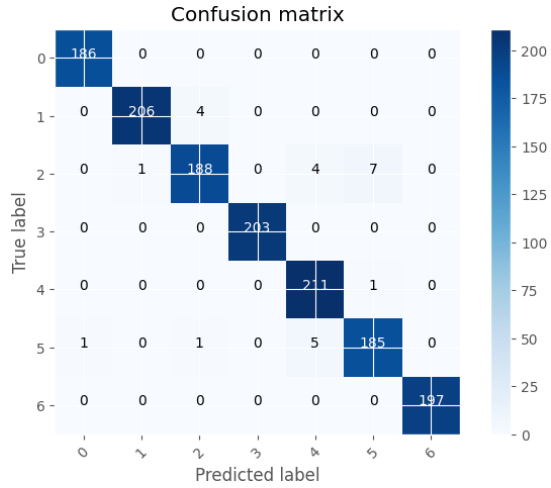


Figure 6.22: Confusion Matrix for the predictedvalues of the Test Set using Xception

Table 6.1: Comparison of Results

Sr. No.	Model	Epoch No. of weights for testing	Parameters	CPU Time	Training Accuracy after epoch 10	Validation Accuracy after epoch 10	Testing Accuracy	Testing Loss
1.	DenseNet 121	10	15495751	54min 52s	0.8932	0.9057	0.9693	0.1127
2.	DenseNet 121 Arch2	8	15628359	2h 14min 15s	0.9611	0.9057	0.9879	0.0627
3.	DenseNet 201	8	34120263	1h 28min 2s	0.9259	0.9143	0.9629	0.1136
4.	DenseNet 201 Arch2	7	34310215	3h 17min 6s	0.9659	0.8914	0.9829	0.0773
5.	Inception V3	9	26066727	42min	0.8430	0.8171	0.9279	0.2547
6.	Inception V3 Arch2	10	26267943	41min 5s	0.8368	0.8514	0.9429	0.2210
7.	MobileNetV2	7	12813383	24min 40s	0.9116	0.8629	0.9443	0.1826
8.	MobileNetV2 Arch2	9	7603271	9min 23s	0.9343	0.9029	0.9650	0.1077
9.	ResNet152V2	7	75178503	3h 11min 11s	0.9573	0.9029	0.9843	0.0690
10.	VGG19	10	24288327	5h 51min 11s	0.7262	0.7486	0.8199	0.5119

Best Performance: DenseNet121
Worst Performance: VGG19 Most
CPU Time: VGG19
Least CPU Time: MobileNetV2

V. CONCLUSION

The findings reveal that while each model exhibits distinctive performance characteristics, there is no one-size-fits-all solution. DenseNET121 and ResNet152V2 emerged as top performers in terms of accuracy, demonstrating superior capability in accurately identifying dermatological conditions. However, this accuracy often came at the expense of longer computation times, making them less practical for real-time clinical application.

On the other hand, MobileNetV2 showcased commendable accuracy while significantly reducing computation time, offering a balance between performance and efficiency. These models could serve as viable options for applications where rapid diagnosis is essential, such as telemedicine or point-of-care diagnostics.

Furthermore, Inception V3 and Xception demonstrated competitive performance, albeit with varying trade-offs between accuracy and computation time, thereby highlighting the importance of considering specific use-case requirements.

In summary, the selection of an AI-based model for preliminary diagnosis of dermatological manifestations should be made based on a thorough understanding of the trade-offs between accuracy and computation time, as well as the specific needs of the clinical setting. Future research focuses on optimizing existing models or developing hybrid approaches to further enhance performance while minimizing computation overhead, ultimately advancing the utility of AI in dermatological diagnosis and patient care.

VI. FUTURE SCOPE

While our current study has yielded valuable insights, there exists a promising avenue for further exploration and improvement. One of the limitations encountered during

this project was the constraint imposed by hardware limitations, which restricted the size of the dataset used for analysis. This constraint necessitated a careful selection process for dataset inclusion, potentially leading to biases in the analysis, particularly in relation to skin colour, texture, and other skin characteristics.

In the future, with access to better computing power and resources, there is an opportunity to address these limitations by leveraging larger and more diverse datasets. By expanding the dataset, encompassing a broader range of skin types, colours, and textures, we can strive for greater representativeness and inclusivity in our analyses. This expanded dataset would enable more comprehensive exploration and understanding of dermatological conditions and treatments acrossvarious demographic groups.

In this study, we present a comprehensive methodology for skin disease classification utilizing various convolutional neural network (CNN) models trained on a diverse dataset of dermatological images. Moving forward, our research suggests several promising avenues for future exploration. Firstly, we propose fine-tuningand optimization strategies to enhance the performance of our CNN models, potentially leading to improved classification accuracy. Furthermore, the integration of multi-modal data fusion techniques, such as patient demographics and clinical history, holds promise for augmentingthe robustness of our classification system. We alsoemphasize the importance of interpretability in AI-driven diagnosis by exploring methods to elucidate the decision-making process of our models, aiding clinicians in understanding and trusting the classification outcomes. Moreover, we advocate for real-time deployment of our models in clinical settings and their integration with telemedicine platforms to improve access to dermatological care, particularly in underserved regions. Collaboration with dermatologists and validation of our system in real-world clinical settings are essential steps in ensuring its clinical relevance and usability. Lastly, we suggest extending our research to include longitudinal analysis and disease progression monitoring, as well as generalizing our findings to other medical imaging tasks, thus contributing to the advancement of computer-aided diagnosis in dermatology and beyond.

REFERENCES

[1] Zhu, C., Wang, Y., Chen, H., Kunlun, G., Shu, C., Wang, J., Yan, L., Yang, Y., Xie, F., & Liu, J. (2021, April 16). A Deep Learning Based Framework for Diagnosing Multiple Skin Diseases in a Clinical Environment. Frontiers in Medicine. https://doi.org/10.3389/fmed.2021.626369

[2] Wei, M., Wu, Q., Ji, H., Wang, J., Lyu, T., Liu, J., & Li, Z. (2023, January 14). A Skin Disease Classification Model Based on DenseNet and ConvNeXt Fusion. Electronics. https://doi.org/10.3390/electronics12020438

[3] Zare, R., & Pourkazemi, A. (2021, October 9). DenseNet approach to segmentation and classification of dermatoscopic skin lesions images. arXiv.org. https://arxiv.org/abs/2110.04632

[4] N. Gouda and J. Amudha, "Skin Cancer Classification using ResNet," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2020, pp. 536-541, doi: 10.1109/ICCCA49541.2020.9250855

[5] Mehra, A., Bhati, A., Kumar, A., & Malhotra, R. (2021, January 1). Skin Cancer Classification Through Transfer Learning Using ResNet-50. Advances in Intelligent Systems and Computing. https://doi.org/10.1007/978-981-33-4367-2_6

[6] Zhang, B., Zhou, X., Luo, Y. et al. Opportunities and Challenges: Classification of Skin Disease Based on Deep Learning. Chin. J. Mech. Eng. 34, 112 (2021). https://doi.org/10.1186/s10033-021-00629-5

[7] Saja Salim Mohammed and Jamal Mustafa Al-Tuwaijari 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1076 012045

[8] Mamun, M. A., Mamun, M. A., & Uddin, M. S. (2022, November 1). A machine learning approach for skin disease detection and classification using image segmentation. Healthcare Analytics. https://doi.org/10.1016/j.health.2022.100122

[9] Agarwal, R., & Godavarthi, D. (2023, October 2). Skin Disease Classification Using CNN Algorithms. EUDL. https://eudl.eu/doi/10.4108/eetpht.9.4039

[10] De A, Mishra N, Chang H. 2024. An approach to the dermatological classification ofhistopathological skin images using a hybridized CNN-DenseNet model. PeerJ Computer Science 10:e1884 https://doi.org/10.7717/peerj-cs.1884

[11] Sun J, Yao K, Huang G, Zhang C, Leach M, Huang K, Yang X. Machine Learning Methods in Skin Disease Recognition: A Systematic Review. Processes. 2023; 11(4):1003. https://doi.org/10.3390/pr11041003

[12] Srinivasu, P.N.; SivaSai, J.G.; Ijaz, M.F.; Bhoi, A.K.; Kim, W.; Kang, J.J. Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM.Sensors 2021, 21, 2852. https://doi.org/10.3390/s21082852

❑ ❑ ❑ ❑ ❑ ❑

AET 009

Plant Disease Detection

Mandar Thakur
Dept. of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India mandar.thakur@vit.edu.in

Pranav Tambre
Dept. of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India pranav.tambre@vit.edu.in

Ganesh Bhagat
Dept. of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India ganesh.bhagat@vit.edu.in

Akshata Kadam
Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India akshata.kadam@vit.edu.in

Prof. Suvarna Bhatt
Computer Engineering Vidyalankar
Institute of Technology
Mumbai, India suvarna.bhatt@vit.edu.in

***Abstract*—Plants have become a crucial energy source in the fight against global warming, but their survival is threatened by diseases. Utilizing convolutional neural networks (CNNs), known for their superior performance in image classification, this study investigates their potential in identifying plant diseases from leaf photos taken in natural settings. Highlighting the significance of agriculture and the challenges posed by plant ailments, early and accurate disease detection is emphasized to minimize crop damage. Traditional methods by farmers are deemed inadequate, necessitating computerized image processing techniques. Various machine learning and deep learning methods have been explored, with CNNs being favored for their ability to extract relevant im- age features autonomously. The choice between machine learning and deep learning depends on factors like the specific problem and data availability, with deep learning, particularly through CNNs, preferred when ample data and computational resources are available. This paper focuses on various plant diseases, providing insights into the performance, evaluation metrics, and outcomes of techniques previously used for detecting and classifying them using AI-based image processing.**

***Keywords*— Agriculture, Image processing techniques, Deep learning, Convolutional neural network (CNN)**

I. INTRODUCTION

Agricultural biodiversity is essential for providing humans with food and raw materials and is an essential component of human civilization [1], [2]. Limited accessibility to agricultural offices ham-pers their ability to promptly determine the nature of infections in their harvests. This research project seeks to tackle this issue by employing image processing and machine learning techniques, specifically focusing on Convolutional Neural Networks (CNN), to identify plant diseases based on their morphology. Pests and diseases significantly contribute to crop destruction, resulting in reduced food production and food insecurity, particularly in less developed countries with limited knowledge about pest management and disease control. Traditional laboratory-based approaches, such as polymerase chain reaction, gas chromatography, mass spectrometry, thermography, and hyper-spectral techniques, are effective but prove to be costly and time-consuming. In recent times, server-based and mobile-based approaches have gained prominence for their advantages, including high-resolution cameras, high-performance processing, and automatic disease recognition capabilities. This research leverages CNNs, a deep learning algorithm, to enhance disease recognition rates and accuracy. CNNs have proven to be effective in various image recognition tasks, making them particularly suitable for plant disease detection. Within the realm of CNNs, the study incorporates three key component descriptors: Hu moments for extracting leaf shapes, Haralick texture for capturing leaf textures, and Color Histogram for representing the color distribution in images.

By amalgamating these advanced technologies and methodologies, this research aims to create an innovative and efficient solution for plant disease detection, with a primary focus on early disease identification in regions with limited agricultural resources. Beyond immediate crop destruction, the impact of plant diseases extends to the overall economic stability of communities. The reduced food production resulting from unchecked diseases exacerbates the prevalent issue of food insecurity in less developed countries. The integration of image processing and CNNs in this research not only addresses the immediate need for accurate disease identification but also presents a proactive approach to mitigating the broader socio-economic implications associated with crop loss. By focusing solely on the CNN model, the research aims to contribute to the ongoing discourse within the scientific community on the most effective approaches to plant disease detection.

II. LITERATURE SURVEY

A. Existing Systems: In the realm of plant disease detection, traditional machine learning (ML) methods have long been relied upon, employing techniques like feature extraction and classification. These methods analyze various aspects of images, such as color, texture, and shape, to train algorithms capable of discerning between healthy and diseased plants. While effective for common diseases like leaf blotch and powdery mildew, these approaches often struggle when it comes to identifying subtle symptoms and catching diseases at their earliest stages. Moreover, they face challenges in processing high-resolution images, limiting their applicability in certain scenarios [3]–[5].

On the other hand, deep learning (DL) techniques, particularly convolutional neural networks (CNNs) and deep belief networks (DBNs), have emerged as powerful alternatives [6], [7]. DL models have the capability to learn intricate patterns and features from images, enabling them to detect even the most subtle signs of diseases that traditional methods may overlook. They also exhibit proficiency in handling complex and high-resolution images, making them well-suited for a wide range of applications in plant disease detection. However, the effectiveness of DL models heavily relies on the availability of large labeled datasets and significant computational resources, which can be limiting factors in certain contexts [8]–[10]. To address these challenges and enhance the performance of DL models, recent trends in plant disease detection have seen the adoption of transfer learning and ensemble methods [11]. Transfer learning involves fine-tuning pre-trained models on specific datasets, thereby leveraging existing knowledge to improve performance on new tasks. Ensemble methods, on the other hand, combine multiple models to achieve better overall performance and reduce reliance on any single model. These approaches not only enhance the robustness and accuracy of disease detection models but also help mitigate

issues such as overfitting, where models perform well on training data but poorly on unseen data. Additionally, data augmentation techniques are employed to increase the diversity of datasets and reduce the need for extensive labeled data, further improving the performance and generalization ability of models. Various deep learning architectures have been explored in the context of plant disease detection, each with its unique strengths and applications. Deep belief networks (DBNs), characterized by multiple layers of Restricted Boltzmann Machines (RBMs), have demonstrated high accuracy in classifying images of plant leaves affected by diseases and pests [12]. Similarly, Boltzmann's Deep Machine (DBM) has shown promise in unsupervised classification for detecting plant lesions [13]. Deep Denoising Autoencoder (DDA) serves dual purposes, effectively removing noise from plant leaf data and accurately predicting diseases [14]. However, among these architectures, deep convolutional neural networks (CNNs) stand out as particularly effective for plant disease classification. Their specialized architecture, consisting of convolutional and pooling layers, allows them to extract intricate features from images, leading to remarkably high accuracy rates ranging from 99% to 99.2% in classifying images of plant leaves affected by diseases and pests [15].

B. Convolutional Neural Network: Traditional fixed feature extraction techniques often fail to acquire valuable features for image segmentation [20]. The CNN led to improved feature extraction and comprised several integration and layering stages, culminating in completely integrated layers. CNN has the capability to harvest picture features and complete image classification simultaneously and has the benefit of immediately publishing featured photographs [21]. The convolution layer, pooling layer, and fully connected layer are the three primary components of CNN. These layers cooperate to perform classification. During the process, the model's features are first acquired and then added to the fully linked layer for classification. CNN's local perception and weight sharing are the primary distinctions between it and the general neural network. It offers deep layers of learning and is used repeatedly to classify images [22]. The convolution kernel extracts the local features of the image, and each region of the image uses the same convolution kernel, which helps to get good accuracy. The convolutional layer is the structure of center square of a CNN. The parameters layer consist of a group of channels (or pieces) that are learnable, and having a small open field, on the other hand get to full info volume complexity. Therefore, the system learns channels which are initiated once it classifies some sort of specific highlight at some spatial positioning information [23].

1. VGGNet: The runner-up in the ILSVRC 2014 challenge was VGGNet [24], developed by Karen Simonyan and Andrew Zisserman from the Visual Geometry Group (VGG) research lab at Oxford University. It had a very simple and classical architecture, with 2 or 3 convolutional layers and a pooling layer, then again 2 or 3 convolutional layers and a pooling layer, and so on (reaching a total of just 16 or 19 convolutional layers, depending on the VGG variant), plus a final dense network with 2 hidden layers and the output layer. It used only 3 × 3 filters, but many filters.

2. AlexNet: AlexNet is a type of neural network with many parts for processing images. It has layers for spotting patterns, like edges or shapes, in pictures. Some layers help make the picture smaller to speed things up. It's big, with over 60 million settings for making sense of images. It uses a special math formula called ReLU to make decisions about what's in the picture. Also, it uses a certain amount of pictures (128 at a time) to learn better and adjusts its understanding as it goes. Plus, it has tricks to make pictures look different so it can learn from them better, like flipping them or changing their colors a bit.

3. GoogleNet: Google Net's standout feature is its inception module, a fundamental component of its design. This module incorporates various convolutional layers with different sizes of patterns to capture diverse features simultaneously. By doing this, Google Net can grasp both small details and broader features in images effectively. An important aspect of the inception module is its ability to reduce computational load by first using smaller convolutions to shrink data dimensions. This not only speeds up processing but also helps prevent the model from memorizing too much information (overfitting). Google Net stacks multiple inception modules together, creating a deep and accurate network architecture. Additionally, the inclusion of auxiliary classifiers at intermediate stages aids in tackling the problem of gradients vanishing during training, contributing to improved performance.

C. Open Source Platforms: TensorFlow is a powerful library for dataflow and differentiable programming [16], [17], which allows for efficient computation on a set of devices with powerful hardware's, that include memory, GPUs and TPUs. Its ability to create dataflow graphs, which describe how data moves through a computation, makes it a popular choice for ML and DL applications. In contrast, Keras is a high-end DL library that operates atop TensorFlow (also some other libraries). It simplifies the creation of DL models by providing a user-friendly API, and it provides a number of pre-built layers and functions, such as convolutional layers and pooling layers, which can be easily added to a model. In recent versions of Tensorflow (2.4 and above). TensorFlow is used to provide low level operations for building and training models, while Keras is used to provide a higher-level API for building and training models more easily. The use of TensorFlow and Keras together in this research has allowed us to effectively and efficiently solve the problem at hand.

PyTorch is also from the open-source community which has lot of capabilities for developing ML and DL applications [18], [19]. PyTorch is a powerful library for building and training DL models. It is known for its flexibility and ease of use, making it a popular choice among researchers and practitioners. One of the key features of PyTorch is its dynamic computational graph. Unlike other libraries, such as TensorFlow, which uses a static computational graph, PyTorch allows for the modification of the graph on-the-fly, making it more suitable for research and experimentation.

III. DATASET

The dataset utilized in this research, obtained from the Kaggle website, is titled "Plant Village Dataset" and "New Plant Disease Dataset", encompasses a wide array of augmented images representing diverse plant species such as tomato, potato, Corn, Maize, Apple, Raspberry, Strawberry, Grapes and many more. Among these images, various symptoms of diseased leaves are depicted, including early blight, late blight, gray mold, powdery mildew, and bacterial spot. Additionally, the dataset includes images of healthy leaves, serving as crucial reference points for training machine learning models. Through the analysis of these images, models can be trained to effectively differentiate between healthy and

diseased leaves, thereby facilitating early detection and diagnosis of plant ailments. The dataset consists of about 87K rgb images of healthy and diseasedcrop leaves which is categorized into 38 different classes.

IV. METHODOLOGY

Our project aimed to develop a system for detecting various leaf diseases using image processing and deep learning techniques. This section details the methodology employed throughout the research process.



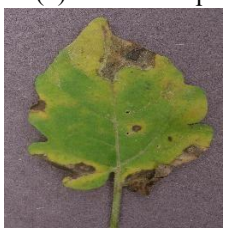
(a) Healthy Tomato



(b) Bacterial Spot



(c) Early Blind



(d) Septoria Spot

A. Data Collection: We fetched a dataset consisting of over 87k augmented imagesof various healthy as well as infected leaves. These images were categorized into 38 classes. The dataset was subsequently split into 80% training data, 10% validation data, and 10% testing data to facilitate model training and evaluation.

B. Model Selection and Training: We opted for the pre-trained VGG19 model as the foundation for our disease detection system. This model's established proficiencyin image classification tasks made it a suitable choice. The pre-trained VGG19 possesses knowledge of generic features extracted from a vast dataset (ImageNet), providing a robust base for ourspecific task of classifying tomato leaf diseases. To adapt the model to this new domain, we fine-tuned its final layers through retraining on our curated dataset of augmented tomato leaf images. This processenabled the model to refine its previously learned features and adapt them to the specific task of plant disease classification. Techniques like gradient descent and backpropagation were employed during the training phase to optimize the model's performance and minimize classification errors on the training data.

C. Hyperparameter Tuning : We performed hyperparameter tuning to optimize the performanceof our model. This iterative process involved training and evaluating the model on the validation data while adjusting various parameters. The parameters we tuned included:

1) Batch size: We utilized a batch size of 16. This refers tothe number of images processed by the model during each training iteration.
2) Optimizer: The Adam optimizer was chosen due to its efficiency and effectiveness observed in various deep learning ap- plications.
3) Learning rate: We employed a learning rate of 0.01 to control the pace of learning and prevent the model from overfittingor underfitting the training data.

B. Model Evaluation

To comprehensively evaluate the model's performance, we employed various metrics beyond accuracy. We utilized accuracy, precision, recall, and F1 score by examining the confusion matrix. This matrix provides valuable insights into the model's ability to correctlyclassify each disease class, allowing us to identify potential areasfor improvement and gain a deeper understanding of the model's strengths and weaknesses.

V. RESULTS AND DISCUSSIONS

A. Evaluation Metrics

1) Confusion Matrix: A confusion matrix is a matrix that summarizes the performance of a machine learning model on a setof test data. It is a means of displaying the number of accurate and inaccurate instances based on the model's predictions. It is often usedto measure the performance of classification models, which aim to predict a categorical label for each input instance. The matrix displays the number of instances produced by the model on the test data; True positives (TP): occur when the model accurately predicts a positive data point. True negatives (TN): occur when the model accurately predicts a negative data point. False positives (FP): occur when the model predicts a positive data point incorrectly. False negatives (FN):occur when the model mispredicts a negative data point.
2) F1 Score: F1 score is an alternative machine learning evalua- tion metric that assesses the predictive skill of a model by elaboratingon its class-wise performance rather than an overall performanceas done by accuracy. F1 score combines two competing metrics- precision and recall scores of a model, leading to its widespreaduse in recent literature.
3) Precision: Precision is defined as the ratio of correctly classified positive samples (True Positive) to a total number of classified positive samples (either correctly or incorrectly). Precision = TP/TP + FP
4) Recall: The recall is calculated as the ratio between the numbers of Positive samples correctly classified as Positive to the total number of Positive samples. The recall measures the model's ability to detect positive samples. The higher the recall, the more positive samples detected. Recall = True Positive/True Positive + False Negative
5) Accuracy: Accuracy is a metric that measures how often a machine learning model correctly predicts the outcome. You can calculate accuracy by dividing the number of correct predictions by the total number of predictions. Accuracy = TP + TN / TP + TN + FP + FN

B. Figures

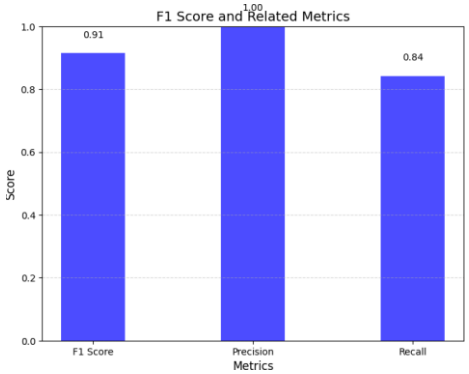


Fig. 1. F1 score, Precision, Recall

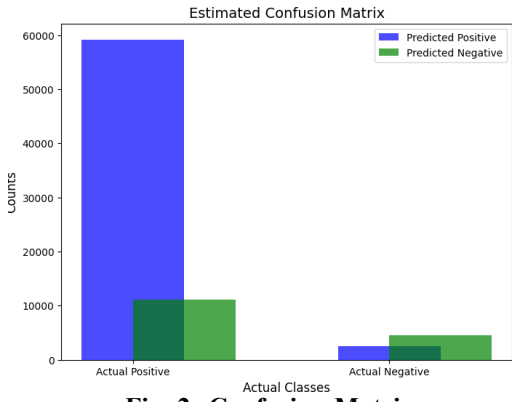


Fig. 2. Confusion Matrix

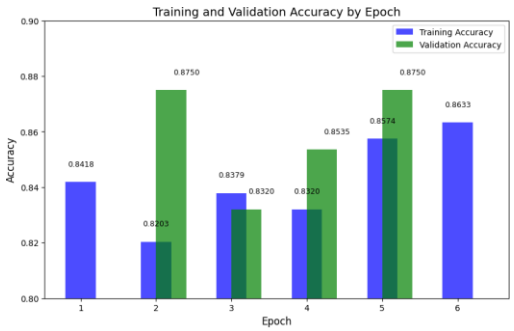


Fig. 3. Accuracy

VI. LIMITATIONS

A. **Hardware Requirements:** Deep learning models, especially large ones like VGG19, require significant computational resources for training and inference. Deploying such models in resource-constrained environments, such as on embedded systems or mobile devices, may be challenging due to memory and processing power constraints.

B. **Imbalanced Classes:** Plant disease datasets often suffer from class imbalance, where certain diseases may have many more examples than others. Thiscan lead to biased model predictions, where the model may be more accurate at detecting prevalent diseases but less accurate for rare diseases.

C. **Transferability to Other Architectures:** While VGG19 is a powerful architecture, there's no guarantee thatit will be the best choice for your specific task. Experimentationwith other architectures, such as ResNet, Inception, or EfficientNet, may be necessary to find the optimal model for your dataset and requirements.

D. **Need for Regular Updates:** Plant diseases and their manifestations can evolve over time due tovarious factors such as climate change, introduction of new cultivars,and emergence of new pathogens. Therefore, the model may need regular updates and retraining to maintain its effec*tiveness in disease detection.*

VII. CONCLUSION

In conclusion, our study has explored the efficacy of the existing system's feature-based approach, which relies on image processing techniques to detect and classify leaf diseases. This method encompasses crucial stages such as image acquisition, preprocessing, segmentation, feature extraction, and classification. Through this systematic approach, we have aimed to develop a reliable systemfor identifying and categorizing plant diseases.

In our proposed system, we employed Convolutional Neural Net- works (CNN) to accomplish leaf disease detection. The CNN model demonstrated remarkable performance, achieving an accuracy rate of 88.67

Looking forward, there are several avenues for enhancing our system. One significant area is the expansion of the dataset to include more classes of leaves and disease types. By incorporatinga broader range of plant species and diseases, we can improve the system's ability to recognize and classify various issues affecting plants. Additionally, we can explore more advanced deep learning architectures and techniques to further boost the system's accuracy and robustness.

In summary, our research demonstrates the potential of CNN-basedapproaches for leaf disease detection. By continuing to refine our models and expand our dataset, we can contribute to the developmentof more effective and versatile solutions for addressing plant health issues and promoting sustainable agriculture practices.

REFERENCES

[1] Tirkey D, Singh KK, Tripathi S. Performance analysis of AI-basedsolutions for crop disease identification detection, and classification. Smart Agric Technol. 2023. https://doi.org/10.1016/j.atech.2023.100238.

[2] Ramanjot, et al. Plant disease detection and classification: a systematic literature review". Sensors. 2023. https://doi.org/10.3390/s23104769.

[3] Mohanty, S. P., Hughes, D. P., Salathé, M. (2016). Using deep learning for image-based plant disease detection. Front. Plant Sci. 7 (September).

[4] Anjna, Sood, M. Singh, P. K. (2020). Hybrid system for detec- tion and classification of plant disease using qualitative texture fea-tures analysis. Proc. Computer Sci. 167 (2019), 1056–1065. doi: 10.1016/j.procs.2020.03.404

[5] Genaev, M. A., Skolotneva, E. S., Gultyaeva, E. I., Orlova, E. A., Bechtold, N. P., Afonnikov, D. A.. (2021). Image-based wheat fungi diseases identification by deep learning. Plants 10 (8), 1–21. doi: 10.3390/ plants10081500

[6] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F. E. (2017). Neurocomputing a survey of deep neural network architectures and their applications. Neurocomputing 234 (November 2016), 11–26. doi: 10.1016/j.neucom.2016.12.038

[7] Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., Menaka, R. (2020). Attention embedded residual CNN for disease detection in tomato leaves. Appl. Soft Comput. 86, 105933. doi: 10.1016/j.asoc.2019.105933

[8] Singh, V., Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. Inf. Process. Agric. 4 (1), 41–49. doi: 10.1016/j.inpa.2016.10.005

[9] Khan, R. U., Khan, K., Albattah, W., Qamar, A. M. (2021). Image-baseddetection of plant diseases: from classical machine learning to deep learning journey. Wireless Commun. Mobile Computing 2021, 1–13. doi: 10.1155/2021/5541859

[10] Ullah, A., Muhammad, K., Haq, I. U., Baik, S. W. (2019). Actionrecognition using optimized deep autoencoder and CNN for surveillance data streams of non-stationary environments. Future Generation Comput. Syst. 96, 386–397. doi: 10.1016/j.future.2019.01.029

[11] Muhammad Shoaib, Babar Shah, Shaker EI-Sappagh, Akhtar Ali, Asad Ullah, Fayadh Alenezi, Tsanko Gechev, Tariq Hussain, Farman Ali (2023). An advanced deep learning models-based plant disease detection: A review of recent research

[12] Hasan, R. I., Yusuf, S. M., Alzubaidi, L. (2020). Review of the stateof the art of deep learning for plant Diseases: A broad analysis and discussion. Plants 9, 1–25. doi: 10.3390/plants 9101302.

❑ ❑ ❑ ❑ ❑ ❑

AET 010

Blockchain-based Copyright Detection System: A Decentralized Approach to Protecting Intellectual Property

Karthi Devendra
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
karthi.devendra@vit.edu.in

Tanuj Ghag
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
tanuj.ghag@vit.edu.in

Hrishikesh Shinde
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
hrishikesh.shinde@vit.edu.in

Purvesh Kanade
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
purvesh.kanade@vit.edu.in

Prof. Divya Nimbalkar
Vidyalankar Institute of Technology
University of Mumbai
Mumbai, India
divya.nimbalkar@vit.edu.in

***Abstract*—In an increasingly digital world, the protection of intellectual property rights has become a paramount concern for creators and content owners. Traditional copyright detection systems often struggle to keep pace with the rapid dissemination of digital content and the challenges posed by piracy. This paper presents a novel solution: a blockchain-based copyright detection system (BCDS) designed to address these challenges.**

The proposed BCDS leverages blockchain technology's core features, including decentralization, transparency, immutability, and cryptographic security, to establish a trustless and efficient copyright detection ecosystem. Through a decentralized ledger, smart contracts, and decentralized storage, BCDS offers a seam- less and tamper-resistant platform for registering, tracking, and enforcing copyright ownership.

I. Introduction

The paper explores the implementation and potential benefits of a blockchain-based copyright detection system to address the challenges associated with protecting intellectual property in the digital age. With the advent of the internet and digital media, content creators face significant hurdles in safeguarding their creative works from unauthorized dis- tribution and infringement. Traditional copyright management systems are often centralized, cumbersome, and susceptible to vulnerabilities.

In response to these challenges, this study proposes a de-centralized Blockchain based copyright management system. The System leverages the transparency, immutability, and security features inherent in blockchain technology to create an immutable record of copyright ownership and usage rights, ensuring a secure and unalterable ledger.

Key aspects of this research paper include:

1. **Blockchain Technology:** An comprehensive gist of technology related to blockchain and its suitability for copyright management, highlighting its decentralized nature, crypto- graphic security, and transparency.
2. **Smart Contracts:** Explanation of how smart contracts are utilized to facilitate automatic copyright enforcement, including royalty distribution, licensing, and content access control.
3. **Copyright Ownership Verification:** A discussion on how blockchain enables efficient and immutable tracking of copy- right ownership, reducing disputes and ambiguities surrounding intellectual property.
4. **Transparency and Trust:** Emphasis on the role of blockchain in establishing trust and transparency within the copyright ecosystem, benefiting both creators and consumers.
5. **Case Studies:** Real-world examples and case studies of blockchain-based copyright management systems in operation, highlighting their successes, challenges, and potential for industry-wide adoption.
6. **Challenges and Future Prospects:** A critical examination of the current limitations and challenges facing blockchain- based copyright management systems, along with suggestions for future research and improvements.

In the ever-evolving landscape of digital content creation and dissemination, protecting intellectual property rights has become an increasingly complex and critical challenge. The rise of the internet and digital technologies has empowered creators to share their works with global audiences more easily than ever before. However, this newfound accessibility has also given rise to issues related to copyright infringement, piracy, and the fair compensation of creators. Traditional copyright management systems, often centralized and riddled with inefficiencies, are struggling to keep pace with the rapid digitization of content.

In response to these challenges, blockchain technology has become a prominent and disruptive answer for copyright management. Blockchain, the decentralized and immutable ledger underlying cryptocurrencies like Bitcoin, offers a new paradigm for managing and protecting intellectual property rights. By leveraging the core attributes of blockchain, such as transparency, security, and decentralization, a copyright management system based on blockchain technology can address several shortcomings of conventional approaches.

II. Related Work

The following examines the existing literature pertaining to blockchain technology and its utilization within the realm of copyright.

A. Preliminary Introduction to blockchain

In this digital age, data and information have become

some of the most valuable assets, and the need for secure, transparent, and tamper-proof systems to manage this data has never been more apparent. The formulation of blockchain technology was initially limited to cryptocurrencies like Bitcoin. Even though it was a groundbreaking technology it misses revolutionizing different domains other than finance.
The concept of blockchain can be likened to a digital ledger or a chain of interconnected blocks, where each block contains a set of transactions or data entries. What sets blockchain apart is its unique structure and characteristics:

1. **Decentralization:** In contrast to conventional centralized systems reliant on a singular point of control, blockchain operates across a decentralized network of computers, often termed nodes. These nodes collaborate to authenticate and log transactions. This decentralized framework eradicates the necessity for intermediaries like banks or clearinghouses, bolstering transaction efficiency and mitigating susceptibility to manipulation.
2. **Immutability:** Once data is inscribed onto a blockchain, it becomes virtually unchangeable. Every block in the chain references the prior one, establishing an irreversible and chronological record. This inherent immutability safeguards data integrity and presents formidable barriers against unauthorized tampering.
3. **Transparency:** Blockchain transactions are openly accessible to all network participants. Though user identities might remain pseudonymous, the transaction history remains visible and verifiable. This transparency fosters trust among participants and diminishes opportunities for fraudulent activities.
4. **Security:** Blockchain relies on cryptographic methodologies to secure data and verify transactions. Consensus mechanisms like Proof of Work and Proof of Stake ascertain that only legitimate transactions enter the ledger. The amalgamation of cryptographic security and consensus mechanisms fortifies blockchain against illicit access and fraudulent endeavors.
5. **Smart Contracts:** Blockchain enables the execution of self-executing contracts, termed smart contracts. These programmable scripts automatically trigger predetermined actions upon meeting specific conditions. Smart contracts facilitate automation within a trustless environment, diminishing reliance on intermediaries across various domains, from supply chain management to financial services.

   The potential applications of blockchain technology are expansive, extending into domains such as finance, healthcare, supply chain management, voting systems, and more. Its capacity to disrupt established industries by augmenting transparency, curbing expenses, and heightening security renders blockchain a subject of considerable interest and ongoing exploration.

B. Traditional Copyright Detection System

Traditional copyright detection systems have been the cornerstone of intellectual property protection for decades. These systems rely on centralized authorities, such as copyright offices and collecting societies, to manage copyright registrations, licensing, and royalty distribution. While effective to some extent, they suffer from issues related to transparency, efficiency, and susceptibility to human errors and fraud. Various studies have explored the limitations of these system and the need for more efficient, automated, and transparent alternatives

C. Blockchain Technology in Copyright Management

Blockchain technology has gained significant attention for its potential to revolutionize copyright management. Early research has laid the foundation for using blockchain to establish immutable records of copyright ownership. Smart contracts, a key feature of blockchain, have been proposed as a means to automate licensing agreements, ensuring that creators receive compensation for their work.

D. Intellectual Property and Blockchain Consortium

Several consortiums and industry groups have emerged to explore the intersection of blockchain and intellectual property rights. Notable initiatives include the Open Music Initiative (OMI) and the Copyright Hub. These initiatives aim to standardize data formats and processes for copyright management on the blockchain, fostering interoperability and wider adoption within the creative industries.

E. Challenges and Limitations

While the promise of blockchain in copyright management is substantial, researchers have not ignored the associated challenges. Scalability issues, energy consumption concerns in some blockchain implementations, and legal complexities related to jurisdiction and enforcement are among the challenges discussed in the literature.

F. How is Blockchain Based Copyright Detection System Different from NFTs

While there may be some similarities between BCDS and Non-Fungible Tokens (NFTs), they are fundamentally different concepts with distinct purposes and functionalities:

1. Blockchain-Based Copyright Detection System (BCDS):

- **Purpose:** BCDS is designed primarily for the detection, tracking, and enforcement of copyright ownership and usage rights for digital content. It leverages blockchain technology's features like decentralization, transparency, immutability, and smart contracts to establish a trustless and efficient copyright management ecosystem.
- **Functionality:** BCDS aims to provide a tamper-resistant platform for registering, tracking, and enforcing copyright ownership. It focuses on ensuring that content creators receive fair compensation for their work and that consumers can access content in compliance with copyright regulations.
- **Use Cases:** BCDS is used to protect intellectual property rights, prevent copyright infringement, and streamline copyright-related transactions in various digital content industries.

2. Non-Fungible Tokens (NFTs):

- **Purpose:** NFTs define the cryptographic tokens which show ownership of unique assets which may be present in physical or digital form like artistic models, music, collections, virtual real estate, and more. They are used to prove the authenticity, ownership, and provenance of these assets.

- **Functionality**: NFTs are created, bought, and sold on blockchain platforms, often using smart contracts. They establish a firm and clear way so that ownership can be established to transfer the unique digital assets. NFTs can be linked to digital content, but they do not inherently enforce copyright or intellectual property rights.
- **Use Cases:** NFTs have gained popularity in various fields present overall like digital art, gaming, and entertainment. They enable artists and creators to tokenize their work, sell it to collectors, and receive royalties from secondary sales.
  In summary, BCDS is primarily focused on copyright protection and management, whereas NFTs are a broader concept related to the ownership and provenance of digital or physical assets. While both use blockchain technology, they serve different purposes within the digital content ecosystem. BCDS aims to enforce copyright, while NFTs are more about proving ownership and uniqueness of digital assets.

III. DESIGN

A. Copyright

Copyright means the important and intelligent property which has the duty for the protection of an individual's private work as soon as the task is completed in a palpable means of expression. The work can be divided into various categories like any type of digital architecture, books, videos, images, recordings, software designs, and important system codes. This phrase of "original work of authorship" means that the task is completed and created by the genuine writer/author singly having a standard of creativity, banning its subject, name, expressions, and shades. Copyrighting doesn't cover any conception, principle, and administration of content creation; instead, it helps in the protection from the effects in the way they are addressed. Copyrighting preserves the content format but it doesn't preserve the details of the contentment.

B. P2P Network

A peer-to-peer (P2P) network within blockchain technology is a foundational component that drives the decentralized nature of blockchain systems. Unlike traditional centralized networks, where a single authority controls data and transactions, a P2P network in blockchain operates on a distributed model where all participants, or nodes, interact directly with one another. This decentralization is a core principle of blockchain technology and brings about several fundamental characteristics that underpin its functionality and security. One of the primary purposes of a P2P network in blockchain is to enable consensus among participants. Consensus mechanisms are employed to guarantee that every node concurs with the state of the blockchain and verifies transactions. This decentralized validation roadmap is critical for keeping the uprightness of the ledger and preventing fraudulent activities. Security is a paramount concern in blockchain, and P2P networks contribute significantly to its robustness. Due to the decentralized architecture, tampering with data on one node does not compromise the security or reliability of the network as a whole. The majority of nodes must reach a consensus for any changes to be accepted, providing a high level of security against malicious activities. Scalability is another advantage of P2P networks in blockchain. To accommodate a growing number of users and transactions, blockchain networks can scale horizontally by adding more nodes to the network. This scalability is crucial for the widespread adoption and use of blockchain technology across various industries.

C. Proof of Work (PoW)

Proof of Work is a mechanism used in technology to verify and secure transactions and maintain the uprightness of the ledger. It serves as a method for achieving harmony among network participants in a decentralized and trustless manner.
In a PoW system, network participants, known as miners, compete to find an answer to some complex puzzles. The tasks are computationally intensive and require significant computational power to solve. Miners use their computing resources, such as specialized hardware which are very efficient in finding accurate answers to attempt to solve these puzzles.
The very first miner to successfully solve the puzzle distributes their solution in the entire network, while also adding a new fresh block of transactions to the blockchain network. The remaining nodes present in the network then verify the justifiability of the solution and the transactions of the new block. Whether the solution is correct and transactions are valid, the next successive block is then added to the ledger. This process is "mining".
Despite its advantages, PoW has been criticized for its high energy consumption & potential for centralization as mining becomes increasingly competitive & resource intensive.

D. Cryptographic Primitives

Cryptographic primitives are essential components of blockchain technology, providing the foundation for security, privacy, and trust within the network. These cryptographic techniques are used to achieve various objectives in blockchain, including data integrity, authentication, confidentiality, and consensus. Here are some of the key cryptographic primitives commonly employed in blockchain systems:

1. **Hash Functions:** These are some methods that play a central criteria in blockchain technology. They are used to transform data (like transactions or blocks) into fixed-size strings of characters called hash values. Hash functions in blockchain provide data integrity and ensure that any modification to the real data points results in a completely dissimilar hash value. This property is crucial for verifying the contents of blocks and transactions.
2. **Digital Signatures:** Digital signatures enable participants in a blockchain network to prove their identity and authenticate their transactions. When the user generates a transaction, they validate it with their private key. The transaction's authenticity can be verified by others using the sender's public key. If the signature is valid, it serves as proof that the transaction was indeed produced by the owner of the private key.
3. **Public Key Cryptography:** It is also known as asymmetric cryptography. Which involves the use of key pairs (i.e. public and private keys). The public key is employed for encryption and verifying digital signatures, whereas the private key is kept confidential and used for decryption and signing. Public key cryptography ensures confidentiality, authenticity, & data integrity in blockchain transactions.

4. **Merkle Trees:** Merkle trees are a hierarchical data structure that plays a significant role in blockchain's efficiency and security. They are used to summarize large amounts of data (e.g., transactions) into a compact, fixed-size hash, known as the Merkle root. This root is included in a block's header and allows for efficient verification of individual transactions within a block.
5. **Symmetric Encryption:** Although less common in blockchain, symmetric encryption may be used in some cases to secure data at rest or in transit. It involves a single key shared between parties for both encryption and decryption. However, managing and securely exchanging these keys can be challenging in decentralized blockchain networks.
6. **Zero-Knowledge Proofs:** Zero-knowledge proofs, like zk-SNARKs (Zero-Knowledge Succinct Non-Interactive Argument of Knowledge), are advanced cryptographic techniques used to prove the validity of a statement without revealing either specific information about the statement itself, thereby enhancing privacy in blockchain transactions by allowing parties to prove they have knowledge of certain data without disclosing the data itself.

   These cryptographic primitives collectively contribute to the security, transparency, and trustworthiness of blockchain systems. They are crucial tools for safeguarding data, verifying transactions, and ensuring the integrity of the blockchain ledger. As blockchain technology continues to evolve, cryptographic techniques will remain a cornerstone of its design and functionality.

IV. IMPLEMENTATION

The system architecture consists of multiple components, including a blockchain network, a media storage system, and a verification mechanism. The blockchain network serves as the underlying infrastructure, ensuring the immutability and integrity of media files. The media storage system securely stores the media files, while the verification mechanism allows users to verify the authenticity and uniqueness of published media.

A. Blockchain Implementation

The blockchain implementation utilizes a custom blockchain framework that supports the creation of new blocks, mining of transactions, and consensus mechanisms. Each block is having a hash value of the previous block, ensuring the integrity and unchangeable nature of the entire chain. The system also incorporates a proof-of-work consensus algorithm to prevent malicious actors from tampering with the blockchain.

B. Media Publishing Process

The media publishing process involves uploading media files to the system, which are then securely stored in the media storage system. Each uploaded file is assigned a unique identifier and undergoes a lookup process to detect and prevent the publication of duplicate or similar media. Once verified, the media file is added to the blockchain as a new transaction, ensuring its authenticity and traceability.

C. Media Verification Process

The media verification process allows users to verify the authenticity and uniqueness of published media files. By querying the blockchain, users can retrieve the transaction history of a specific media file, including information about the author, title, and genre. This process ensures transparency and trust in the media publishing ecosystem.

1. **Image Verification:** In case if the media type chosen is 'image', then the following process takes place. Initially the images are opened and loaded in the program using necessary libraries. The program then computes the perceptual hash of each image. The perceptual hash is a compact representation of the image that captures its visual features. Generation of perceptual hash takes place in following steps:

- Image Resizing: The input image is resized to a small, consistent size, typically to a smaller resolution, to simplify the hashing process. The resizing helps to reduce the complexity and computational load of subsequent operations.
- Grayscale Conversion: The resized image is converted into a grayscale representation. This step ensures that the hash calculation focuses on luminance values rather than color, which can improve the robustness of the hash against color variations.
- Pixel Intensity Calculation: The average pixel intensity is calculated for the resized grayscale image. This average value serves as a threshold for the creation of the hash.
- Hash Generation: For each pixel in the resized and grayscale image, the algorithm compares the pixel's intensity with the calculated average intensity. If the pixel's intensity is greater than the average, it's represented as '1' in the hash. Conversely, if the intensity of pixel is ¡= to the average, it's represented as '0' in the hash.

  After obtaining the perceptual hashes, the program calculates the Hamming distance between them. The Hamming distance is a measure of the number of differing bits between two strings. In our case, it represents the dissimilarity between the two images. Finally, if the calculated Hamming distance is less than given threshold, the program notifies the author.

2. **Audio Verification:** The audio data is taken as input and a fingerprint is generated based on several parameters, including the sample rate, number of channels, an iterable containing blocks of PCM (Pulse Code Modulation) data, and an optional maximum audio length. The fingerprint generation process typically involves several steps:

- **Audio Data Preparation:** The input audio data is typically in PCM (Pulse Code Modulation) format, which represents the raw audio waveform as a series of amplitude samples. Parameters like sample rate (e.g., 44100 Hz), number of channels (mono or stereo), and aniterable containing blocks of PCM data are provided.
- **Preprocessing and Frame Blocking:** The audio data might undergo preprocessing steps, which could include normalization, filtering, or resampling to ensure consistency in format and quality. The audio is divided into smaller frames or blocks of fixed duration (e.g., 1-2 seconds).

- **Feature Extraction:** For each block of PCM data within the specified duration, features are extracted to capture essential information about the audio. This can involve techniques such as computing spectrograms, calculating frequency-domain representations (e.g., using FFT - Fast Fourier Transform), or other signal processing methods.
- Hashing or Encoding: The extracted features from each block of audio data are encoded or hashed to create a compact representation. Hashing typically involves applying a hash function to the extracted features to generate a fixed-size hash or code that represents the characteristics of that specific audio segment.
- Optional Maximum Audio Length: In some systems, an optional maximum audio length parameter might be used to limit the duration of the audio being fingerprinted. This parameter helps in processing large audio files efficiently without overwhelming the system resources.

Overall, the process of generating audio fingerprints involves breaking the audio into smaller segments, extracting relevant features from each segment, encoding these features into compact representations (hashes), and potentially aggregating them to create a unique fingerprint that represents the audio track.

These fingerprints are then compared against the ledger in blockchain for audio identification purposes. If a match is found, the user is notified about the same.

3. **Text Verification:** Initially the TLSH hash of file is generated. TLSH (Trend Micro Locality-Sensitive Hashing) is an algorithm used for generating a hash (digest) for data, particularly for similarity comparisons between data items. It's specifically designed to calculate a hash that reflects the similarity between different datasets while providing a low collision rate for similar data. The TLSH algorithm involves several key steps:

- **Data Preprocessing:** TLSH operates on byte-level data. The input data(file), is first processed by dividing it into chunks of a fixed size.
- **Checksum Calculation:** TLSH computes checksums from the input data chunks. Checksums are calculated for various aspects of the data, including: Checksum A: Based on the numbers of characters in certain ranges (0-31, 32-63, 64-95, etc.). Checksum B: Using pairwise XOR of bytes. Checksum C: Counting the number of consecutive NULL bytes.
- **Buckets and Bucketing:** TLSH further processes the calculated checksums by assigning them to different buckets. This process aims to create a hash that is robust against certain types of changes while sensitive to structural similarities between datasets.
- **Final Hash Generation:** The generated checksums and bucket information are combined to form the final TLSH hash. The hash typically consists of a fixed-length hexadecimal string, representing the fingerprint of the input data.

TLSH focuses on producing a hash that is sensitive to minor changes in data which makes it very useful in copyright detection.

The code then calculates the difference between the TLSH hashes of the uploaded file and the files on the blockchain ledger. The lower the difference, the more similar the files are. If the difference is found to be lower than the given threshold then the uploaded file is titled as duplicate.

D. Experimental Evaluation

To evaluate the outcome and efficiency of the proposed system, we performed experiments using a dataset of media files. The experiments focused on measuring the system's throughput, latency, and resistance to attacks. The results demonstrate the system's efficiency in handling a large number of transactions while maintaining the security and authenticity of the media files.

V. Future Scope

Development of a blockchain-based copyright detection system represents a significant step forward in addressing the challenges of digital copyright protection. However, This technology is constantly advancing, and there are number of promising directions for future research and development in this field. The following are some potential directions for future work:

1. **Integration with Emerging Blockchain Platforms:** As blockchain technology continues to advance, it is essential to explore and integrate with emerging blockchain platforms. Evaluating the compatibility of the copyright detection system with newer blockchain ecosystems like Polka dot, Ethereum 2.0, or Cardano could enhance scalability and interoperability.
2. **Machine Learning and AI Integration:** Leveraging machine learning and artificial intelligence (AI) algorithms for more accurate copyright detection is a promising direction. Future research can focus on improving the system's ability to recognize copyrighted content, even in modified or derivative forms.
3. **Decentralized Content Hosting:** Extending the capabilities of the system to include decentralized content hosting can be explored. This would enable content creators to not only protect their work but also securely host it on decentralized networks, reducing reliance on centralized hosting platforms.
4. **Blockchain Consensus Mechanisms:** Exploring alternative consensus mechanisms beyond Proof of Work and Proof of Stake could enhance the efficiency and sustainability of the copyright detection system. Consensus mechanisms such as Proof of Authority and Delegated Proof of Stake could be assessed.
5. **Cross-Border Copyright Enforcement:** International copy- right enforcement remains a complex challenge. Research can delve into the development of blockchain-based tools for cross-border copyright enforcement and dispute resolution, fostering collaboration between legal jurisdictions.
6. **Legal and Regulatory Frameworks:** The legal and regulatory landscape surrounding blockchain-based copyright protection is still evolving. Researchers can engage with policy- makers and legal experts to help shape the legal framework and ensure that it aligns with technological advancements.

7. **Collaboration with Publishers and Content creators:** Collaboration with publishers, Content creators and copyright organizations is paramount. Future research should involve these stakeholders in the design and testing phases to ensure that the system meets their specific needs and concerns.

In conclusion, the development of a blockchain-based copy- right detection system holds great promise for safeguarding intellectual property rights in the digital age. Future research and development efforts should focus on enhancing technology, ensuring legal compliance, and fostering user adoption to realize the full potential of this innovative approach to copyright protection.

VI. CONCLUSION

The suggested system utilizes blockchain technology to enhance the security and dependability of the copyright process when compared to traditional copyright detection systems. This paper proposes a decentralized copyright system for cross-border copyright protection, based on a P2P blockchain using the PoW consensus algorithm. Unlike traditional copyright systems, the proposed system doesn't rely on a centralized server to store the copyrighted content. Instead, blockchain technology is used to store the data (metadata) of copyrighted content. This establishes a synchronized plat- form for investigating copyright of digital content worldwide. Also in copyright management it discards the need for an encryption-decryption process. Interested individuals can readily inquire about content to ascertain its copyright status. Hence, they can become aware of copyrighted works and will avoid unintentionally violating copyright laws. Based on these findings, we can conclude that this system is scalable and suitable for commercialization in international copyright management.

REFERENCES

[1] Yang, Wentao He, Ping and Yang, Zhen and Yi, Xiaochun and Chen, Chang. (2020). Digital Copyright Depository System Enhanced by Blockchain. 198-202. 10.1109/ICCST50977.2020.00044.

[2] Islam, Md. Mainul and IN, Hoh. (2021). Decentralized Global Copy- right System Based on Consortium Blockchain with Proof-of-Authority. 10.36227/techrxiv.16564218.v1.

[3] Jing, Nan and Liu, Qi and Sugumaran, Vijayan. (2021). A blockchain- based code copyright management system. Information Processing and Management. 58. 102518. 10.1016/j.ipm.2021.102518.

[4] Liu, Liming and Shang, Wenqian and Lin, Weiguo and Huang, Wei. (2021). A Decentralized Copyright Protection, Transaction and Content Distribution System Based on Blockchain 3.0. 45-50. 10.1109/SNPDWinter52325.2021.00018.

[5] Xiao, Xiangli and He, Xiaotong and Zhang, Yushu and Dong, Xuewen and Yang, Lu-Xing and Xiang, Yong. (2023). Blockchain-based reliable image copyright protection. IET Blockchain. 10.1049/blc2.12027.

[6] Fahim, Shahriar and Mahmood, Sharfuddin and Rahman, S M. (2023). Blockchain: A Comparative Study of Consensus Algorithms PoW, PoS, PoA, PoV. 10.5815/ijmsc.2023.03.04.

[7] Zhang, Linlu and Liu, Shuxian and Ma, Chengji and Su, Tingt- ing. (2023). Innovative Application of Blockchain Technology for Digital Recipe Copyright Protection. Applied Sciences. 13. 9803. 10.3390/app13179803.

[8] Y. Ding, L. Yang, W. Shi and X. Duan, " The Digital Copy- right Management System Based on Blockchain," 2019 IEEE 2nd International Conference on Computer and Communication Engineering Technology (CCET), Beijing, China, 2019, pp. 63-68, doi: 10.1109/CCET48361.2019.8989370.

[9] J. Guo, H. Zhou, L. Yang and X. Chen, "Research on digital copy- right blockchain technology," 2020 3rd International Conference on Smart BlockChain (SmartBlock), Zhengzhou, China, 2020, pp. 1-5, doi: 10.1109/SmartBlock52591.2020.00028.

[10] Y. Zeng, "Digital Music Resource Copyright Management Mechanism Based on Blockchain," 2020 3rd International Conference on Smart BlockChain (SmartBlock), Zhengzhou, China, 2020, pp. 158-162, doi: 10.1109/SmartBlock52591.2020.00036.

[11] R. A. Dobre, R. O. Preda, R. A. Badea, M. Stanciu and A. Brumaru, "Blockchain-Based Image Copyright Protection System using JPEG Re- sistant Digital Signature," 2020 IEEE 26th International Symposium for Design and Technology in Electronic Packaging (SIITME), Pitesti, Ro-mania, 2020, pp. 206-210, doi: 10.1109/SIITME50350.2020.9292296.

[12] Q. Chen, Y. Kong and L. Cheng, "A Digital Copyright Protection System Based on Blockchain and with Sharding Network," 2023 IEEE 10th International Conference on Cyber Security and Cloud Computing (CSCloud)/2023 IEEE 9th International Conference on Edge Computing and Scalable Cloud (EdgeCom), Xiangtan, Hunan, China, 2023, pp. 393- 398, doi: 10.1109/CSCloud-EdgeCom58631.2023.00073.

[13] V. A. Kolesov, "Approach to the Creation of a Digital Content Copy- right Protection System," 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), St. Petersburg, Moscow, Russia, 2021, pp. 2349-2355, doi: 10.1109/ElCon-Rus51938.2021.9396153.

[14] W. Liang, D. Zhang, X. Lei, M. Tang, K. -C. Li and A. Y. Zomaya, "Circuit Copyright Blockchain: Blockchain-Based Homomorphic En- cryption for IP Circuit Protection," in IEEE Transactions on Emerging Topics in Computing, vol. 9, no. 3, pp. 1410-1420, 1 July-Sept. 2021, doi: 10.1109/TETC.2020.2993032.

[15] X. Chen, A. Yang, J. Weng, Y. Tong, C. Huang and T. Li, "A Blockchain- Based Copyright Protection Scheme With Proactive Defense," in IEEE Transactions on Services Computing, vol. 16, no. 4, pp. 2316-2329, 1 July-Aug. 2023, doi: 10.1109/TSC.2023.3246476.

❑ ❑ ❑ ❑ ❑ ❑

AET 011

Allocatrix: Revolutionizing Investment Strategy Optimization

Gaurav Bhadoria
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
gaurav.bhadoria@vit.edu.in

Pratham Tetgure
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
pratham.tetgure@vit.edu.in

Harsh Purohit
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
harsh.purohit@vit.edu.in

Kedar Vaidya
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India. kedar.vaidya@vit.edu.in

Dr. Nayana Mahajan
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India. nayana.mahajan@vit.edu.in

***Abstract*– In the rapidly evolving landscape of financial markets, Allocatrix emerges as a pioneering force, seamlessly melding human expertise with cutting-edge technology to redefine investment strategy optimization. This paper explores how Allocatrix embodies a transformative paradigm shift, representing the convergence of traditional wisdom and innovative methodologies. By integrating data analytics and machine learning algorithms, Allocatrix empowers investors to navigate the intricacies of financial markets with unparalleled precision and foresight. This abstract provides an overview of Allocatrix's role as a catalyst for augmenting decision-making capabilities, distilling vast financial data into actionable insights, and heralding a new era of investment strategy optimization.**

I. INTRODUCTION

In the ever-changing landscape of financial markets, marked by escalating complexities and the relentless influx of data, the necessity for shrewd decision-making stands as an imperative. Enter Allocatrix, a trailblazing entity that seamlessly blends the profound insights derived from human expertise with the cutting-edge capabilities offered by technological advancements. Within the domain of investment strategy optimization, Allocatrix embodies a revolutionary paradigm shift, symbolizing the fusion of traditional wisdom and pioneering methodologies. At its essence, Allocatrix serves as a testament to the harmonious coalescence of human intellect and technological acumen. In essence, it acts as a catalyst for enhancing the decision-making prowess of investors, presenting an intricate framework that transcends conventional methods. Through the strategic amalgamation of data analytics and machine learning algorithms, Allocatrix empowers investors to navigate the complexities of financial markets with unparalleled precision and foresight.

At the heart of Allocatrix's effectiveness lies its capacity to distill extensive volumes of financial data into actionable insights, enabling investors to make well-informed decisions in real-time. Leveraging advanced algorithms and predictive analytics, Allocatrix sheds light on market trends, identifies nascent opportunities, and assesses risk factors with unprecedented accuracy. Furthermore, by harnessing historical data and deploying sophisticated forecasting techniques, Allocatrix equips investors with the ability to anticipate market shifts and strategically position their portfolios for optimal outcomes.

Beyond its technological prowess, Allocatrix serves as a beacon of innovation, ushering in a new era of investment strategy optimization. By amalgamating the collective wisdom of human experts with the computational capabilities of machine learning algorithms, Allocatrix transcends the constraints of traditional methodologies, offering a dynamic and adaptive framework that evolves in tandem with the ever-fluctuating dynamics of financial markets.

In essence, Allocatrix epitomizes the synergy between human ingenuity and technological innovation, propelling investors towards unprecedented realms of success in the realm of financial markets.

II. EASE OF USE

Allocatrix prioritizes user experience, ensuring that its advanced capabilities are accessible and intuitive for investors of all levels of expertise. The platform features a user-friendly interface designed to streamline the investment process, allowing users to navigate seamlessly through various tools and functionalities.

Allocatrix offers a visually appealing and intuitive dashboard that provides users with a comprehensive overview of their investment portfolios, market trends, and performance metrics. Key information is presented in a clear and organized manner, enabling users to make informed decisions at a glance.

To cater to individual preferences and investment goals, Allocatrix allows users to customize their settings and preferences. Whether adjusting risk tolerance levels, setting investment objectives, or specifying asset allocation preferences, users have the flexibility to tailor the platform to their specific needs.

Allocatrix offers a range of interactive tools and features to facilitate investment decision-making. From interactive charts and graphs to real-time market data feeds, users have access to the tools they need to conduct thorough analysis and research.

Recognizing the importance of ongoing learning and skill development, Allocatrix provides users with access to educational resources and tutorials. Whether it's learning how to interpret financial ratios or understanding the principles of portfolio optimization, users can enhance their knowledge and skills within the platform.

In the event of any questions or issues, Allocatrix offers responsive customer support to assist users. Whether through live chat, email, or phone support, users can rely on timely assistance to address their concerns and optimize their experience on the platform. Overall, Allocatrix prioritizes ease of use to empower users with the tools, resources, and support they need to navigate the complexities of financial markets confidently and effectively.

III. REVIEW OF LITERATURE

By synthesizing and critically analyzing the existing body of literature, this review aims to provide a solid foundation for understanding the context, significance, and implications of the current study. Through a systematic examination of prior research and theoretical frameworks, we endeavor to identify gaps, inconsistencies, and opportunities for further exploration, thereby contributing to the advancement of knowledge in the field.

Here are several articles related to our project:
"What Happened To The Quants In August 2007?"
Authors: Amir E. Khandani and Andrew W. Lo
Published: September 20, 2007

This paper provides an in-depth analysis of the significant events that transpired during the pivotal week of August 6th, a period that sent shockwaves through the hedge fund industry. The study examines the unprecedented losses experienced by numerous quantitative hedge funds, attributing them largely to reliance on long/short equity strategies, particularly involving short-selling practices. By shedding light on the systemic risks inherent in the quantitative fund industry, this research raises broader questions about the stability and resilience of quantitative investment strategies in the face of evolving market dynamics.

"The Cross-Section of Expected Stock Returns"
Authors: Eugene F. Fama and Kenneth R. French
Published: June 1992

Fama and French's (1992) three-factor model, which includes market risk, size, and value, laid the foundation for understanding stock returns beyond market risk alone. Subsequent research has further refined our understanding, with factors like momentum, profitability, investment, and low volatility being explored. Recent advancements in empirical methodologies, including machine learning, have uncovered new patterns in stock returns, emphasizing the multidimensional nature of stock risks.

"5 Factor Asset Pricing Model"
Authors: Eugene F. Fama and Kenneth R. French
Published: April 2015

This paper introduces a five-factor model aimed at comprehensively capturing size, value, profitability, and investment patterns in average stock returns, advancing beyond the three-factor model proposed in 1993. The new model outperforms its predecessor by providing a more nuanced understanding of factors driving stock returns. However, challenges remain in capturing low average returns observed in small stocks, highlighting the need for further refinement in modeling approaches. These articles collectively contribute to our understanding of investment strategies, market dynamics, and the challenges and opportunities inherent in quantitative finance. Through critical analysis and synthesis of these works, we aim to inform and guide our own research efforts in this field.

IV. METHODOLOGY

The flowchart describes the flow of the system.

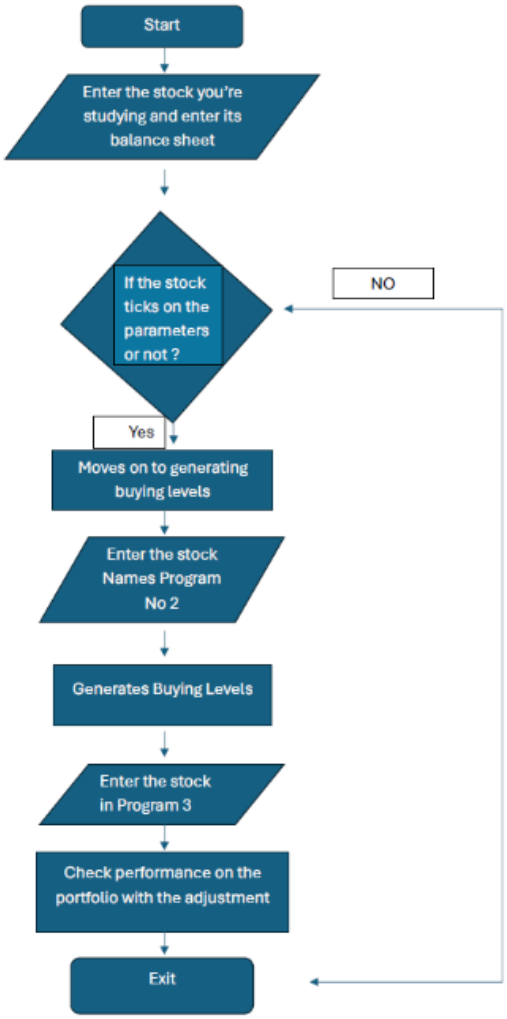


Figure 1: Block Diagram for Allocatrix

As mentioned in Figure 1, the sequential steps involved in the program's operation. The program initiates its execution from the "Start" point, signaling the beginning of the process. Users interact with the program through the "User Input" block, where they provide inputs such as the stock they wish to analyze and key financial ratios. The "Evaluation" block conducts a comprehensive analysis, including fundamental and optional technical analyses, to assess the selected stock's viability.

Based on this analysis, the program determines whether the stock meets its criteria. If the stock passes the evaluation, the program proceeds to provide recommendations in the "Generate Buying Levels" block, suggesting potential buy-in prices and allocation percentages for the stock purchase. However, if the stock fails to meet the criteria, the program may either reject it and exit or offer alternative stock suggestions.

Following the recommendation generation, the program may conduct a "Performance Check," which involves tasks such as portfolio integration and optional back testing. Portfolio integration allows users to monitor how the stock's performance affects their overall investment portfolio, while back testing simulates past performance to evaluate recommendation effectiveness.

Finally, the program concludes its operation at the "Exit" block, terminating its execution. This structured overview

delineates the sequence of actions and decisions involved in the stock analysis and recommendation process, providing a clear understanding of the program's workflow.

V. EXPERIMENTATION LSTM INVESTMENT MODEL

The LSTM (Long Short-Term Memory) Model, a variant of recurrent neural networks (RNNs), stands out for its ability to process sequential data while capturing long-term dependencies. In finance, LSTM finds applications in time-series analysis, where it excels at uncovering complex patterns and relationships in data such as stock prices, exchange rates, and economic indicators.

At its core, the LSTM architecture comprises memory cells equipped with gates—input, output, and forget gates. These gates regulate the flow of information within the network, allowing it to selectively update and access stored information over time. For instance, the input gate controls the flow of new information into the memory cell, while the forget gate determines which information to discard from the cell's memory.

In a real-life example, consider a financial analyst using LSTM to predict stock prices. The analyst feeds historical stock price data into the LSTM model, which then learns to identify patterns and trends in the data. By analyzing sequential data points over time, LSTM can capture complex relationships between various factors influencing stock prices, such as market trends, company performance, and macroeconomic indicators.

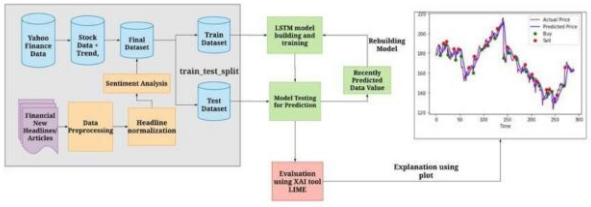


Figure 2: LSTM Model

Once trained, the LSTM model can generate predictions for future stock prices based on new input data. For instance, if the model detects a pattern indicating a potential upward trend in a company's stock price, it may suggest buying the stock in anticipation of future gains. Conversely, if the model identifies a pattern indicative of a downward trend, it may recommend selling the stock to avoid potential losses.

In this way, LSTM empowers investors and financial institutions to make informed investment decisions by providing insights into future market movements and trends. By leveraging historical data and capturing temporal dependencies, LSTM enhances the accuracy and reliability of investment predictions, ultimately contributing to better portfolio management and risk mitigation strategies.

VI. MARKOWITZ INVESTMENT MODEL

The Markowitz Investment Model, developed by Harry Markowitz in the 1950s, revolutionized portfolio theory by introducing the concept of diversification as a means of optimizing risk adjusted returns. At its core, the model aims to construct efficient portfolios that offer the highest expected return for a given level of risk or the lowest risk for a given level of return.

The model starts by identifying the expected returns, volatilities, and correlations of various assets within the investment universe. These inputs are crucial for calculating the risk and return characteristics of different portfolio allocations. Additionally, investor risk preferences, often captured by utility functions, play a significant role in shaping portfolio construction.

In a real-life scenario, imagine an investor seeking to construct a diversified investment portfolio using the Markowitz Model. The investor begins by analyzing the historical performance of different asset classes, such as stocks, bonds, and real estate. Based on their risk tolerance and investment objectives, the investor selects a combination of assets that maximizes expected return while minimizing overall portfolio risk.

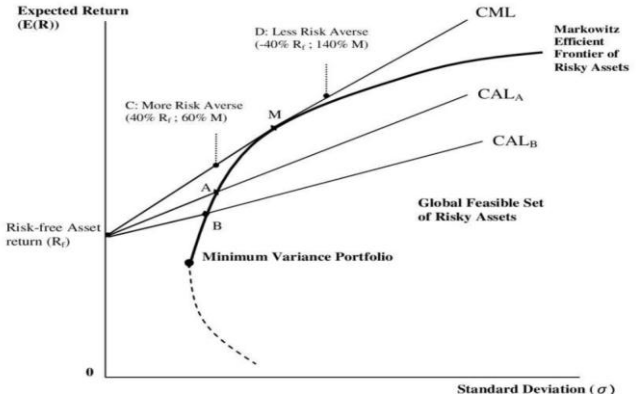


Figure 3: MARKOWITZ Model

For example, suppose the investor has a moderate risk tolerance and aims to achieve a balance between growth and stability. In this case, the Markowitz Model may recommend a portfolio allocation consisting of a mix of stocks and bonds, with the proportion of each asset class determined by their expected returns, volatilities, and correlations.

Once the portfolio allocation is determined, the investor monitors its performance over time and adjusts as needed to maintain the desired risk-return profile. By systematically weighing the risk-return trade-offs of different asset combinations, the Markowitz Model provides a disciplined approach to portfolio management that aims to optimize returns while managing risk effectively.

In summary, the Markowitz Investment Model offers a systematic framework for constructing diversified investment portfolios tailored to investor preferences and objectives. By considering the interplay between risk and return across different asset classes, the model empowers investors to make informed decisions that align with their financial goals and risk tolerance levels.

Summary of LSTM and Markowitz Investment Models:

Both the LSTM (Long Short-Term Memory) Investment Model and the Markowitz Investment Model represent powerful tools in the realm of finance, each offering unique insights and methodologies for optimizing investment strategies. The LSTM Investment Model, a form of recurrent neural network (RNN), excels at processing sequential data while capturing long-term dependencies. In finance, LSTM finds extensive applications in time-series analysis, where it uncovers complex patterns and relationships in data such as stock prices, exchange rates, and economic indicators. At its core, LSTM comprises memory cells equipped with gates—input, output, and forget gates—that regulate the flow of information within the network. By selectively updating and accessing stored information over time, LSTM learns to identify subtle trends and nonlinear relationships often missed

by traditional statistical methods. In contrast, the Markowitz Investment Model, developed by Harry Markowitz, is grounded in modern portfolio theory and emphasizes diversification as a means of optimizing risk-adjusted returns. The model constructs efficient portfolios that offer the highest expected return for a given level of risk or the lowest risk for a given level of return. It starts by identifying the expected returns, volatilities, and correlations of various assets within the investment universe, considering investor risk preferences captured by utility functions. By systematically weighing the risk-return trade-offs of different asset combinations, the Markowitz Model provides a disciplined approach to portfolio management that aims to optimize returns while managing risk effectively. In a real-life context, both models offer valuable insights and guidance for investors and financial institutions. For example, a financial analyst may use LSTM to predict future stock prices based on historical data, leveraging its ability to capture temporal dependencies and adapt to changing market dynamics. Meanwhile, the Markowitz Model can guide portfolio construction by identifying optimal asset allocations that balance risk and return according to investor preferences. Overall, the LSTM and Markowitz Investment Models complement each other, with LSTM providing granular insights into market trends and patterns, while the Markowitz Model offers a systematic framework for constructing diversified portfolios tailored to investor objectives. By leveraging the strengths of both models, investors can make more informed decisions and navigate the complexities of financial markets with confidence and precision.

VII. RESULT AND DISCUSSION

The results and discussions pertaining to investment strategies and models offer a comprehensive exploration into the multifaceted landscape of financial markets, providing investors with valuable insights and considerations to enhance their decision-making processes and optimize portfolio performance while managing risk effectively. Firstly, in-depth examinations of the LSTM (Long Short-Term Memory) Model underscore its potential to significantly improve predictive accuracy within financial markets. By harnessing the power of historical data and effectively capturing nonlinear relationships, the LSTM Model emerges as a potent tool for forecasting future price movements and identifying prevailing market trends. Its inherent capability to process sequential data while retaining long-term dependencies renders it particularly well-suited for time-series analysis in finance, thereby offering a robust foundation for informed decision-making. Nevertheless, the scholarly discourse surrounding the LSTM Model also illuminates certain challenges that warrant careful consideration. These challenges include issues such as data scarcity, the risk of overfitting, and the interpretability of model outputs. To address these concerns, researchers advocate for the adoption of various strategies, including the incorporation of additional features, the implementation of regularization techniques, and the development of frameworks to enhance model interpretability.

Conversely, discussions revolving around the Markowitz Investment Model underscore its enduring relevance in the realm of portfolio construction and risk management. Through the mechanism of mean-variance optimization, the Markowitz Model facilitates the creation of portfolios that effectively balance risk and return according to investors' preferences, thereby contributing to the attainment of an efficient frontier. Despite its intuitive appeal and theoretical robustness, criticisms directed towards the Markowitz Model shed light on certain limitations that necessitate careful consideration. These limitations encompass issues such as the model's sensitivity to input parameters, the underlying assumptions of normality, and the potential impact of estimation error. In response to these challenges, researchers advocate for the exploration of alternative optimization techniques, the adoption of robust estimation methodologies, and the incorporation of considerations for non-normal distributions.

Furthermore, the literature underscores the significance of integrating diverse approaches and models into the decision-making processes of investors. The synergistic integration of the predictive capabilities offered by the LSTM Model with the diversification benefits provided by the Markowitz Model, for instance, holds the promise of enhancing portfolio performance and optimizing risk management strategies. Moreover, scholars advocate for the adoption of a dynamic approach to investment strategies, which involves the continuous adaptation of models and methodologies to effectively navigate evolving market conditions and accommodate shifting investor preferences.

In essence, the results and discussions encapsulate the intricate interplay between different investment models, elucidating their respective strengths, limitations, and potential synergies. Through a thorough examination of these models and the insights gleaned from empirical research, investors are empowered to make informed decisions that align with their financial objectives, while effectively managing risk amidst the dynamic fluctuations of financial markets. It is through this iterative process of evaluation and refinement that investors can navigate the complexities of financial markets with confidence and precision, ultimately optimizing their investment outcomes and fostering long-term financial success.

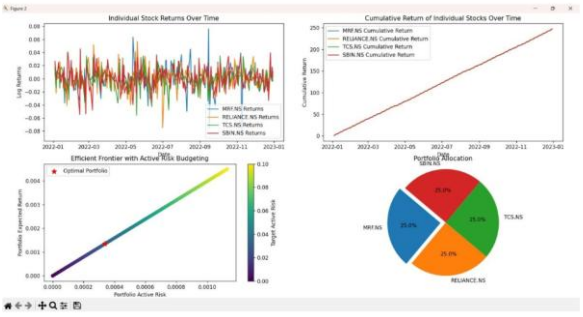


Figure 4: Hedged Portfolio's Output

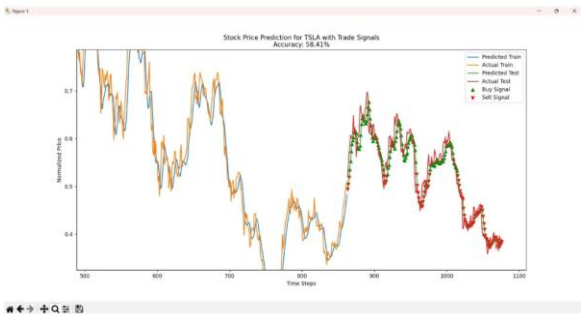


Figure 5: Output Graph

CONCLUSION

In summary, delving into investment strategies and models, particularly focusing on the LSTM (Long Short-Term Memory) Model and the Markowitz Model, provides illuminating insights into the dynamic realm of financial decision-making. Each model offers distinct avenues for maximizing returns and mitigating risks within investment portfolios, yet they also present their own unique challenges and considerations.

The LSTM Model exhibits significant potential by leveraging historical data to forecast future market trends, offering valuable insights into nonlinear relationships and sequential patterns. However, its efficacy relies on addressing hurdles such as data scarcity, overfitting, and model interpretability. Strategies aimed at enhancing the LSTM Model's performance include integrating additional features, employing regularization techniques, and refining interpretability frameworks.

Conversely, the Markowitz Model remains a cornerstone in portfolio theory, underscoring the significance of diversification in minimizing risk while optimizing returns. Despite its widespread adoption, critiques of the model spotlight issue regarding sensitivity to input parameters, assumptions of normality, and estimation error. Exploring alternative optimization methods, robust estimation techniques, and considerations for non-normal distributions present avenues for bolstering the Markowitz Model's effectiveness.

Furthermore, literature emphasizes the value of integrating diverse approaches and models into investment decision-making. By harnessing the predictive capabilities of the LSTM Model alongside the diversification advantages of the Markowitz Model, investors can strive for enhanced portfolio performance and risk management. Additionally, embracing a dynamic approach to investment strategies, which adapts models to evolving market conditions and investor preferences, proves vital for effectively navigating financial markets' intricacies.

In essence, the exploration of investment strategies and models unveils a nuanced comprehension of the trade-offs, opportunities, and challenges intrinsic to portfolio management. By critically assessing these models, incorporating insights from empirical research, and embracing a flexible and adaptive stance, investors can make informed decisions to pursue their financial objectives while adeptly managing risk amidst a constantly evolving market landscape. Through continuous refinement and adaptation, investors can navigate financial markets with assurance and precision, ultimately optimizing their investment outcomes and fostering enduring financial prosperity.
(i.e. Human vs. Machine).

REFERENCES

[1] De Prado, Marcos Lopez. (2018). "Advances in Financial Machine Learning." Marcos Lopez de Prado, True Positive Technologies.

[2] Markowitz, Harry. (1952). "Portfolio Selection." The Journal of Finance, 7(1), 77-91.

[3] Hochreiter, Sepp, & Schmidhuber, Jürgen. (1997). "Long Short-Term Memory." Neural Computation, 9(8), 1735-1780.

[4] Sharpe, William F. (1964). "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." The Journal of Finance, 19(3), 425-442.

[5] Fabozzi, Frank J., & Markowitz, Harry. (2011). "The Theory and Practice of Investment Management." John Wiley & Sons.

[6] Cho, Kyunghyun, et al. (2014). "Learning Phrase Representations using RNN Encoder– Decoder for Statistical Machine Translation." arXiv preprint arXiv:1406.1078.

[7] Campbell, John Y., & Viceira, Luis M. (2005). "Strategic Asset Allocation: Portfolio Choice for Long-Term Investors." Oxford University Press.

[8] Bengio, Yoshua, et al. (2003). "A Neural Probabilistic Language Model." Journal of Machine Learning Research, 3(6), 1137-1155.

[9] Swensen, David F. (2009). "Pioneering Portfolio Management: An Unconventional Approach to Institutional Investment." Simon and Schuster.

[10] Heaton, J.B., et al. (2017). "A Deep Learning Approach to Predicting Stock Price Movements." Financial Innovation, 3(1), 1-23

❑ ❑ ❑ ❑ ❑ ❑

AET 012

Saga – Gauging Seconds

Smruti Yadav
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
smruti.yadav@vit.edu.in

Siddhi Sawant
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
siddhi.sawant20@vit.edu.in

Janhvi Gupta
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
janhvi.gupta@vit.edu.in

Chhaya Sharma
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
chhaya.sharma@vit.edu.in

Dr. Nayana Mahajan
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
nayana.mahajan@vit.edu.in

***Abstract*– In the age of digitalization, where timepieces have transcended their traditional utility to become symbols of style, luxury and personality, the need for a comprehensive watch website is of great importance. An e-commerce platform designed for watch enthusiasts, where a variety of durable products from different brands and designers are offered. The platform aims to offer users hassle-free shopping with intuitive navigation, detailed product information and transaction security. This website aims to redefine the online watch shopping experience by appealing to a variety of interests and providing users with the information they need to place orders. HTML forms a backbone, defining the structure and content elements of the webpage. CSS is employed for styling, to specify the visual presentation of elements. Java-script adds interactivity and dynamic behavior to the site. Python is commonly employed for backend tasks. The website aims to change the way customers buy watches by offering a wide range of products, unmatched convenience, and an unmatched market experience.**

I. INTRODUCTION

E-commerce involves buying and selling goods and services online through digital platforms. It uses the internet to conduct business, manage inventory and attract customers, changing traditional business practices. Elements of e-commerce include online stores for businesses to showcase their products and services, digital payment systems that support secure transactions, supply chain management to ensure efficient transportation, digital marketing strategies to attract and retain customers, and customer satisfaction. This is a website where you can discover a variety of stylish watches from the comfort of your home. We have a collection of watches from the best brands and designers to suit every taste and budget. Whether you are looking for a classic watch for daily use, or a watch that will enhance your style, everything you are looking for is here. Buying a watch has never been easier thanks to easy navigation and secure transactions. The basis of the project is to spend time and understand the needs of today's buyers. With the rapid growth of e-commerce, we recognize the importance of providing a digital space for watch lovers to find, compare and buy their favorite watches. Whether the choice is a purchase or a first-time buyer. By carefully selecting watches from famous and new brands, from classics to new designs, we ensure that every customer finds a watch that suits their style and budget. In summary, the Saga watch website project aims to set a new standard for online watch platforms, offering a curated blend of information, interactivity, and community engagement. This website seeks to inspire, educate, and delight watch enthusiasts of all backgrounds.

II. EASE OF USE

Efficient Selection: This website features a wide selection of watches, including luxury watches, affordable options. Whatever the customer's preference this website has something for everyone.

User-friendly interface: This website understands the importance of search links. Our user-friendly interface allows users to easily browse categories, filter options, and product details, making the shopping process efficient and enjoyable.

Product description: All listings contain detailed information, including technical specifications, watch features, high-resolution images. The website believes in giving the customers the information they need to make a purchasing decision.

Customer Experience: Providing a seamless & personalized shopping experience is crucial in e-commerce. This includes user-friendly website design, personalized recommendations.

Security: We prioritize the security of sensitive information. The website uses strong encryption techniques to ensure all transactions are safe and secure.

III. REVIEW OF LITERATURE

For the designing of Saga, we referred various literatures, papers, etc. [1] The idea behind the overall working, the website is an e-commerce watch website which included the same features, the customers can buy watches at discounted prices. It even has various additional features, mainly it includes the virtual try-on feature of different watches. The customers can virtually try on the watches at the comfort of their home. This will help customers to enhance their shopping experience. These websites typically use advanced augmented reality (AR) technology to provide an immersive experience for users, allowing them to virtually try on watches in real-time or through uploaded photos. [2] Unlike the conventional watch websites, and the same old features this website includes the customization feature, with the help of which customers can customize their own watches according to their likes and dislikes. There is a separate section present for the users to customize the watch starting from the watch strap, color, dial and everything else as per the customer needs. [3] The website has a unique option of selling your old watches and getting new ones, many watch websites even provide watches on EMI. Several additional features were tested to determine which is the best website to purchase watches from. Following factors make each website unique:

- Virtual try-on

Each of these watch websites which we studied offers a virtual try-on feature with its own unique strengths, whether it's in terms of realism, user experience, convenience, versatility, brand prestige, or selection. These features contribute to an

enhanced online shopping experience for customers, helping them make more informed decisions when purchasing watches online.

- Customization

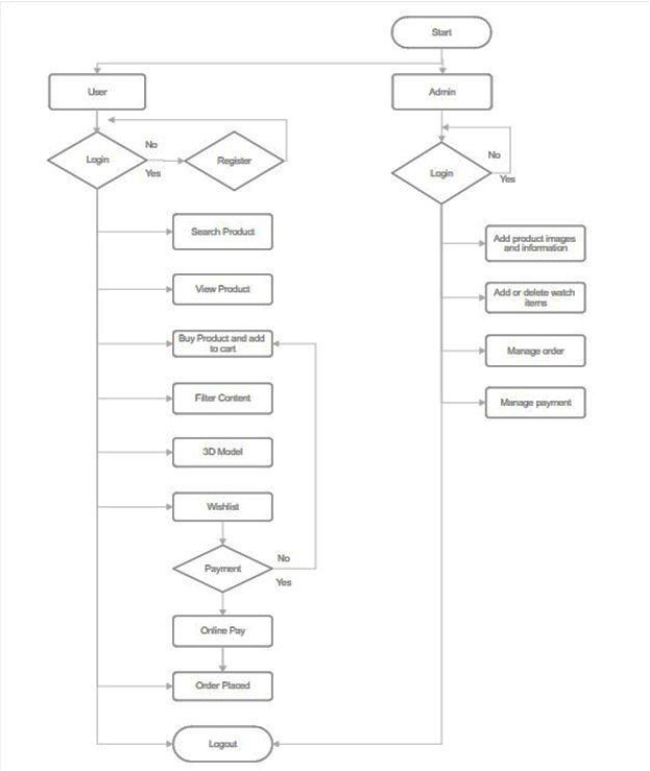
Several watch brands offer online configurators or bespoke services that allow you to visualize your custom watch design before placing an order. Whether you're designing a watch for everyday wear, a special occasion, or as a collector's piece, customizing your own watch allows you to express your individuality and create a truly unique accessory that reflects your style and personality.

- Exchange

Several websites and retailers offer trade-in programs where you can exchange your old watch for credit towards the purchase of a new one. Before participating in a trade-in program, be sure to carefully review the terms and conditions, including any eligibility requirements, appraisal processes, and valuation criteria. Additionally, compare offers from different retailers to ensure you're getting the best value for your trade-in.

IV. METHODOLOGY

The flowchart describes the flow of the system.



The above flowchart depicts the entire flow of the website. The website is divided into two sections comprising of the front end and the back end. The front-end part is visible to the user. The user can search for the products, view products, buy products and add them to the cart. For the payment procedure online transaction can be done. After adding the product to cart and after the payment is done. It will notify you that the order is placed. The backend part mainly handles the work related to adding anything and everything about the product. We can add the watch images and information from the backend. It also handles the delivery of the product, whether it is in transit or delivered. The access to the backend is only for the admin.

The entire project:



Figure 1. Home Page

As you land on the homepage, you'll be greeted with a captivating showcase of our latest and most popular timepieces, each one meticulously selected to cater to diverse tastes and preferences as shown in fig.3. Scroll down to explore our curated collections, featuring everything from classic designs to cutting-edge innovations including separate sections for women's and men's watches.

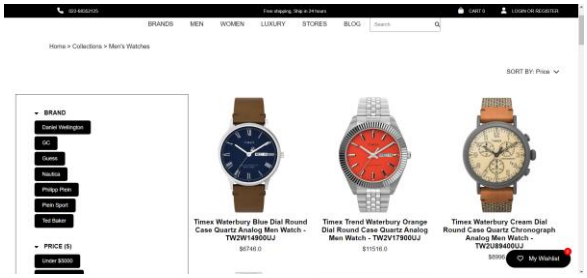


Figure 2. Men's Section

As shown in the above figure, it represents the men's section. This page offers more than just an impressive collection of watches. Men's department consists of watches curated specifically for men. Take benefits of our intuitive search and filtering options to quickly get a perfect timepiece to match your style and budget.

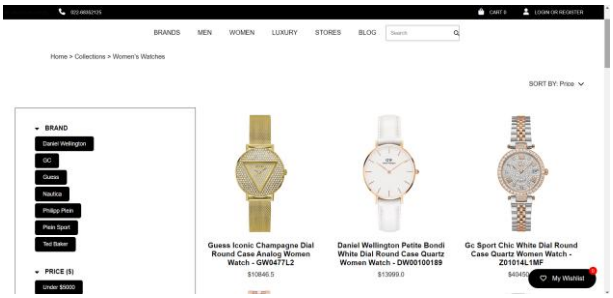


Figure 3. Women's Section

As shown in the above figure, it represents the women's section. This page offers more than just an impressive collection of watches. Women's section consists of watches specifically for women. Take advantage of our intuitive search and filtering options to quickly find the perfect timepiece to match your style and also they are budget friendly.

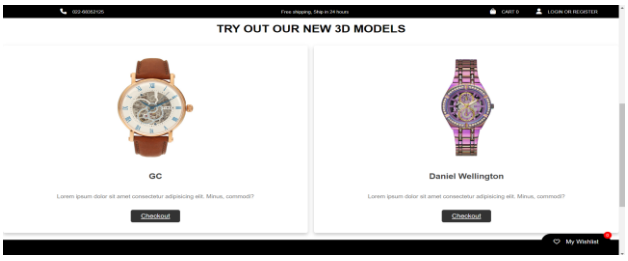


Figure 4. Luxury

The above figure, is the luxury section of the website. People can subscribe to the luxury section by adding the email id of the respective user. The 3D Model feature enhances the online shopping experience, providing users with a clear understanding of watch design and aesthetics before making a purchasing decision.



Figure 5. 3D Models

3D models are utilized to create realistic and dynamic product images of watches in an e-commerce website. This 3D model serves as visual representation of the watches, allowing users to visualize and interact with the product from various angles. First, special software is used to create a 3D model of the watch. The software accurately captures the model, shape and details of the watch, including the dial, hands and strap. These models are textured using high-resolution images to replicate materials such as metal, leather or glass, adding realism to the visualization. This modification allows the creation of beautiful images that present the watch in its best light, highlighting its functionality. Interact with it with features like zoom, rotate, rotate & examine the watch from any angle. Overall, using 3D models to create visuals for ecommerce websites adds a layer of realism and interactivity, eliminating the difference between online shopping and experience sales.



Figure 6. Our Stores

The figure above depicts the stores of the website and its location in the nearby area. In the shop the customer can find exclusive curated timepieces from top brands and trendy designs. Whether they are into fancy watches or formal ones for daily chores, the website has something for everyone.

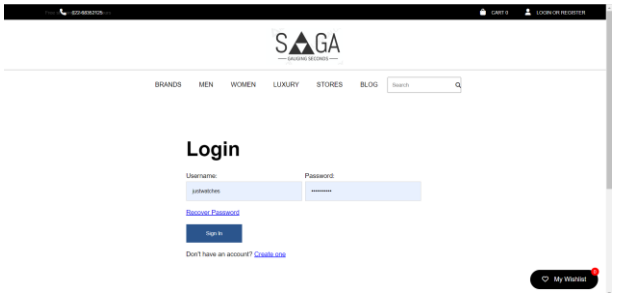


Figure 7. Login

The above figure shows the login page of the website. In order to stay up to date with the latest trends, offers and releases, sign up for the newsletter and be the first to know about exclusive offers, new arrivals, and exciting trends

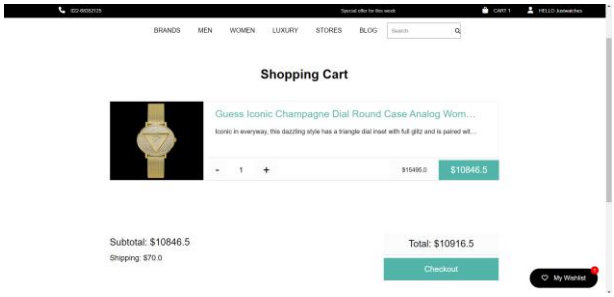


Figure 8. Shopping Cart

Fig. 8 shows the shopping cart of the website. As the user explores the wide range collection of watches, they can simply click the "Add to Cart" tab next to their chosen items to place them here. The intuitive interface allows the user to review their selections, update quantities, and proceed to checkout. The Cart Section is designed to maintain the shopping journey, with secure payment options and efficient order processing.

V. SOFTWARE REQUIRED

The project utilized technologies like HTML, CSS, Python and JavaScript. HTML forms a backbone, defining the structure and content elements of the webpage. CSS is employed for styling, to specify the visual presentation of HTML elements. Java-script adds interactivity and dynamic behavior to the site. Python is commonly employed for backend tasks.

VI. CONCLUSION

In accordance, this watch website has focused on providing a seamless and satisfying experience for clients looking to explore and purchase time gadgets. We've mindfully curated a wide range of watches to cater to different tastes and needs. Our aim is to provide not just a platform for shopping, but also a place where watch enthusiasts can discover new variety in styles, learn about watch craftsmanship, and connect with others who share the same interest. We have enhanced our website with various extraordinary features like filters, 3D models, easy transactions for the website and providing seamless and reliable customer service.

REFERENCES

[1] Vladimir Zwass, “International Journal of Electronic Commerce”, Volume 28, Issue 1 (2024).

[2] Hongqiang Yue, “Research on E-Commerce Data Standard System in the Era of Digital Economy from the Perspective of Organizational Psychology”, Volume 13 – 2022.

[3] Sarah S. Alrumiah, Mohammed Hadwan “Implementing Big Data Analytics in E-Commerce: Vendor and Customer View” Year: 2021 Volume: 9.

[4] Yidan Wang, “Research on E-commerce Platform of Online Shopping Consumers”, Year: 2021.

[5] Zhijie Zhao, Jiaying Wang, Huadong Sun, Yang Liu, Zhipeng Fan, Fuhua Xuan, “What Factors Influence Online Product Sales? Online Reviews, Review System Curation, Online Promotional Marketing and Seller Guarantees Analysis”, Year: 2020, Volume: 8.

❑ ❑ ❑ ❑ ❑ ❑

AET 013

Car crash Detection and Assistance

Shritej Magare
Vidyalankar Institute of Technology
Wadala, Mumbai, India

Meet Limbani
Vidyalankar Institute of Technology
Wadala, Mumbai, India

Dr. Nayana Mahajan
(Associate Professor)
Vidyalankar Institute of Technology

Sahil Sharma
Vidyalankar Institute of Technology
Wadala, Mumbai, India

Bhargav Rane
Vidyalankar Institute of Technology
Wadala, Mumbai, India

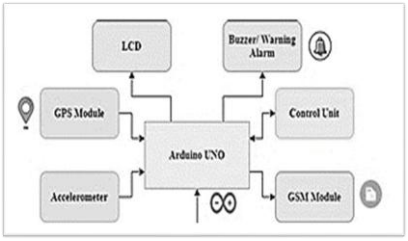
***Abstract*– Our paper aims to address the critical issue of individuals being left withoutassistance in the event of an accident while riding their vehicle. The GPS Receiver identifies the vehicle's direction, while the SMS containing the emergency text is sent. The system can detect severe accidents using an Impact sensor. The information is sent by the GSM module, which transmits the data, including the victim's precise location, to the assigned emergency/personal contact. This System offers a professional and innovative solution to a significant problem, ensuring the safety of individuals. This is a highly effective solution, particularly in developing nations less signal availability in remote area. However, this System can prevent uncertain deaths by sending a message alert to a registered mobile number, providing the precise location ofthe accident through a Google map link. This system is a valuable investment in ensuring the safety of drivers and passengers alike.**

I. INTRODUCTION

This project detects sudden change in the axis of vehicle due to a crash or accident. And then the project sends a call alert by making call to your mobile phone. Accident Alert System Project also send an alert SMS on your MobilePhone with the location of the accident. We have Added the Emergency phone number in the code for receiving SMSalerts as well as call alerts.

The GPS Location of the accident is sent to your phone in the form of GoogleMap link along with the latitude and longitude. you just tap on the link, to visualize the real time GPS data on Google maps.

This Accident alert system can also be used as GPS Tracking System. You just send the "get GPS" command from your phone. and you will receive the location of the vehicle on your phone. If the project detects false crash due to noise or some other reason. Then youdon't want to receive that false accident alert on your phone. So, we have added a push button. If you want to cancel the Alert SMS then you just pressthis push button within 30 seconds and it will cancel the SMS and Call Alert. The whole project is based on GSM network. So, no internet connection isrequired for this project. you just need a sim card and GSM module for sending SMS alerts and making calls. This Algorithm detects accident by getting the x, y and z axis values from accelerometer. This algorithm differentiates between parking mode and driving mode. So, it can detect crashes while you are driving. And it can also detect the crash when the car is parked. Promotion of Technological Innovation: Encourage the adoption ofadvanced technology in the automotive industry, promoting innovation and job creation in the field. Accessibility for Vulnerable Populations: Provide a safety net forvulnerable populations, such as the elderly or individuals with medical conditions, to ensure their mobility and safety while driving. Accountability: Promote accountability among drivers, reducingincidents of hit-and run accidents and encouraging responsible driving practices. User Education and Awareness: Raise awareness and educate users about the system's capabilities and limitations to ensure effective utilization and maximize its benefits. Ethical Considerations: Address ethical and social considerations, such as the responsible use of data and privacy protection, to prevent negative impacts on users and society.



System Implementation Fig.1

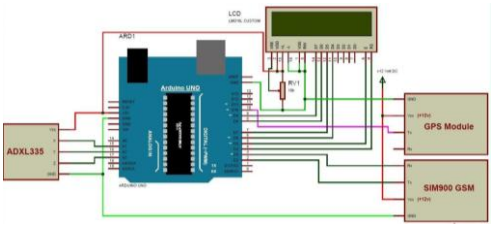


Fig.2: Virtual simulation Circuit Diagram

This paper refers to the GSM-based SMS warning system of accident detection using Arduino. The GSM module plays an important role in receiving the signalfrom the Arduino UNO in the above block diagram. When the power supply is supplied to the circuit until the Accelerometer sensor senses sudden axis shift, it sends the signal to the Arduino and the position of the vehicle obtained by the GPS module, and the vehicle coordinates are sent to the GSM module. The received data is gathered and sent via SMS to the valued individual.

Circuit connections of this Vehicle Accident Alert System Here The GPSmodule's TX pin is directly connected to the Arduino digital pin number 10. The GPS module is powered by a 3.3-volt supply. The Tx and Rx pins are directly related to Arduino pins D2 and D3. In this accident detection system, an Accelerometer is added and its x,y & z-axis ADC output pins are directly connected to Arduino ADC pins A1, A2, A3.

SOFTWARE

Arduino IDE

This is an Arduino code for a vehicle accident alert system. The system uses a 3-axis accelerometer to detect impacts or accidents. When an impact is detected, the system records the magnitudeof the impact and checks if it is greater than or equal to a predefined sensitivity level. If the impact magnitude is greater than or equal to the sensitivity level, the system considers it as an accident and triggers an alert.

The alert system consists of the following components:

- A 3-axis accelerometer to detect impacts or accidents.
- A GPS module to get the location data.
- A GSM module to send SMS alerts & make emergency calls.
- A buzzer to indicate the impact detection.
- A button to stop the alert.

The code initializes the required libraries, pins, and modules. It sets up the LCD screen, GSM module, GPS module, and accelerometer. The code then enters an infinite loop where it continuously checks for impacts using the accelerometer. When an impact is detected, the code records the magnitude of the impact and checks if it is greater than or equal to the sensitivity level. If the impact is considered an accident, the code gets the GPS location data and sends SMS alerts to the emergency contact numbers stored in the EMERGENCY_PHONE constant. The code also makes an emergency call to the emergency contact number.

Result

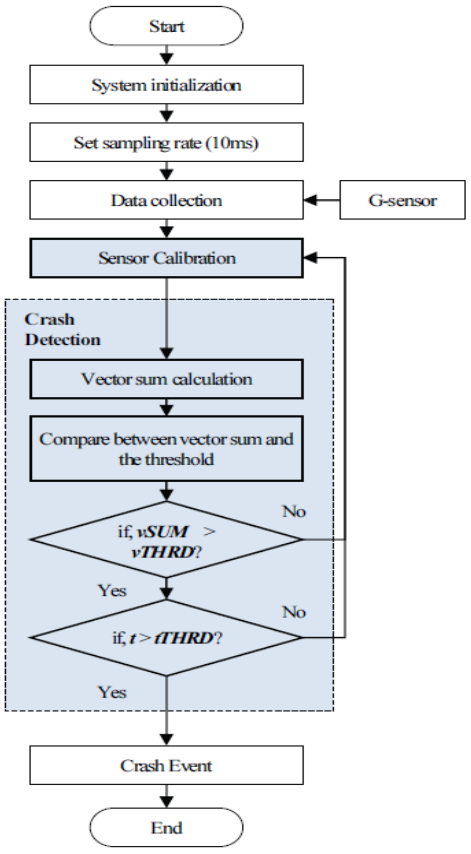


Fig 3: Purposed flow chart of Crash Detection



Fig. 4: PCB mounted system

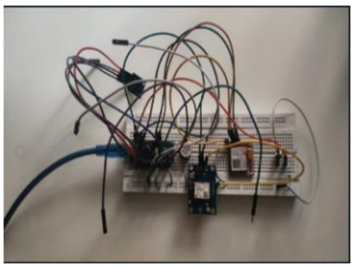


Fig.5: Breadboard design

Fig.4 shows the PCB design of the system in which Here The GPSmodule's TX pin is directly connected to the Arduino digital pin number The GPS module is powered by a 3.3-volt supply. The Tx and Rx pins are directly related to Arduino pins D2 and D3.In this accident detection system, an Accelerometer is added and its x,y and z-axis ADC output pins are directly connected to Arduino ADC pins A1, A2, and A3.Similary in Fig.5 the same connection is on the breadboard.

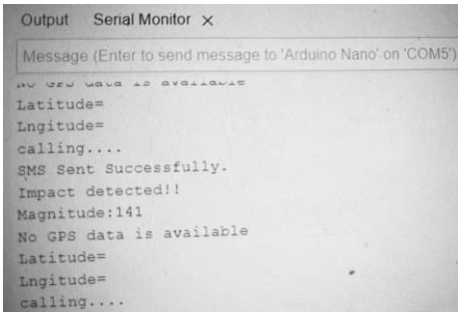


Fig.6: Virtual simulation output

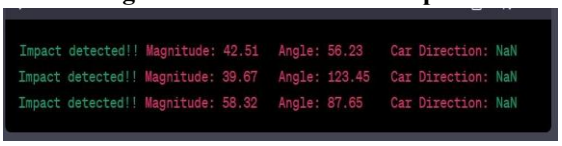


Fig.7: Virtual Impact Detection with "15"mag. threshold

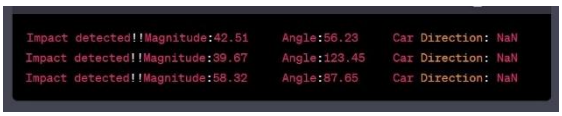


Fig.8: Virtual Impact Detection with "12"mag. Threshold

Impact Threshold	Impact Magnitude	Detection
15	42.51	Yes
15	12.14	No
15	39.69	Yes

In the above table the threshold value is set to be 15. The impact value detected is 42.51 which exceeds the threshold value. Hence, the output is "Yes" as the impact detected is greater than the threshold value. But when the impact value is 12.14 i.e. less than the threshold value the output shown is "No".

Impact Threshold	Impact Magnitude	Detection
12	42.51	Yes
12	10.11	No
12	39.69	Yes

In the above table the threshold value is set to be 12. The impact value detected is 42.51 which exceeds the threshold value. Hence, the output is "Yes" as the impact detected is greater than the threshold value.
But when the impact value is 10.11 i.e. less than the threshold value the output shown is "No".

- The code generates random values for the simulated accelerometer readings and triggers impact detection when the magnitude exceeds the sensitivity threshold.
- This is simulated output for testing and debugging purposes,and it's not based on real sensor data
- The exact values for magnitude and angle will be different eachrun the code because the accelerometer-like data is generated randomly.

CONCLUSION

This paper has proposed a novel crash detection algorithm applicable to general driving and parking mode scenarios that consider a various direction of car and that extracts as much information as possible from the calibrated data. The algorithm successfully detected in various directions of crash. It is expected that the proposed crash detection algorithm is reliable although further evaluations are required. The proposed algorithm can be used as an integrated collision detection algorithm by integrating tracking information from multiple sources for collision warning, avoidance and mitigation. This system offers the most realistic alternative to the inadequate emergencyservices given to victims of road accidents. With the aid of this technology, when an accident happens, prompt action can be taken by alerting the appropriate individuals by sending a message. Throwing new light on existing methods, we hope that the proposed algorithm will complement others and help to reduce overfitting to datasets with little test examples or training as well as contribute to the development of algorithms that work well in the crash detection system of vehicle.

REFERENCE

[1] Tanushree Dalai, "Emergency Alert and Service for Automotives for India", International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE) Mysore India, vol. 2, no. 5, pp. 08-12, 2013.

[2] Manuel Fogue, Piedad Garrido, Francisco J. Martinez, JuanCarlos Cano, Carlos

[3] T. Calafate, and Pietro Manzoni (2012) "Assistance through Communication Technologies and Vehicle", IEEE vehicular technology magazine.

[4] PL Needham, Collision prevention: The role of an accident data recorder. Automated Emergency Call for Road Accident, European Commission Press G. Singh and H. Song, Using Hidden Markov Models in Vehicular crash detection, IEEE Transactions.

[5] Niranjan Kumar K., Rama Narasimha Dattu C.H., Vishnu S., Jino Ramson S.R.(2019) Automatic Accident Rescue System Using IoT. In: Smys S., Bestak R., Chen JZ., Kotuliak I. (eds).

[6] Niranjan Kumar K., Rama Narasimha Dattu C.H., Vishnu S., Jino Ramson S.R.(2019) Automatic Accident Rescue System Using IoT. In: Smys S., Bestak R., Chen JZ., Kotuliak I. (eds)

[7] R. R. Knipling et al., "Assessment of IVHS countermeasures for collision avoidance systems," National Highway Traffic Safety Administration, Washington, DC, USA, Tech. Rep. DOT HS 807 995, May 1993.

[8] T. Kim and H. Jeong, "A Novel Algorithm for Crash Detection Under General Road Scenes Using Crash Probabilities and an Interactive Multiple Model Particle Filter," in IEEE Transactions on Intelligent Transportation Systems, vol. 15, no.6, pp. 2480-2490, Dec. 2014.

[9] V. Naumov, "Analysis of Time and Distance Delays in Car Following Models," 2010 International Conference on Intelligent Systems, Modelling and Simulation, Liverpool, 2010, pp. 296-299.

[10] Geiger, P. Lenz and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, 2012, pp. 3354- 3361.

[11] Cismas, M. Ioana, C. Vlad and G. Casu, "Crash Detection Using IMU Sensors," 2017 21st International Conference on Control Systems and Computer Science (CSCS), Bucharest, 2017, pp. 672-676.

❑❑❑❑❑❑

AET 014

Performance Analysis of 16 Bit RISC Processor Using Vedic Multiplier

Mithilesh Kavade
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
Mumbai, India
mithilesh.kavade@vit.edu.in

Mayank Patil
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
Mumbai, India
mayank.patil@vit.edu.in

Prasad Deshmukh
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
Mumbai, India
prasad.deshmukh@vit.edu.in

Sankalp Mhatre
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
Mumbai, India
sankalp.mhatre@vit.edu.in

Akhil Masurkar
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
Mumbai, India
akhil.masurkar@vit.edu.in

***Abstract*— In this work a comprehensive performance analysis of a 16-bit RISC processor integrated with two different multiplier architectures namely Vedic multiplier and shift and add multiplier. The objective was to evaluate and compare the efficiency and effectiveness of these multiplier designs in terms of performance and power consumption. The designs were implemented and synthesized using Verilog in Xilinx Vivado. The tested architecture was implemented on a Spartan 7 FPGA. The above two multipliers were implemented inside a 16-bit RISC processor and the processor performance was measured. Fourteen different operations using the RISC processor are considered in the analysis. The performance was measured using parameters like WNS (Worst Negative Slack), TNS (Total Negative Slack), WHS (Worst Hold Slack), THS (Total Hold Slack), TPWS (Total Pulse Width Slack), Total Power, Number of LUT's used and Number of Flip Flops used. It is observed that the RISC processor performance improves by around 53.08% when a Vedic multiplier is used instead of Shift and Add multiplier.**

***Keywords*—16-bit RISC, Vedic multiplier, shift and add multiplier, Xilinx Vivado, FPGA, Spartan 7.**

I. INTRODUCTION

Multiplication is a crucial basic operation within an ALU. Given that multiplication consumes a significant portion of the execution time for most of the DSP algorithms, the need for a high performance multiplier is paramount. Presently, the time required for multiplication remains a key factor in determining the instruction cycle duration of a DSP chip. As the demand for enhanced computing power on battery-powered mobile devices grows, the focus in design has shifted from merely optimizing traditional metrics like delay, time, area, and size to also minimizing power dissipation while maintaining high performance. This paper aims to assess the performance of a 16-bit (RISC) processor that incorporates two distinct multiplier architectures: the Vedic multiplier and the shift-and-add multiplier. The objective is to compare these multiplier designs based on their area utilization, power consumption, and operating speed, implemented on an FPGA board using Verilog in Xilinx Vivado.

The advantage of Vedic mathematics lies in its ability to simplify conventional mathematical calculations significantly. This simplification is attributed to the Vedic formulas, which are believed to align with the natural functioning of the human mind. Vedic Mathematics encompasses a set of arithmetic principles that enable more efficient computational speed . The Urdhva Tiryakbhyam Sutra, which translates to "Vertically and Crosswise," serves as a universal multiplication formula suitable for all multiplication scenarios . In this method, digits at the ends of the line are multiplied, and their product is added to the preceding carry. When multiple lines are involved in one step, all resulting products are summed with the previous carry. The least significant digit of this sum becomes part of the final result, while the remaining digits serve as the carry for the next step. Initially, the carry is assumed to be zero.

II. LITERATURE REVIEW

A. 16 BIT RISC PROCESSOR

A 16-bit Reduced Instruction Set Computing (RISC) processor is a type of microprocessor architecture that emphasizes a streamlined instruction set aimed at enhancing simplicity, efficiency, and performance. This processor operates on 16-bit wide data units. Its architecture typically includes a limited set of instructions, which are optimized for common tasks to achieve high performance and efficient execution. Unlike Complex Instruction Set Computing (CISC) processors, RISC processors feature a simplified instruction set. In a 16-bit RISC architecture, instructions are usually fixed-length and are designed to execute in a single clock cycle, thus simplifying the instruction decoding process.

These processors typically incorporate a small number of general-purpose registers (GPRs) for data storage and manipulation. These registers are easily accessible by the instructions, enabling efficient data processing and reducing the latency of memory access. RISC processors often utilize a load-store architecture, where data transfer between memory and registers is explicitly managed by load (LD) and store (ST) instructions. This method streamlines the instruction set and enhances efficiency by separating data access operations from arithmetic operations.

The compact size and efficient operation of 16-bit RISC processors make them ideal for applications where power consumption, space, & cost are critical factors. Consequently, they are commonly found in consumer electronics, automotive systems, industrial control, and IoT devices. When designing a 16-bit RISC processor, it is essential to balance performance with constraints on area & power consumption. Optimizations in microarchitecture, instruction scheduling, & data forwarding techniques are key to achieving optimal performance within the constraints of a limited instruction set width.

Table I: Sample Instruction set for 16 bit RISC processor.

Operation	Opcode				Result
LW	0	0	0	0	Read from memory
SW	0	0	0	1	Write to memory
ADD	0	0	1	0	$R_c = R_a + R_b$
SUB	0	0	1	1	$R_c = R_a - R_b$
MUL	0	1	0	0	$R_c = R_a * R_b$
AND	0	1	0	1	$R_c = R_a \& R_b$
OR	0	1	1	0	$R_c = R_a \mid R_b$
XOR	0	1	1	1	$R_c = R_a \wedge R_b$
NOT	1	0	0	0	$R_z = \sim R_a$
SHL	1	0	0	1	Shift left R_a
SHR	1	0	1	0	Shift right R_a
INC	1	0	1	1	Increment R_a
DEC	1	1	0	0	Decrement
HLT	1	1	0	1	Stop
JMP	1	1	1	0	Jump
MVI	1	1	1	1	Move Immediate

A 16-bit RISC (Reduced Instruction Set Computer) processor operates efficiently by utilizing a streamlined instruction cycle. This cycle begins with the program counter (PC) fetching the next instruction from memory. Each instruction, being 16 bits in size, is then decoded to identify the operation and its operands. The subsequent stage involves fetching the necessary operands from registers or memory.

With the operands available, the processor proceeds to the execute stage, where the Arithmetic Logic Unit (ALU) performs the specified operation, which can include arithmetic, logical, or control functions. If the instruction involves memory access, such as loading or storing data, this is handled during the memory access stage. The final step in the cycle is the write back stage, where the result of the operation is stored in the destination register or memory location, thereby completing the instruction cycle.

The design of a 16-bit RISC processor emphasizes simplicity and efficiency. It employs a reduced instruction set and a limited number of general-purpose registers. To boost performance, these processors frequently use pipelining, which allows different instruction processing stages to overlap. This approach ensures that most instructions are executed within a single clock cycle, enhancing predictability and efficiency.

The minimalist architecture of the 16-bit RISC processor prioritizes speed and efficiency over complexity. In the paper titled "Design and 16-bit RISC," the authors describe the development and implementation of a custom 16-bit RISC processor optimized for embedded applications. This processor features a reduced instruction set, efficient pipelining, and hardware support for essential operations, showcasing improvements in performance and area utilization compared to traditional designs.

B. VEDIC MATHEMATICS

Vedic Mathematics is a collection of techniques and sutras designed to solve arithmetic problems quickly and easily. It comprises 16 Sutras (formulas) and 13 sub-sutras (sub-formulas) that can be applied to arithmetic, algebra, geometry, calculus, and conics. This system of mathematics was discovered by the Indian mathematician Jagadguru Shri Bharathi Krishna Tirthaji between 1911 and 1918, and his findings were published in his book "Vedic Mathematics."

The term "Veda" is Sanskrit for "knowledge." Traditional mathematical methods can be complex and time-consuming, but Vedic Mathematics offers both general techniques (applicable to any set of data) and specific techniques (applicable to particular sets of data) that enable rapid numerical calculations.

Vedic Mathematics is renowned for its ability to expedite mathematical computations. While we agree with this to some extent, as certain methods in Vedic Mathematics are indeed very fast, it should be noted that some techniques are specific to certain types of numbers and are thus called specific methods.

Here are some operations that can be performed using Vedic Mathematics:

1. Division Shortcuts
2. Multiplication Shortcuts
3. Calculating Squares
4. Calculating Square Roots
5. Calculating Cubes
6. Calculating Cube Roots

Tirthaji Maharaj has categorized the methods for multiplying numbers in Vedic Mathematics into Specific and General Methods. Specific Multiplication Methods are applicable when numbers meet certain conditions, such as both numbers being close to 100, the numbers being close to each other, or the sum of the last digits of both numbers equaling 10. On the other hand, General Multiplication Methods can be applied to any numbers.

In Vedic Mathematics, multiplication techniques are classified into Sutras based on these Specific and General Methods:

1. Nikhilam Sutra (Specific Technique)
2. Anurupyena Sutra (Specific Technique)
3. Urdhva Tiryagbhyam Sutra (General Technique)
4. Ekayunena Purvena (Specific Technique)
5. Antyaordaske'pi (Specific Technique)

In our project, we have employed the Urdhva Tiryagbhyam Sutra, as it is the general technique for multiplication.

C. URDHVA TRIYAGBHYAM SUTRA

The Urdhva Tiryak Sutra is an effective shortcut method for multiplication in Vedic Mathematics. This general method allows for the quick multiplication of any type of numbers and can be easily applied to multiply 3-digit, 4-digit, or even larger numbers.

Here is a shortcut for multiplying 2-digit numbers using the Urdhva Tiryak Sutra: Shortcut to Multiply 2-digit numbers:
Formula: $(ax+b)(cx+d) = acx2 + (ad+bc)x + bd$.

Steps:

1. Vertical Multiplication of the First Digits: Multiply the first digits of both numbers.
2. Crosswise Multiplication Addition: Perform crosswise multiplication (i.e., multiply the first digit of the first number with the second digit of the second number and vice versa) and add the results.
3. Vertical Multiplication of the Last Digits: Multiply the last digits of both numbers.
4. Digit Management: For all steps except the first, ensure each result compartment contains only one digit. If the result has more than one digit, carry the initial digits to the previous compartment.

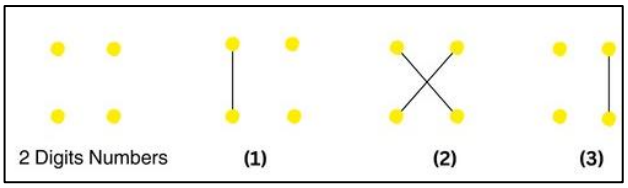


Fig. 1. Multiplication of 2-digit numbers using Urdhva Triyagbhyam Sutra.

Shortcut to Multiply 3-digit numbers:

Formula: (ax2 + bx + c) (dx2 + ex + f) = adx4 + (ae + bd) × 3 + (af+be+cd)x2 + (bf+ce)x + cf.

Steps:

1. Multiply the first digits of the two numbers vertically.
2. Add the result of crosswise multiplication of the first two digits of both numbers.
3. Add the result of crosswise multiplication of all three digits of both numbers.
4. Add the result of crosswise multiplication of the last two digits of both numbers.
5. Multiply the last digits of the two numbers vertically.
6. Ensure that each compartment, except the first one, contains only one digit. If there are more digits, carry forward the initial digits to the previous compartment (Refer to the examples below for clarity).

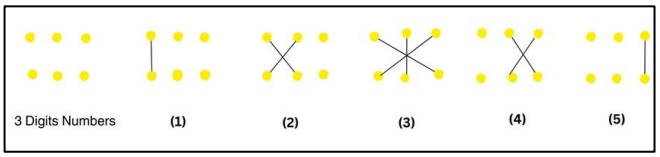


Fig. 2. Multiplication of 3-digit numbers using Urdhva Triyagbhyam Sutra.

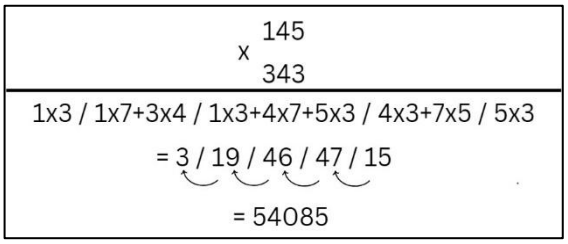


Fig. 3. Example 1 for Urdhva Triyagbhyam Sutra.

Shortcut to Multiply 4-digit numbers:

Steps:

1. Multiply the first digits vertically.
2. Add the result of crosswise multiplication of the first two digits.
3. Add the result of crosswise multiplication of the first three digits.
4. Add the result of crosswise multiplication of all digits.
5. Add the result of crosswise multiplication of the last three digits.
6. Add the result of crosswise multiplication of the last two digits.
7. Multiply the last digits vertically.

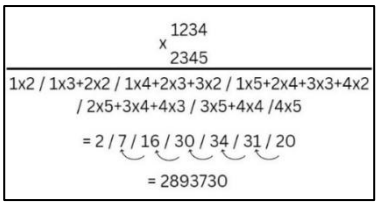


Fig. 4. Example 2 for Urdhva Triyagbhyam Sutra.

Similar techniques can be used for multiplication of n-digit numbers.

D. VEDIC MULTIPLIERS

2x2 Vedic Multiplier:

The 2 x 2 Vedic multiplier operates by employing the ancient Vedic mathematics technique called "Urdhva-Tiryak" (vertically and crosswise). In this method, two 2-bit binary numbers, denoted as A and B, are taken as inputs, and their product is efficiently computed through a series of straightforward arithmetic operations. This process involves multiplying the digits of A and B in a crosswise manner to generate partial products, which are then combined using simple addition operations to derive the final result. By utilizing parallelism and optimized arithmetic, this approach eliminates the necessity for complex multiplication circuits, making it well-suited for hardware implementations where speed and simplicity are paramount. The 2 x 2 Vedic multiplier demonstrates how ancient mathematical techniques remain relevant in modern computational systems, providing a compact and efficient solution for fundamental multiplication tasks.

4x4 Vedic Multiplier:

The 4 x 4 Vedic multiplier, rooted in the principles of Vedic mathematics, efficiently computes the product of two 4-bit binary numbers using a structured and parallel approach. This multiplier simplifies the multiplication process by breaking it down into several steps. Initially, it generates partial products by multiplying individual bits of the two input numbers in a crosswise manner. Each multiplication operation yields a partial product associated with a specific position in the final product. Subsequently, these partial products are aligned and added together using straightforward addition operations, considering their respective positional weights. This addition process resembles traditional multiplication but is structured differently. Finally, carry propagation is managed to ensure the accurate result. The 4 x 4 Vedic multiplier optimizes multiplication by leveraging parallel processing of partial products and streamlined addition steps, rendering it well-suited for hardware implementations where speed, simplicity, and efficiency are paramount.

8x8 Vedic Multiplier:

The 8 x 8 Vedic multiplier is constructed by employing four 4 x 4 Vedic multipliers. Initially, the two 8-bit inputs are divided into two parts, each 4 bits wide. These 4-bit inputs are subsequently fed into the four 4 x 4 multipliers. The products obtained from these multipliers are then forwarded to three adders, where the data is combined through addition operations. The result of these additions yields a final 16-bit output, representing the product of the two 8-bit inputs. This approach enables the multiplication of two 8-bit numbers by leveraging parallel processing and structured addition operations, ensuring efficient computation while maintaining simplicity and speed.

A. COMPARATIVE PERFORMANCE ANALYSIS OF VEDIC MULTIPLER IN EXISTING LITERATURE

In their study, referenced as [3], the authors underscore the efficiency of the Urdhva Triyagbhyam Sutra, emphasizing its minimal delay in multiplication operations across various numerical types and magnitudes. They illustrate the practical applicability and performance of this method through its implementation using Verilog HDL for 8x8 bit multiplication, followed by FPGA synthesis utilizing Xilinx tools on the Spartan 3E kit.

The synthesis results presented by the authors reveal a notable computation time of 15.05 ns for computing the product of 8x8 bits, underscoring the speed advantage of the Urdhva Triyagbhyam method over conventional multiplier architectures. Additionally, the study introduces innovative modifications to ripple carry adders, optimizing hardware utilization and further enhancing computation speeds.

Here are the results reported by the authors in their research paper:

Table II: Results found by Authors in [3].

Device Utilization Summary			
Logic Utilization	Used	Available	Utilization
Number of 4 input LUT's	171	7,168	2%
Number of occupied Slices	95	3,584	2%
Number of slices containing only related logic	95	95	100%
Number of slices containing only unrelated logic	0	95	0%
Total Number of 4 input LUT's	171	7,168	2%
Number of bonded IOBs	33	141	23%
Average Fanout of Non-Clock Nets	3.40	-	-

In a study referenced as [1], detailed in the research paper titled "Performance Comparison of Vedic Multiplier and Booth Multiplier," the authors emphasize the critical importance of reducing power consumption and delay in processors to enhance overall performance. The research highlights the need for optimizing multiplier architecture to achieve efficient processor design. By implementing both Vedic Multiplier and Booth Multiplier on FPGA platforms, the study conducts a comparative analysis to evaluate their respective power and delay characteristics.

The methodology in [1] involves implementing the Vedic Multiplier and Booth Multiplier using VHDL in Modelsim 10.1b, followed by synthesizing the designs in Xilinx 8.2i ISE. This closely aligns with our paper's approach, enabling a direct comparison between their findings and our experimental results. Table 2.5.2 in their study compares the area, delay, and power consumption of both Booth and Urdhava Tiryakbhyam (Vedic) multipliers. The results indicate that the Urdhava Tiryakbhyam multiplier outperforms the Booth multiplier in terms of area, delay, and power consumption.

Following are the results given by the authors in their research paper:

Table III: Results found by Authors in [1].

FPGA Device Package (XC2550 TQ144)	Area in LUT's	Delay (ns)	Speed (MHz)	Memory (Kb)	Power (mW)
Booth's Multiplier	9	11.176	89.477	156440	11.30
Urdhava Multiplier	86	8.460	118.203	166744	9.29

METHODOLOGY AND WORKING

B. METHODOLOGY

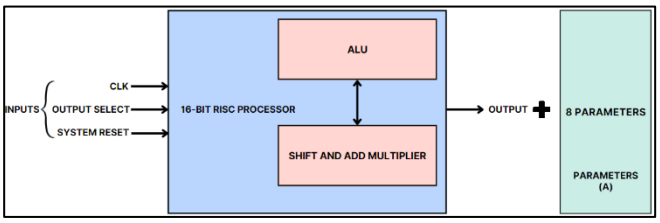


Fig. 5. Block Diagram of 16-Bit RISC Processor with 8 x 8 Vedic multiplier, Inputs, Outputs, and Comparison Parameters.

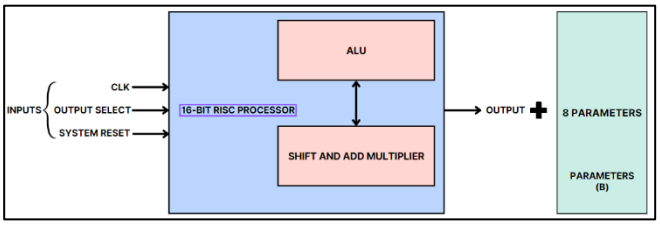


Fig. 6. Block Diagram of 16-Bit RISC Processor with Shift and Add multiplier, Inputs, Outputs, and Comparison Parameters.

The block diagram of the methodology for our project "Performance analysis of 16-bit RISC processor using Vedic multiplier" illustrates the operational flow and integration of components within the processor.

The 16-bit RISC processor is driven by a clock signal and a system reset input. The clock signal provides timing synchronization for all processor operations, while the system reset (active-low) initializes the processor state. When the clock signal is active and the system reset is low, the processor activates. The Program Counter (PC) determines the address of the next instruction in the program memory. Instructions stored in the program memory are fetched and loaded into the Instruction Register (IR). The IR holds the current instruction being executed.

The instruction stored in the IR is decoded by the Control Unit. The Control Unit interprets the instruction opcode to determine the corresponding operation to be executed. Based on the decoded instruction, the Control Unit directs the processor to execute specific arithmetic or logical operations using data stored in the General-Purpose Registers (GPRs) or memory. Within the set of instructions handled by the Control Unit, there exists a specific instruction for multiplication (mul). When this instruction is encountered during execution, it triggers the Vedic multiplier module / Shift and Add multiplier module. The Vedic multiplier module / Shift and Add multiplier module performs multiplication operations on the specified operands, leveraging its unique design for efficient and rapid computation.

The operands required for the multiplication operation are selected from the GPRs or memory locations based on the instruction's addressing mode. Upon completion of instruction execution, the results of arithmetic or logical operations can be accessed based on the output select input provided to the processor. Post-synthesis and implementation using Xilinx Vivado, the processor's performance is evaluated through various parameters:

- Worst Negative Slack (WNS): Measures the critical path delay, indicating the longest delay in meeting timing requirements. WNS = Data Required Time – Data Arrival Time.

Register to Register Setup Slack:

Arrival Time = Tck->q + Combo Delay

Required Time = Clock Period - Setup Time

Setup Slack = Required Time - Arrival Time

Setup Slack = Clock Period - Setup Time - Tck->q – Combo Delay

- Total Negative Slack (TNS): Sum of all negative slack values, reflecting overall timing constraints.
- Worst Hold Slack (WHS): Represents the minimum hold time margin for timing stability. WHS = Data Arrival Time – Data Required Time.

Register to Register Hold Slack:

Required Time = Hold Time

Arrival Time = Tck->q + Combo Delay

Hold Slack = Arrival Time - Required Time

Hold Slack = Tck->q + Combo Delay - Hold Time

- Total Hold Slack (THS): Aggregate hold time margin across all timing paths.
- Total Path Worst Slack (TPWS): Quantifies the most critical timing path within the design.
- Total Power Consumption: Reflects the dynamic and static power consumption of the synthesized design.
- Number of Look-Up Tables (LUTs) and Flip-Flops (FFs): Indicate the hardware resources utilized in the design.

The performance metrics obtained from both Vedic multiplier and shift-and-add multiplier implementations are compared. This comparative analysis evaluates their respective strengths and weaknesses in terms of timing, power consumption, and hardware utilization. The goal is to identify which multiplier architecture (Vedic or shift-and-add) offers superior performance characteristics for arithmetic operations within the 16-bit RISC processor.

In conclusion, this methodology integrates the 16-bit RISC processor with a Vedic multiplier to explore enhancements in arithmetic performance and efficiency. By leveraging Xilinx Vivado tools for synthesis and analysis, the paper aims to provide insights into optimal design choices for multiplication operations within resource-constrained environments. The comparative analysis facilitates informed decision-making regarding multiplier selection to optimize overall processor performance.

C. WORKING OF 16-bit RISC PROCESSOR:

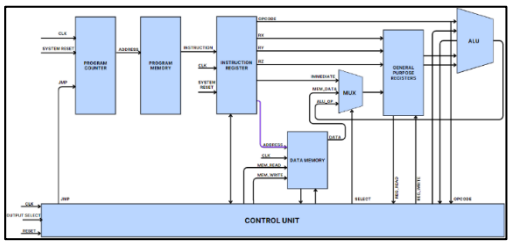


Fig. 7. Block diagram of 16-bit RISC processor

The block diagram of the 16-bit RISC processor designed for this project encompasses various interconnected components to enable efficient data processing and control.

The Program memory unit can store 16 instructions, each 16 bits wide, serving as the primary storage for program instructions. 16-bit wide instruction register holds the currently fetched instruction from the program memory for decoding and execution. The control unit interprets instructions stored in the instruction register, generating control signals to coordinate operations within the processor components. A bank of 16 registers, each capable of holding 16-bit data, provides temporary storage for operands and intermediate results during computation.

The ALU performs a range of arithmetic and logical operations (14 operations in total) on data from the register bank based on control signals from the control unit. Notably, the ALU includes a specialized multiplier module that is instantiated when a multiplication instruction is executed. When a multiplication instruction is decoded, the multiplier module within the ALU is activated. This module performs multiplication operations efficiently, leveraging the Vedic multiplier technique for enhanced computational speed and efficiency. Similar to program memory, the data memory unit stores 16 lines of data, each 16 bits wide, providing storage for variables and data utilized during program execution. The input buffer facilitates data transfer from the input pins of the FPGA board into the processor. Conversely, the output buffer enables the transfer of processed data or results from the processor to the output pins of the FPGA board.

The processor's operation follows a sequential processing model: instructions are fetched from the program memory into the instruction register, decoded by the control unit, and executed by the ALU. The general-purpose register bank acts as a workspace for data manipulation, while the input and output buffers facilitate efficient data exchange with the input and output pins of the FPGA board.

Table IV: Instruction Register.

16-BIT INSTRUCTION REGISTER			
15 - 12 bits	11 - 8 bits	7 - 4 bits	3 - 0 bits
Operation Type	Destination Register	Source Register 1	Source Register 2

The 16-bit Instruction Register (IR) plays a crucial role in decoding and executing instructions within the 16-bit RISC processor architecture. The IR is structured to accommodate various components essential for instruction execution. Specifically, the IR is filled with a 16-bit instruction fetched from the program memory, to which the program counter (PC) points during instruction fetch.

The IR is divided into distinct fields:

Operation Code (Opcode) Field (Bits 15 to 12): The last 4 bits of the IR are dedicated to encoding the operation to be

performed by the Arithmetic Logic Unit (ALU). These bits select one of the 14 different operations supported by the ALU.

Destination Register Address Field (Bits 11 to 8): The subsequent 4 bits of the IR specify the address of the destination General Purpose Register where the result of the operation will be stored after execution.

Source Register 1 Address Field (Bits 7 to 4): The next 4 bits of the IR indicate the address of the first source General Purpose Register 1(rsrc1), representing Operand 1 used in the specified operation. This field identifies the register holding the first operand for the ALU operation.

Source Register 2 Address Field (Bits 3 to 0): The final 4 bits of the IR specify the address of the second source General Purpose Register 2 (rsrc2), denoting Operand 2 required for the ALU operation. This field identifies the register containing the second operand used in the specified operation.

During instruction execution, the control unit decodes the contents of the IR to determine the specific operation to be executed by the ALU, along with the associated operands and destination register.

Table V: Instruction Set.

OPERATION TYPE	OPCODE
LOAD WORD (LW)	0000
STORE WORD (SW)	0001
ADDITION (ADD)	0010
SUBTRACTION (SUB)	0011
MULTIPLICATION (MUL)	0100
LOGICAL AND (AND)	0101
LOGICAL OR (OR)	0110
LOGICAL NOT (NOT)	0111
LOGICAL EX-OR (XOR)	1000
LOGICAL NAND (NAND)	1001
LOGICAL NOR (NOR)	1010
LOGICAL EX-NOR (XNOR)	1011
JUMP (JUMP)	1100
HALT (HALT)	1101

Load Word (LW): Used to load a word from data memory into a general-purpose register.

Store Word (SW): Used to store a word from a general-purpose register into a data memory.

Addition (ADD): Performs addition of two operands.

Subtraction (SUB): Performs subtraction of two operands.

Multiplication (MUL): Initiates multiplication of two operands.

Logical and (AND): Performs bitwise AND operation between two operands.

Logical or (OR): Performs bitwise OR operation between two operands.

Logical not (NOT): Performs bitwise NOT operation on a single operand.

Logical ex-or (XOR): Performs bitwise XOR (exclusive OR) operation between two operands.

Logical nand (NAND): Performs bitwise NAND operation between two operands.

Logical nor (NOR): Performs bitwise NOR operation between two operands.

Logical ex-nor (XNOR): Performs bitwise XNOR (exclusive NOR) operation between two operands.

Jump (JUMP): Changes the program counter to jump to a specified memory address.

Halt (HALT): Stops the processor's operation, effectively halting further execution.

These opcode assignments enable the control unit of your processor to decode instructions fetched from memory (stored in the Instruction Register) and determine the appropriate operation to be executed by the Arithmetic Logic Unit (ALU) or other processor components. Each instruction is represented by a 16-bit word, with the opcode field specifying the type of operation and additional fields indicating operand addresses or immediate values as required by the instruction.

Table VI: Data Given to Program Memory

Sr no	INSTRUCTION	HEXADECIMAL FORM
1	0000_0000_0000_0000	0000
2	0000_0001_0001_0000	0110
3	0000_1111_1111_0000	0FF0
4	0000_1110_1110_0000	0EE0
5	0010_0010_0000_0001	2201
6	0001_0010_0010_0000	1220
7	0100_1101_1111_1110	4DFE
8	0001_1101_1101_0000	1DD0
9	0011_0011_1110_1111	33EF
10	0001_0011_0011_0000	1330
11	0101_0100_1110_1111	54EF
12	0001_0100_0100_0000	1440
13	0110_0101_1110_1111	65EF
14	0001_0101_0101_0000	1550
15	1000_0110_1110_1111	86EF
16	1101_0000_0000_0000	D000

Instruction 1: 0000_0000_0000_0000 (0000): In this instruction GPR [0] is loaded with data from data memory [0] which is 16-bit hexadecimal 0002.

Instruction 2: 0000_0001_0001_0000 (0110): In this instruction GPR [1] is loaded with data from data memory [1] which is 16-bit hexadecimal 0003.

Instruction 3: 0000_1111_1111_0000 (0FF0): In this instruction GPR [15] is loaded with data from data memory [15] which is 16-bit hexadecimal 00AB.

Instruction 4: 0000_1110_1110_0000 (0EE0): In this instruction GPR [14] is loaded with data from data memory [14] which is 16-bit hexadecimal 00CD.

Instruction 5: 0010_0010_0000_0001 (2201): In this instruction we perform addition operation on GPR [0] & GPR [1] and the result is stored in destination register GPR [2].

Instruction 6: 0001_0010_0010_0000 (1220): In this instruction data is stored from GPR [2] to data memory [2] which is the result of the addition operation.

Instruction 7: 0100_1101_1111_1110 (4DFE): This is the instruction when the vedic and shift and add multipliers are instantiated. In this instruction we perform multiplication operation on GPR [15] and GPR [14] and the result is stored in destination register GPR [13].

Instruction 8: 0001_1101_1101_0000 (1DD0): In this instruction data is stored from GPR [13] to data memory [13] which is the result of the multiplication operation.

Instruction 9: 0011_0011_1110_1111 (33EF): In this instruction we perform subtraction operation on GPR [14] and GPR [15] and the result is stored in destination register GPR [3].

Instruction 10: 0001_0011_0011_0000 (1330): In this instruction data is stored from GPR [3] to data memory [3] which is the result of the subtraction operation.

Instruction 11: 0101_0100_1110_1111 (54EF): In this instruction we perform bitwise and operation on GPR [14] and GPR [15] and the result is stored in destination register GPR [4].

Instruction 12: 0001_0100_0100_0000 (1440): In this instruction data is stored from GPR [4] to data memory [4] which is the result of the bitwise and operation.

Instruction 13: 0110_0101_1110_1111 (65EF): In this instruction we perform bitwise or operation on GPR [14] and GPR [15] and the result is stored in destination register GPR [5].

Instruction 14: 0001_0101_0101_0000 (1550): In this instruction data is stored from GPR [5] to data memory [5] which is the result of the bitwise or operation.

Instruction 15: 1000_0110_1110_1111 (86EF): In this instruction we perform bitwise xor operation on GPR [14] and GPR [15] and the result is stored in destination register GPR [6].

Instruction 16: 1101_0000_0000_0000 (D000): In this instruction we perform halt operation which indicates the completion of execution of all instructions.

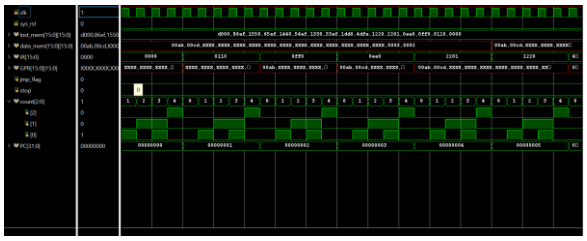


Fig. 8. Instructions going Inside the Instruction Register as the Program Counter Increments.

As illustrated in Figure 3.1.2, the program counter continuously increments its value, advancing to the subsequent instructions stored in the program memory. This process is crucial for the sequential execution of a program. Each time the program counter increments, it points to the next instruction, which is then loaded into the 16-bit Instruction Register (IR). The IR temporarily holds the instruction that the program counter is addressing, allowing the processor to decode and execute it.

To ensure that each instruction is executed thoroughly before fetching the next one, a 3-bit counter is employed to introduce a brief delay. This counter provides a delay of 5 counts, which is sufficient to guarantee that the current instruction has been fully executed before the next instruction is loaded into the IR. This delay is critical in synchronous digital circuits, where precise timing ensures proper operation and prevents errors due to overlapping instruction execution.

The sequence begins with the program counter pointing to the initial instruction in the program memory. This instruction is then fetched and loaded into the IR. During the execution phase, the 3-bit counter starts counting to provide the necessary delay. Once the counter completes its 5 counts, indicating that the current instruction has been executed, the program counter increments, and the process repeats for the next instruction.

This mechanism ensures a smooth and orderly execution of instructions, with each step meticulously timed to avoid conflicts and ensure reliable operation. The use of the 3-bit counter for delay demonstrates the importance of timing control in digital systems, particularly in processors where precise instruction execution is paramount. By incorporating this delay, the system can handle instructions sequentially, maintaining the integrity and accuracy of the operations being performed. WORKING OF 8 X 8 VEDIC MULTIPLIER:

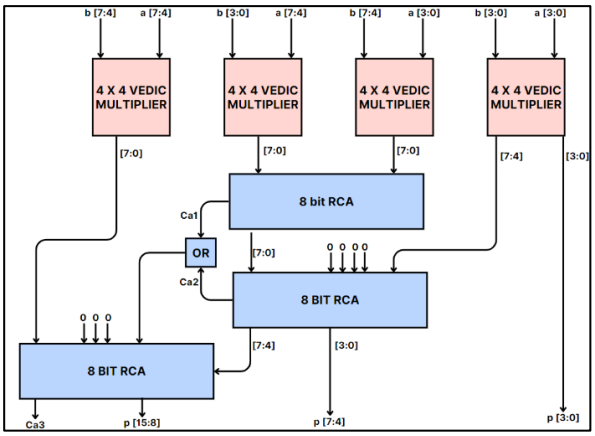


Fig. 9. Block diagram of 8 x 8 vedic multiplier.

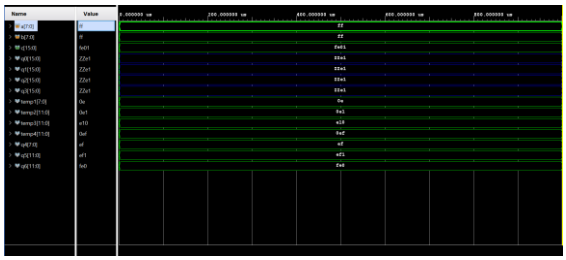


Fig. 10. Simulation of 8 x 8 Vedic multiplier showing the result of two 2-digit hexadecimal numbers ff and ff.

The 8x8 Vedic multiplier is considered a simple combinational circuit due to its inherent design characteristics, which enable it to perform multiplication operations without the need for clock signals. Unlike sequential circuits that rely on clock pulses for synchronization and state transitions, combinational circuits like the Vedic multiplier execute operations based solely on the current input values, producing immediate outputs in response. As shown in Figure 3.2.3, the Verilog code for the 8x8 Vedic multiplier was successfully simulated. In the simulation, the multiplication of two 2-digit hexadecimal numbers “ff” multiplied by “ff” was performed, yielding the result “fe01”.

The Verilog module defines an instruction decoder (decode_inst) responsible for interpreting the type of operation (oper_type) extracted from the instruction register (IR). Each instruction type (LW, SW, add, sub, mul, and, or, not, xor, nand, nor, xnor, jump, halt) corresponds to a distinct case within the instruction decoder. Specifically, for the multiplication (mul) operation, the operands (rsrc1 and rsrc2) are assigned to temporary variables (temp1 and temp2) to prepare them as inputs for the 8x8 Vedic multiplier.

The eight_x_eight_vedic_mul module is instantiated (u0) within the top-level module (top_vedic_RCA), connecting its input (ip1 and ip2) and output (op) ports. The operands (temp1 and temp2) identified during the mul operation are fed into the Vedic multiplier (u0), which computes the product (op) of the two 8-bit numbers.

The results of arithmetic operations (add, sub, mul, and, or) or immediate data (LW, SW) are assigned to the dout output register based on specific control signals (add_op, sub_op, mul_op, and_op, or_op). When the mul_op control signal is asserted, the output from the 8x8 Vedic multiplier (op) is directly assigned to dout, representing the multiplication result. The program memory (inst_mem) is initialized with instructions using the $readmemb function, which loads the sequence of operations to be executed by the processor. The processor reads these instructions sequentially, incrementing the program counter (PC) to load each instruction into the instruction register (IR), thereby triggering the instruction decoding process (decode_inst).

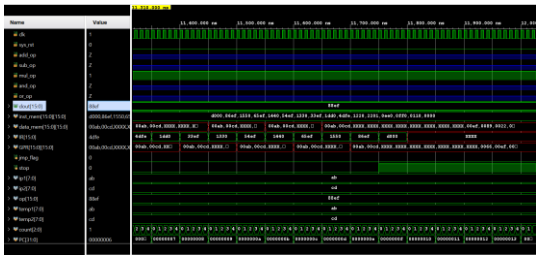


Fig. 11. Simulation of Processor integrated with 8 x 8 vedic multiplier when sys_rst=0 and mul_op=1.

When the system reset (sys_rst) signal is low and the multiplication operation (mul_op) signal is high, it indicates a specific condition where the system is ready to perform a multiplication this scenario triggers the following actions:

The processor module detects the high state of mul_op during a normal operational state (after reset). This condition signals the processor to initiate a multiplication instruction. Inside the processor module, the mul operation case in the decode_inst task is activated. This prepares the necessary operands (temp1 and temp2) by assigning them the values from the source registers (GPR[rsrc1] and GPR[rsrc2]). The vedic multiplier module begins the multiplication operation using the provided operands (ip1 and ip2). The result of the multiplication (op) is calculated and made available once the operation.

D. WORKING OF SHIFT AND ADD MULTIPLIER:

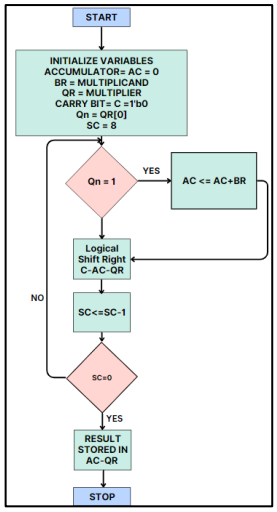


Fig. 12. Algorithm for Shift and Add multiplier.

The shift and add multiplication algorithm, also known as the multiply-and-accumulate method, is a fundamental approach used to perform multiplication operations using only basic arithmetic operations such as addition and shifting (rightward). This algorithm is particularly useful for implementing multiplication in hardware circuits where complex multiplication units may not be feasible or efficient. It allows for straightforward multiplication by leveraging simple and easily implemented operations, making it ideal for digital systems with limited resources.

The algorithm begins with several initializations. The first step is to initialize the Accumulator (AC) to 0, which will be used to store intermediate and final results. The two operands involved in the multiplication process are the multiplier (QR) and the multiplicand (BR). Additionally, a carry bit (C) is initialized to 0. The least significant bit (LSB) of the multiplier QR is denoted as Qn. A counter (SC) is also initialized to be equal to the number of bits (n) of both the multiplier and the multiplicand. This counter will determine the number of iterations needed to complete the multiplication process.

The core of the algorithm is a loop that iterates n times, where n is the bit-width of the operands. In each iteration, the algorithm checks the value of the least significant bit (Qn) of the multiplier (QR). If Qn is equal to 1, the multiplicand (BR) is added to the Accumulator (AC). This addition operation accounts for the contribution of the current bit of the multiplier to the final product. After the addition, a logical shift right operation is performed on the carry bit (C), the Accumulator (AC), and the multiplier (QR). This shift operation effectively divides the combined value of AC and QR by 2, preparing the system for the next bit's contribution.

After the logical shift right operation, the counter (SC) is decremented by 1. The algorithm then checks if the counter has reached zero. If SC is zero, the process terminates, and the 2n-bit result is stored in the Accumulator (AC) and the multiplier (QR). This combined value represents the final product of the two operands. If SC is not zero, the algorithm repeats the loop, starting from the check of the least significant bit of the multiplier.

If the least significant bit (Qn) of the multiplier (QR) is equal to 0, the algorithm bypasses the addition step and directly performs the logical shift right operation on the carry bit, Accumulator (AC), and multiplier (QR). This shift operation ensures that the system correctly handles the scenario where the current bit of the multiplier does not contribute to the product. After shifting, the counter (SC) is decremented, and the algorithm checks if it has reached zero. If SC is zero, the final product is stored as before. If not, the algorithm repeats the process from the initial step of checking Qn.

By repeating these steps, the shift and add multiplication algorithm effectively computes the product of the two operands. Each iteration processes one bit of the multiplier, adding the appropriately shifted multiplicand to the Accumulator when necessary and shifting the combined value of the Accumulator and multiplier rightward. This method, while simple, provides a robust and efficient way to perform multiplication in hardware, especially in environments with limited computational resources. The straightforward nature of the algorithm makes it a fundamental technique in digital design, illustrating the power of basic arithmetic operations in achieving complex computational tasks.

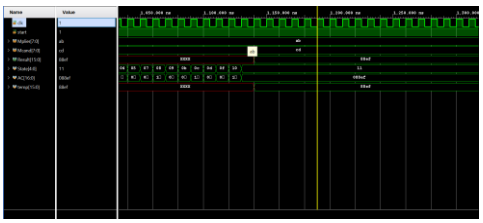


Fig. 13. Simulation of 8 x 8 Shift and Add multiplier showing the result of two 2-digit hexadecimal numbers ab and cd.

The Verilog module defines an instruction decoder (decode_inst) that plays a crucial role in interpreting the operation type (oper_type) extracted from the instruction register (IR). This interpretation is fundamental as it determines the specific arithmetic or logical operations to be executed using operands stored in general-purpose registers (GPR). The instruction decoder essentially functions as the brain of the processor, ensuring that each instruction is correctly understood and executed.

Each instruction type, such as load word (LW), store word (SW), addition (add), subtraction (sub), multiplication (mul), and various logical operations (and, or, not, xor, nand, nor, xnor), as well as control operations like jump and halt, corresponds to a distinct case within the instruction decoder. For the multiplication (mul) operation, the operands (rsrc1 and rsrc2) are assigned to temporary variables (temp1 and temp2). These variables are then used to set the t_start signal, which initiates the multiplication process.

Within the top-level module (top_shift_add_new), the shift_add_mul_new module is instantiated as (u0). This module connects its input ports (ip1 and ip2) to the inputs of the multiplier module, labeled 'Mplier' and 'Mcand,' respectively, and its output port (op) to the output of the multiplier module, labeled 'Result.' During the multiplication operation, the operands (temp1 and temp2) extracted from the general-purpose registers are fed into the Shift and Add Multiplier (u0). This module computes the product (op) of the two 8-bit numbers, temp1 and temp2, using the shift and add algorithm.

The result of various arithmetic operations (add, sub, mul) or logical operations (and, or), as well as immediate data instructions (LW, SW), is assigned to the dout output register based on specific control signals (add_op, sub_op, mul_op, and_op, or_op). When the mul_op control signal is asserted, indicating a multiplication operation, the output from the Shift and Add Multiplier (op) is directly assigned to dout. This direct assignment ensures that the multiplication result is correctly represented in the output register.

The program memory (inst_mem) is initialized with instructions using the $readmemb function. This initialization defines the sequence of operations that the processor will execute. The processor reads these instructions sequentially, with the program counter (PC) incrementing to point to the next instruction. Each instruction is loaded into the instruction register (IR), which triggers the instruction decoding process handled by the decode_inst module.

This structured approach allows the processor to systematically execute a wide range of operations, from simple arithmetic to complex logical functions, ensuring that each step is meticulously carried out as per the encoded instructions. The shift and add multiplication method, integrated into this system, highlights the elegance and efficiency of using fundamental arithmetic techniques to achieve reliable computational results in a hardware environment.

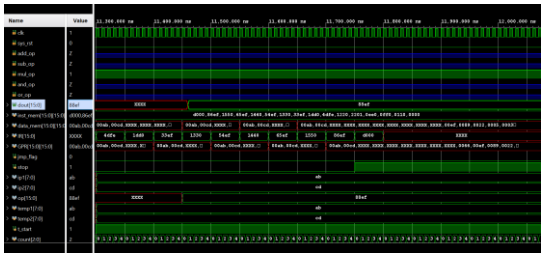


Fig. 14. Simulation of Processor integrated with shift add multiplier when sys_rst=0 and mul_op=1.

When the system reset (sys_rst) signal is low and the multiplication operation (mul_op) signal is high, it indicates a specific condition where the system is ready to perform a multiplication this scenario triggers the following actions:

The processor module detects the high state of mul_op during a normal operational state (after reset). This condition signals the processor to initiate a multiplication instruction. Inside the processor module, the mul operation case in the decode_inst task is activated. This prepares the necessary operands (temp1 and temp2) by assigning them the values from the source registers (GPR[rsrc1] and GPR[rsrc2]).

Additionally, the t_start signal is asserted (t_start <= 1'b1), which serves as a trigger to start the multiplication process in the shift and add multiplier module (shift_add_mul_new). With t_start set high, the shift and add multiplier module begins the multiplication operation using the provided operands (ip1 and ip2). The result of the multiplication (op) will be calculated and made available once the operation completes.

RESULTS

Table VII: Results for Processor integrated with 8 × 8 vedic multiplier.

PARAMETERS	8 X 8 VEDIC MULTIPLIER	8 X 8 SHIFT AND ADD MULTIPLIER
WNS (ns)	8.161	6.050
TNS (ns)	0	0
Total Setup Delay (ns)	1.833	3.907
WHS (ns)	0.197	0.182
THS (ns)	0	0
Total Hold Delay (ns)	0.315	0.269
TPWS (ns)	0	0
TOTAL POWER (W)	0.044	0.045
Number of LUT's	373	307
Number of FF's	27	65

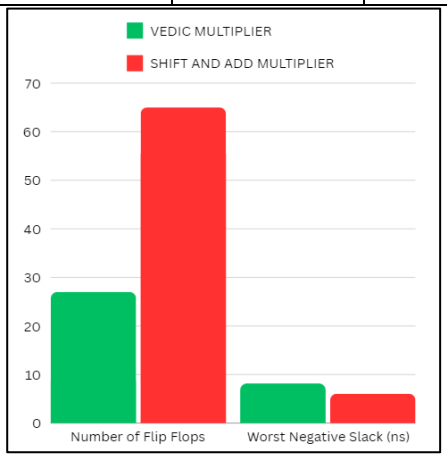


Fig. 15. Bar Graph Comparing Number of flipflops and Worst Negative Slack

The integration of the processor with the Vedic multiplier yields a higher Worst Negative Slack (WNS) compared to the shift and add multiplier. This higher WNS signifies improved setup delay performance, indicating effective meeting of timing requirements. However, achieving superior setup delay performance with the Vedic multiplier comes at the expense of increased area and power consumption. While optimizing setup delay is advantageous, balancing all aspects of digital design performance, area, and power consumption remains a challenge in this scenario.

Flip-flops play a crucial role in determining the critical path delay and overall timing performance of a digital circuit. A higher number of flip-flops typically correlates with a longer critical path and potentially slower timing characteristics due to increased logic depth and routing complexity. The higher WNS for the Vedic multiplier can be attributed to the reduced number of flip-flops. With fewer flip-flops, the critical path delay in the Vedic multiplier design is shorter, leading to a larger negative slack value. Negative slack indicates how far a signal's arrival time is from meeting the required timing constraints.

Conversely, integrating the processor with the shift and add multiplier results in a lower Worst Negative Slack (WNS) compared to the Vedic multiplier. A lower WNS suggests challenges in meeting timing requirements, potentially leading to setup violations and degraded performance. Despite this concern, the shift and add multiplier exhibits advantages in power consumption and area utilization. Consuming a total power of 0.045 W and utilizing fewer resources (307 Look-Up Tables and 65 Flip-Flops compared to the Vedic multiplier's 373 LUTs), the shift and add multiplier offers trade-offs where timing requirements are compromised for improved power efficiency and reduced area utilization. This highlights the nuanced considerations involved in digital design optimization.

As for the other parameters, both the multipliers have Total negative slack (TNS) of 0ns which means all the paths in both the designs are meeting their timing requirements. A Worst Hold Slack (WHS) of zero or positive means that all data inputs are stable for at least the required hold time relative to the clock edge, ensuring proper data capture. Similarly, a Total Hold Slack (THS) of zero or positive across all paths indicates that the hold time constraints are universally met in the design. A Total Pulse Width Slack (TPWS) of zero or positive ensures that all signals maintain their required pulse width (duration) without any violations.



Fig. 16. Implementation on FPGA Board.

As you can see in the above figure 4.1.1 the when the first slide switch which is assigned to sys_rst is turned off (0) and the mul_op switch is turned on (1) the output we get to see on the output LED's is the result of the multiplication of the two hexadecimal numbers AB and CD which is 88EF.

CONCLUSION

In conclusion, the paper focused on the performance analysis of a 16-bit RISC processor integrated with two different multiplier designs: the Vedic multiplier and the shift and add multiplier. The paper involved implementing these multiplier architectures using Verilog on an FPGA board and conducting a comparative analysis based on performance metrics.

Through simulation and synthesis, we evaluated key performance indicators such as Worst Negative Slack (WNS) to assess the timing behavior of each integration. The results showed that while the Vedic multiplier integration exhibited a higher WNS, indicating better setup delay performance, it also consumed more area and power compared to the shift and add multiplier. On the other hand, the shift and add multiplier integration demonstrated a lower WNS, suggesting potential timing challenges, but offered advantages in terms of reduced area utilization and power consumption.

These findings underscore the trade-offs inherent in digital design—balancing performance, area, and power efficiency to achieve optimal system design. Depending on project requirements and priorities, different multiplier architectures may be preferred. The Vedic multiplier excels in timing performance but at the expense of increased resource utilization, while the shift and add multiplier prioritizes efficiency in area and power consumption.

REFERENCES

[1] Anju. "Performance Comparison of Vedic Multiplier and Booth Multiplier." International Journal of Electronics and Communication Engineering & Technology, vol. 2, no. 5, June 2013, pp. 336-339. BEIESP, ISSN: 2249-054X. [Retrieval Number: E1846062513/2013©BEIESP]

[2] Yadav, Ankita & Bendre, Varsha. (2021). Design and Verification of 16 bit RISC Processor Using Vedic Mathematics. 759-764. 10.1109/ESCI50559.2021.9396965.

[3] Kumawat, Jyoti, and Sunil Sharma. "A 8x8 bit multiplier using Vedic Mathematics." International Journal of Engineering and Technical Research (IJETR), vol. 2, no. 3, March 2014, ISSN: 2321-0869.

[4] Chen Ping-hua and Zhao Juan, "High-Speed Parallel 32x32-bit Multiplier Using a Radix-16 Booth Encoder", Third International Symposium on intelligent Information Technology Application Workshops, 2009. IITAW 09, pp.406-409, 21-22 Nov. 2009.

[5] Swami Bharati Krshna Tirthaji, Vedic Mathematics. Delhi: Motilal Banarsidass Publishers, 1965.

[6] Asmita Haveliya "A Novel Design for High Speed Multiplier for Digital Signal Processing Applications (Ancient Indian Vedic mathematics approach)" International Journal of Technology And Engineering System(IJTES):Jan - March 2011- Vo12 .No. l.

[7] Harpreet Singh Dhillon and Abhijit Mitra "A Digital Multiplier Architecture using Urdhava Tiryakbhyam Sutra of Vedic Mathematics" IEEE Conference Proceedings,2008.

[8] Parth Mehta, Dhanashri Gawali "Conventional versus Vedic mathematical method for Hardware implementation of a multiplier" 2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies.

[9] Pushpalata Verma, K.K. Mehta, "Implementation of an efficient multiplier based on Vedic Mathematics using EDA Tool", International Journal of Engineering and Advance Technology (IJEAT) ISSN : 2249-8958, volume-1, Issue -5, June 2012.

[10] Gaonkar, Supraj and M. P. Anitha. "Design Of 16-bit RISC Processor." International journal of engineering research and technology 2 (2013): n. pag.

❑ ❑ ❑ ❑ ❑ ❑

AET 015

Cricket Data Analytics: IPL Score Prediction

Prof. Uma Jaishankar
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
(Affiliation to Mumbai University)
Mumbai, India
uma.jaishankar@vit.edu.in

Rupesh Khanzode
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
(Affiliation to Mumbai University)
Mumbai, India
rupesh.khanzode@vit.edu.in

Sakshi Bhosale
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
(Affiliation to Mumbai University)
Mumbai, India
sakshi.bhosale@vit.edu.in

***Abstract*— Cricket, notably the Indian Premier League (IPL), stands as one of the most captivating sports globally, attracting millions of viewers and enthusiasts. With the unpredictability inherent in cricket matches, predicting outcomes and estimating final scores poses a fascinating challenge. In this paper, we introduce the IPL Score Predictor, a novel web application meticulously crafted to harness the power of historical match data and advanced machine-learning techniques for forecasting final scores in IPL matches. By amalgamating historical match statistics, player performances, and contextual factors, the IPL Score Predictor provides cricket enthusiasts and analysts with invaluable insights into ongoing matches. Through a user-friendly interface, the application empowers users to make informed assessments of match dynamics and potential outcomes, thereby enhancing their engagement and enjoyment of IPL matches.**

***Keywords*—Indian Premier League (IPL), Cricket Score Prediction, Machine Learning, Web Application, Predictive Analytics**

I. INTRODUCTION

Cricket, often described as a religion in India, commands a massive following worldwide, with the IPL emerging as a cornerstone of the sport's popularity. The IPL, characterized by high-intensity matches between franchise teams, captivates audiences with its unpredictability and thrilling performances. Predicting match outcomes and estimating final scores in such a dynamic environment presents a compelling challenge, driving the need for innovative solutions that leverage data and technology.

Cricket enthusiasts worldwide are eager to gain deeper insights into match dynamics and potential outcomes, fuelling the demand for accurate prediction tools like the IPL Score Predictor. Through cutting-edge machine learning techniques, our project seeks to unlock hidden patterns within the vast expanse of cricket data, empowering fans, and analysts alike with actionable intelligence. By bridging the gap between data science and cricket fandom, the IPL Score Predictor promises to revolutionize how cricket matches are experienced and analyzed.

II. LITERATURE REVIEW

Cricket score prediction has garnered significant attention from researchers and practitioners in recent years, with various studies exploring different approaches and techniques to forecast match outcomes and player performances. This section provides a review of relevant literature in the field of cricket score prediction, focusing on machine learning approaches and predictive modelling techniques.

A. Machine Learning Approaches

Several studies have utilized machine learning algorithms to predict cricket match outcomes and player performances. Harsh Mittal et al. (2021) conducted a study on machine learning approaches for player performance and match results prediction in cricket. They explored various techniques for predicting match outcomes, player selections using machine learning and artificial intelligence. Their findings highlighted the importance of feature selection, data preprocessing, and model selection in achieving accurate predictions [3].

B. Dynamic Match Outcome Prediction

Rajesh Goel et al. (2021) proposed a model for dynamic cricket match outcome prediction, where the match outcome is predicted ball by ball during the second inning. Their methodology incorporated real-time game context and relative team strengths to predict match outcomes accurately. By analyzing a dataset spanning multiple seasons, they demonstrated the effectiveness of their approach in predicting match outcomes with high accuracy. Their study laid the foundation for real-time match prediction systems and highlighted the importance of incorporating dynamic game context in predictive modelling [4].

C. Data Mining Techniques

Rana Kashif (2020) conducted a comparative analysis of data mining techniques to predict projected scores and winners in Twenty-20 cricket matches. They applied multiple data mining techniques, including multiple linear regression and classification algorithms, to forecast batting team scores and predict match winners. Their research emphasized the importance of feature selection, data preprocessing, and model evaluation in achieving accurate predictions. Additionally, they highlighted the potential of advanced classification techniques such as support vector machines and random forests in improving prediction accuracy [5].

D. Summary

Overall, the literature review highlights the diverse range of machine-learning approaches and predictive modelling techniques applied to cricket score prediction. While some studies focus on predicting match outcomes and player performances using historical data, others explore dynamic prediction models that adapt to real-time game context. By leveraging advanced data mining techniques and machine learning algorithms, researchers have made significant strides in enhancing the accuracy and reliability of cricket score prediction systems. These insights provide valuable guidance for the development of the IPL Score Predictor and underscore the importance of feature engineering, model selection, and performance evaluation in predictive modelling.

III. PROPOSED METHODOLOGY

The proposed methodology outlines the systematic approach adopted for developing the IPL Score Predictor, emphasizing data processing, predictive modelling, and user interface design.

A. Data Collection and Preprocessing

- **Data Collection:** Historical IPL match data, including match statistics, player performances, and contextual official IPL databases and cricket statistics websites.
- **Data Preprocessing:** The collected data undergoes preprocessing to clean, transform, and prepare it for analysis. This involves handling missing values, encoding categorical variables, and extracting relevant features for predictive modelling. Exploratory data analysis techniques are employed to gain insights into the dataset's characteristics and identify potential patterns.

B. Predictive Modeling

- **Feature Engineering:** Feature engineering techniques are applied to create new features and transform existing ones to improve model performance. This involves extracting meaningful features from the data, including match statistics, player attributes, and contextual factors.
- **Model Development:** Various machine learning algorithms, such as decision tree regressors, random forest regressors, and support vector regressors, are trained on the pre-processed data to develop predictive models for estimating cricket scores. Model selection is based on performance metrics such as accuracy, mean absolute error, and root mean squared error.

C. User Interface Development

- **Designing the Interface:** An intuitive and user-friendly web interface is designed to allow users to input current match information and view score predictions. The interface features input fields for match details such as batting team, bowling team, overs bowled, runs scored, wickets fallen, and runs scored in the last five overs.
- **Real-time Updates:** Mechanisms for providing real-time score predictions during IPL matches are implemented to enable users to stay informed about ongoing matches and make timely decisions. Streaming data processing techniques are employed to handle real-time updates and incorporate live match data into the prediction models.

D. Evaluation and Validation

- **Model Evaluation:** The performance of the predictive models is evaluated using various metrics such as accuracy, mean absolute error, and root mean squared error. Cross-validation techniques are employed to ensure the robustness and generalization capability of the models.
- **User Feedback:** User feedback and usability testing are conducted to assess the effectiveness and user satisfaction of the IPL Score Predictor. Iterative improvements are made based on user input to enhance the application's usability and functionality.

E. Deployment and Integration

- **Deployment Infrastructure:** The trained prediction models are deployed on scalable and reliable cloud infrastructure such as AWS, Azure, or Google Cloud Platform. APIs or microservices are developed to handle prediction requests in real time.
- **Integration with External Platforms:** Opportunities for integrating the prediction tool with external platforms such as fantasy cricket apps, sports betting platforms, and broadcasting channels are explored to extend its reach and utility. APIs are developed to facilitate seamless integration with external systems.

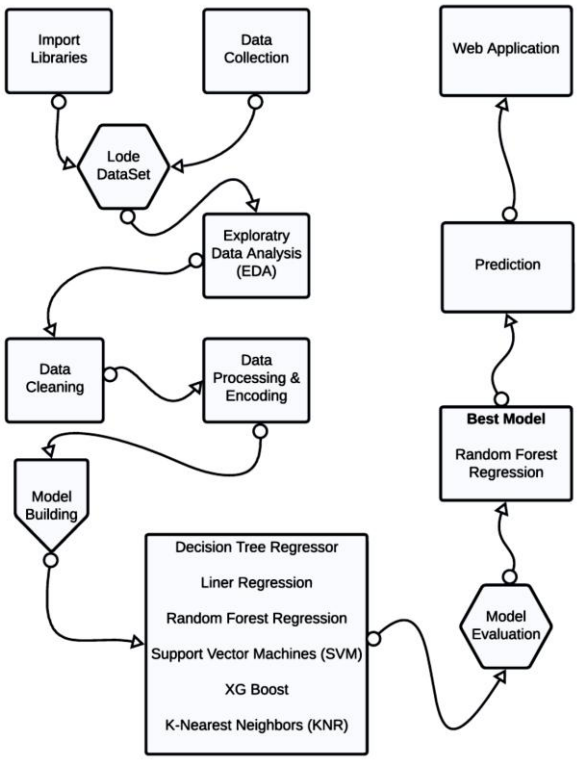


Fig. 1: Block Diagram of Implementation

Overall, the proposed methodology provides a systematic framework for developing the IPL Score Predictor, leveraging advanced data processing techniques, predictive modelling algorithms, and user-centric design principles to deliver accurate and intuitive score predictions for IPL matches.

IV. RESULTS

The results of the IPL Score Predictor project showcase the performance of various regression models in predicting the target variable, i.e., the total score in IPL matches. The evaluation metrics provide insights into the effectiveness and generalization capability of the predictive models.

A. Model Performance Metrics

- **Train and Test Scores:** The train and test scores indicate how well the models fit the data and generalize to unseen data. A high train score suggests that the model fits the training data well, while a high-test score indicates that the model performs well on unseen data.
- **Mean Absolute Error (MAE):** MAE measures the average absolute difference between the predicted values and the actual values. Lower MAE values indicate better model performance.
- **Mean Squared Error (MSE):** MSE measures the average squared difference between the predicted values and the actual values. It penalizes larger errors more heavily than MAE.

- **Root Mean Squared Error (RMSE):** RMSE is the square root of the MSE and provides an interpretable measure of the average error.

B. Model Comparison

The table below summarizes the performance of different regression models evaluated in the IPL Score Predictor project:

Table I: MODEL COMPARISION

Model	Train Score	Test Score	MAE	MSE	RMSE
Random Forest Regression	99.03%	93.87%	4.46	55.51	7.45
Decision Tree Regressor	99.98%	85.59%	4.10	130.65	11.43
XGB Regression	89.11%	85.63%	8.26	130.21	11.41
K Neighbors Regressor	86.66%	77.43%	9.90	204.60	14.30
Linear Regression	65.86%	66.09%	13.09	307.44	17.53
Support Vector Regression	57.47%	57.18%	14.78	388.17	19.70

From the results, it is evident that the Random Forest Regression model outperforms other algorithms, achieving the highest test score of 93.87%. The model demonstrates robust performance and generalization capability, as indicated by its low MAE, MSE, and RMSE values. Conversely, Linear Regression and Support Vector Regression exhibit lower performance metrics, suggesting limitations in capturing the underlying patterns of the data.

C. Interpretation of Results

The results highlight the trade-offs between model complexity, overfitting, and generalization in regression tasks. While some models achieve high train scores, they may suffer from overfitting and perform poorly on unseen data. Conversely, models with a balance between train and test scores demonstrate better generalization capability and are more reliable for score prediction in IPL matches.

Overall, the results underscore the importance of selecting appropriate regression models and evaluating their performance using relevant metrics to develop accurate and reliable prediction systems for cricket scores.

V. FUTURE WORK

In the pursuit of enhancing the IPL Score Predictor project, several avenues for future work can be explored to further improve prediction accuracy and expand the scope of the application:

A. Feature Engineering Enhancements

- **Advanced Feature Engineering:** Delve deeper into feature engineering techniques to extract more meaningful information from the dataset. Explore the creation of new features and the transformation of existing ones to better capture the nuances of cricket match dynamics.
- **Incorporating Domain Knowledge:** Integrate domain knowledge and expert insights into the feature engineering process to enhance the predictive power of the models. Domain-specific features related to player performance, pitch conditions, and match context can provide valuable inputs for improved predictions.

B. Model Optimization and Tuning

- **Hyperparameter Tuning:** Fine-tune the hyperparameters of regression models to optimize their performance further. Employ techniques such as grid search, random search, or Bayesian optimization to identify the best combination of hyperparameters for each model.
- **Ensemble Methods:** Experiment with ensemble learning techniques such as stacking, blending, or boosting to combine the strengths of multiple regression models. Ensemble methods often lead to improved predictive performance by reducing bias and variance.

C. Advanced Regression Models

- **Time-Series Analysis:** Explore time-series analysis techniques if the dataset contains temporal data. Time-series models like ARIMA, SARIMA, or Prophet may provide better insights and predictions, especially if the target variable exhibits temporal patterns.
- **Deep Learning Models:** Investigate the applicability of deep learning models such as recurrent neural networks (RNNs) or long short-term memory networks (LSTMs) for IPL score prediction. Deep learning architectures have shown promise in capturing complex patterns in sequential data and may offer improved prediction accuracy.

D. Evaluation and Validation

- **Cross-Validation Techniques:** Implement advanced cross-validation techniques such as k-fold cross-validation or time-series cross-validation to obtain more reliable estimates of model performance. Ensure robust evaluation metrics that account for model stability and generalization.
- **External Validation:** Validate the predictive models using external datasets or real-world IPL matches to assess their performance in diverse scenarios. External validation provides valuable insights into model robustness and generalizability across different contexts.

E. Scalability and Deployment

- **Scalability Considerations:** Address scalability challenges to accommodate larger datasets and increasing user demand. Explore distributed computing frameworks and cloud-based solutions to handle computation-intensive tasks efficiently.
- **Real-Time Prediction:** Develop mechanisms for real-time score prediction during live IPL matches, enabling users to receive instant updates and insights. Implement streaming data processing techniques to handle real-time data streams and incorporate live match data into prediction models.

By exploring these avenues for future work, the IPL Score Predictor project can evolve into a more sophisticated and reliable tool for predicting cricket scores in IPL matches. Continual refinement and innovation will contribute to the advancement of sports analytics and enhance the viewing experience for cricket enthusiasts and analysts alike.

VI. CONCLUSION

In conclusion, the IPL Score Predictor project significantly leveraged machine learning techniques to predict scores in Indian Premier League (IPL) cricket matches. The project aims to provide valuable insights into match dynamics and enable informed decision-making for cricket enthusiasts, analysts, and stakeholders by harnessing historical match data and employing advanced regression models.

Throughout the project lifecycle, rigorous data preprocessing, feature engineering, and model training procedures were undertaken to ensure the robustness and reliability of the predictive models. The findings highlight the Random Forest regression model as the top performer, demonstrating its ability to capture complex relationships among various match factors and deliver accurate predictions of IPL scores.

Despite challenges encountered during data collection and model development, such as data inconsistency and model complexity, the project has successfully addressed the primary objective of predicting IPL scores with reasonable accuracy. Moving forward, future work will focus on enhancing feature engineering techniques, optimizing model performance, and exploring advanced regression models to further improve prediction accuracy and scalability.

Overall, the IPL Score Predictor project represents a significant step forward in the domain of sports analytics, offering valuable insights and predictive capabilities that contribute to the ongoing evolution of cricket analysis and fan engagement. As the project continues to evolve, it holds promise for revolutionizing how cricket matches are analyzed, understood, and enjoyed by enthusiasts worldwide.

❑ ❑ ❑ ❑ ❑ ❑

REFERENCES

[1] Haseeb Ahmad, Ali Daud, Licheng Wang, 2017, "Prediction of Rising Stars in the Game of Cricket", DOI: 10.1109/ACCESS.2017.2682162

[2] Saranya G, Aravind Swaminathan, Joel Benjamin J, 2023, "IPL Data Analysis and Visualization for Team Selection and Profit Strategy" DOI: 10.1109/ICCMC56507.2023.10083736

[3] Harsh Mittal, Deepak Rikhari, Jitendra Kumar, Ashutosh Kumar Singh, 2021, "A Study on Machine Learning Approaches for Player Performance and Match Results Prediction"

[4] Rajesh Goel, Jerryl Davis, Amit Bhatia, 2021, "Dynamic cricket match outcome prediction" DOI: 10.3233/JSA-200510

[5] Rana Muhammad Kashif, 2020, "A Comparative Analysis of Data Mining Techniques To Predict The Projected Score And Winner In Twenty-20 Cricket", DOI: 10.13140/RG.2.2.10240.03844

AET 016

LevelUp- A Smart Way Towards Safety

Khooshi Tembhurne
Department of Eletronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
khooshi.tembhurne@vit.edu.in

Aditya Jogdand
Department of Eletronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
aditya.jogdand@vit.edu.in

Shivam Choughule
Department of Eletronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
shivam.choughule@vit.edu.in

Madhur Sarda
Department of Electronics Engineering
Vidyalankar Institute of Tehnology
Mumbai, India
madhur.sarda@vit.edu.in

Dr. Akshatha Bhat
Department of Electronics Engineering
Vidyalankar Institute of Technology
Mumbai, India
akshatha.bhat@vit.edu.in

***Abstract*— This project introduces an innovative approach to road safety via the integration of intelligent transportation systems (ITS) into roadways. Using data-driven solutions and the most recent technological developments, our study aims to provide novel approaches to reduce traffic accidents and improve traffic management in general. Our proposed prototype features functionalities capable of seizing challenges related to road safety such as speeding, pedestrian accidents, and driver non-compliance with traffic regulations. Dynamic speed bumps, an integrated challan system, light activated street lighting, and pedestrian-aware crosswalks are important parts of our proposed system. We expect significant improvements in accident prevention, traffic flow optimization, and overall transportation efficiency with the help of this project. Our findings underscore the critical role of smart infrastructure in developing a safer mobility experience and highlight the importance of continued research and innovation in this field.**

***Keywords*—Adjustable automated speedbump, integrated challan system, over speeding, pedestrian safety, motion activated streetlights, vehicle detection, automation, smart road.**

I. INTRODUCTION

Considering the increasing vehicular population on roads and highways, prioritizing road safety is a vital concern. This necessitates the evaluation of innovative methods to increase traffic law regulations and compliance amongst motorists. Speed limits, speedbumps, and rumble strips, however frequent, do little to ensure cooperation on the driver's end. Moreover, the efforts of safety regulators and law enforcement personnel, who regularly jeopardize their own safety for the security of all utilizing the public space, underscores the risks associated to manual oversight. Recognizing the susceptibility of human operators to errors, which may result in devastating incidents, there arises a vital requirement to embrace automated systems as explained within this paper.

The "LevelUp" system is an "all you need" solution for all road safety related matters. However small, the below stated project and all specifications are suitable for a prototype only. Specifications, components and all the necessary data that we reckon would be best for the implementation of this project in the practical world are mentioned further in the paper.

II. LITERATURE REVIEW

After thorough review and analysis of existing literature, notable scholarly contributions have been identified as crucial to this section's development. These references are thoroughly documented in the bibliography provided further into the paper, serving as the foundation for the research conducted in this paper.

The "LevelUp" project is structured into three distinct sections, each with notable contributions to the study.

A. Development of an Automated Height-Adjustable Speed Bump

The current research on this topic, although progressing, is minimal. An intelligent speed bump system designed to operate adaptively based on vehicle identification or driving speed is the goal for this project. Vehicle speed can be measured using pressure sensors, IR sensors or even using a completely software method. After that, image-based license plate recognition identifies approaching vehicles. The speed bump is activated if the vehicle's speed crosses the predefined speed limit, or if a pedestrian is walking on the zebra crossing. The system aims to bring road safety and security compliance amongst drivers, while also digitising the process to prevent human error on roads.

B. Integration of an Iot-Based Challan System using Speed & Number Plate Detection

The proposed system integrates with street cameras to detect and identify vehicles breaking traffic laws. It employs object detection, license plate detection, character segmentation, and recognition to extract license plate numbers from live feeds, recorded footage, or static images. Identified vehicles are matched with a local database to retrieve vehicle information, and electronic challans are generated and sent to offenders via email on the same day. This process streamlines and hastens the current tedious manual process of issuing traffic violations.

C. Design of Light Sensing Streetlight System

This section focuses on creating an intelligent street light control system to reduce power consumption. It utilizes LED technology and incorporates Light Dependent Resistors (LDRs) for it to function.

The system switches streetlights depending on the time of the day. Employment of an LDR makes it possible for the streetlight to switch ON when there is not sufficient light. In our prototype, we have made use of LEDs (Light Emitting Diode) to simulate the effect of streetlights in the project.

III. METHODOLOGY

To understand the method perfectly with the least bit of confusion, we will divide the paper into sections containing: Speed Detection, Pedestrian Detection, Automated Streetlight and Number Plate Detection.

D. Speed Detection

Speed detection is used for a major part of our project, i.e., automated speedbump. We use speed detection as one of the criteria to activate the height of the speedbump. The core idea is the compel the vehicle driver to slow down and follow safety regulations. We experimented with both software and hardware ideas for this part. Both have turned out to be quite promising. We have listed both the methods below.

The first methodology of speed detection, for better understanding, is segregated into six segments.

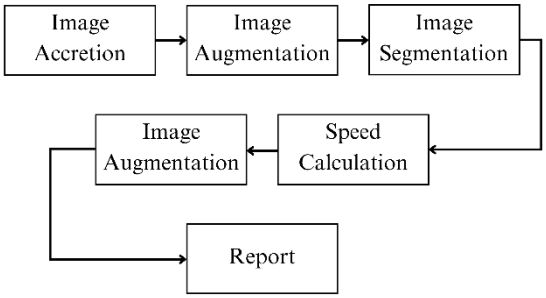


Figure 1

a) Image Accretion

The image accretion section takes video input and buffers the image frame into the memory. We incorporated a 48-megapixel phone camera for image capturing. We decided to use Microsoft DirectShow. This technology eases video streaming input without us having to worry about the format or source of the media.

The Filter Graph Manager within DirectShow manages this process, working at the device driver level to stream multimedia data. The filter graph consists of three main types of filters: source filters, decoder filters, and render filters. These filters work together to receive, process, and output media in a consistent format. In our system, the Image Accretion component calls the filter graph to grab single frames from the video stream and buffer them in memory storage.

b) Image Augmentation

Image Augmentation aims to process and enhance the image for better and efficient computing. Because of our limited computing power, we decided to use Grey Scaling and Image Scaling.

Grey Scaling is the process of turning a RGB colour image into a greyscale image. It converts an approximation of more 36 million colours in to 24 levels of grey. This reduces processing power for future processes. We settled for the Luminosity Method for grey scaling.

The Luminosity Method calculates the grey scale value of a RGB (red, blue, green) pixel weighted by their perceived brightness. The formula for converting a single (R, G, B) pixel to greyscale is:

$$Greyscale = 0.21 \times R + 0.72 \times G + 0.07 \times B \qquad (1)$$

We also implemented image scaling for computational requirement. Image Scaling is the process of resizing (downsizing or upsizing) an image without altering its aspect ratio. For this project, we made the decision to downsize the image. We employed the Bilinear Interpolation Algorithm for reducing the pixel size.

The Bilinear Interpolation Algorithm is a simple method used to downsize an image. It averages the colour value of neighbouring pixels to create a new pixel colour value. The bilinear interpolation formula calculates the intensity value of a target pixel in the downsized image based on the intensities of its four nearest neighbours in the original image. The formula is:

$$I_{new} = (1-f)(1-g)I_{tl} + f(1-g)I_{tr} + (1-f)gI_{bl} + fgI_{br} \qquad (2)$$

where,

- I_{new} = the new intensity calculated.
- I_{tl}, I_{tr}, I_{bl}, I_{br} = intensities of neighbouring pixels (top left, top right, bottom left, bottom right).
- f & g are the fractional parts of the x-coordinate and y-coordinate of the target pixel, respectively.

The formula calculates the weighted average of the intensities of the four neighbouring pixels based on their distances from the target pixel. The terms *(1-f)(1-g)*, *f(1-g)*, *(1-f)g*, *f* & *g* represent the weight assigned to each individual pixel, where:

- (1-f)(1-g) corresponds to the weight assigned to the intensity of the top-left neighbour.
- f(1-g) corresponds to the weight assigned to the intensity of the top-right neighbour.
- (1-f)g corresponds to the weight assigned to the intensity of the bottom-left neighbour.
- fg corresponds to the weight assigned to the intensity of the bottom-right neighbour.

By summing these weighted intensities, the formula computes the interpolated intensity value of the target pixel in the downsized image.

c) Image Segmentation

Image Segmentation is conducted to differentiate the foreground and background from each other. For this step, we utilised the image differentiation approach. The first step to segmenting the image is subtracting it. We take the first image of our video footage as the reference frame and subtract the rest of the images from it. This generates a new image after every subtraction.

After subtraction, the resulting image creates an image with both positive and negative values. In the next step, we exercise the Adaptive Thresholding Algorithm.

Withing adaptive thresholding algorithm, for each pixel, the neighbouring pixels is considered. A threshold value for that pixel is calculated based on the intensity distribution in that neighbourhood. Common methods are mean, median, or average of the neighbourhood pixel values. One a threshold value is chosen; it is compared to the pixels' intensity in the neighbourhood. If the intensity is greater than the threshold value, it is assigned the binary number 1; if smaller, 0.

1, in this section, means foreground and 0 equates to background. The result gives a binary image where pixels are classified as foreground and background.

To measure these intensities and calculate the threshold value, we used the Gaussian Weighted Average. There are some crucial factors that need to be identified before applying gaussian average. First, the size of the neighbourhood is identified and then the Gaussian Kernel.

Gaussian Kernel is a two-dimensional matrix that represents the weight assigned to each pixel within the neighbourhood. The size of the kernel determines the spread of the neighbourhood & the parameters of Gaussian function determine the shape spread of the weights.

Gaussian Weights Calculation: We first calculate the distance of the pixel from the middle of the local neighbourhood. Using these distances, we compute the corresponding weights using the Gaussian function formula, which is given as follows:

$$w(x,y) = \frac{1}{2\pi\sigma^2} \times e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3)$$

where,

- (x,y) are the coordinates of pixel relative to the one in the centre,
- w(x,y) is the weight assigned to the pixel at the position (x,y) in the neighbourhood,
- σ is the standard deviation of the Gaussian function,
- e is the base of natural logarithm

The next step is to normalize weight. We do this to ensure that the sum of all weights equals 1. For this, we divide each weight by the sum of all weights in the kernel.

Now, we compute the weighted average. We multiply each pixel's intensity & weight and add all the weighted intensities to obtain the Gaussian Weight Average. Mathematically, it is represented as:

$$\sum_{i=1}^{N}\sum_{j=1}^{N} Intensity(i,j) \times Weight(i,j) \quad (4)$$

where,

- *Intensity(i,j)* is the intensity value of the pixel position *(i,j),*
- *Weight(i,j)* is the weight of pixel position *(i,j),*
- *N* is the size of the neighbourhood.

d) *Image Analysis*

Further, we need to detect objects in the image. For this, we utilised the Connected Component Analysis (CCA) technique.

Once the foreground (1) & background (0) is identified, CCA scans through the image from top left corner, moving row-by-row. When it encounters a foreground pixel, it checks all neighbouring pixels to determine whether it is an object or not. If it is an object, the CCA assigns a unique label to each connected component onto the image. The labels start from 1 and keep on increasing with every cluster of pixels that seems like an object. This is the first pass of the Two Pass Algorithm.

If the neighbouring pixel is already labelled, the pixel is assigned the smallest value amongst them. Equivalence relationship is established during this pass.

Then the Two Pass Algorithm will make its second pass where it replaces the provisional label with a final label based on the equivalence relationship. Throughout both passes, the algorithm maintains an array to keep track of provisional & final labels. Then, we filter the connected components based on size, shape, or intensity to identify components likely to be the mark point.

This step is conducted using an algorithm called Centroid Calculation.

Each pixel within a connected component has coordinates (x,y). To calculate centroid, we sum up all the x-coordinates & y-coordinates. Next, we divide the sums of x-coordinates & y-coordinates by the total number of pixels in the component. This results in an average value of all x-coordinates & all y-coordinates within the component. Mathematically, centroid formula is given as:

$$C_x = \frac{\sum x_i}{N} \quad (5)$$

$$C_y = \frac{\sum y_i}{N} \quad (6)$$

where,

- x & y are the coordinates of each pixel,
- N is the total number of pixels.

This step gives us a mark point to track.

e) Speed Calculation

After centroids of the clustered components are marked, which in this case is a vehicle, the change of position between the consecutive frames/images is tracked & stored. After the position of the car is tracked, we use the Euclidean Distance formula to measure the distance travelled between two points. The Euclidean distance between two points in a two-dimensional space is given by:

$$Distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (7)$$

This formula calculates the straight-line distance between two centroids of the car.

To calculate the speed, we divide the distance (calculated above) by the time interval between frames. The time interval is found by taking the reciprocal of the frame rate. Frame rate is the measure of the how many frames are displayed or captured per second. Mathematically,

$$Time\ Interval = \frac{1}{Frame\ Rate} \quad (8)$$

Centroid tracking and Euclidean distance give us the speed of the vehicle for a pair of consecutive frames. This gives us individual speed measurements between consecutive frames. Further, we add all the distances travelled by the vehicle over the entire duration of the image sequence. That gives us the total distance travelled. For the time elapsed, product of the number of frames in the sequence & time interval between frames is taken.

Now, that we have the total distance & time elapsed, we detect the average speed (V_{avg}), which is mathematically given as:

$$V_{avg} = \frac{Total\ Distance\ Travelled}{Total\ Time\ Elapsed} \quad (9)$$

The result gives the speed of the vehicle.

The second method involved hardware equipment. An Arduino UNO R3 SMD, 2 IR sensors and a few jumper wires. To prevent uncalled confusion, we will refer to the IR sensors as IR1 and IR2.

The Arduino Uno R3 SMD (Surface Mounter Device) is a microcontroller board that we have used for this project. The necessary specifications are as follows:

TABLE I.

Microcontroller	ATmega328P
Operating Voltage	+5 V
Input Voltage	7V - 12 V
Output Voltage	6V – 20 V
Digital I/O Pins	14 (including 6 PWM output)
PWM Digital Output Pins	6
Analog Input Pins	6
DC Current per I/O pin	20 mA
DC Current for 3.3V Pin	50 mA
Flash Memory	32 KB (ATmega328P) of which 0.5 KB is used by bootloader
SRAM	2 KB (ATmega328P)
EEPROM	1 KB (ATmega328P)
Clock Speed	16 MHz
LED BUILTIN	13
Dimensions	68.6 × 58.4 mm

The pin diagram of Arduino Uno R3 SMD (Surface Mount Device) is:

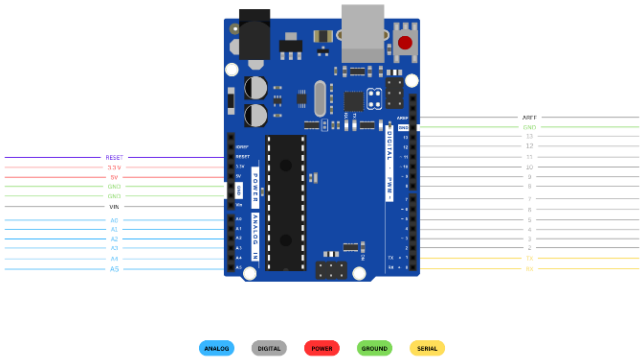


Figure 2

TABLE II.

Power	V_{in}	Input Voltage when using external source for Arduino.
	3.3 V	Generated by an onboard voltage regulator. Maximum current drawn is 50 mA.
	5 V	Regulated power supply used to power the microcontroller& other components on the board.
	GND	Ground pins.
Reset	Reset	Resets the microcontroller.
Analog Pins	A_0-A_5	Used to provide analog input from 0 to 5V.
I/O Pins	Digital pins: 0-13	Can be used as input or output pins.
Serial	0 (Rx), 1 (Tx)	Receives & transmits TTL serial data.
External Interrupts	2, 3	Triggers an interrupt.
PWM	3, 5, 6, 9, 11	Provides 8-bit PWM output.
SPI	10(SS), 11(MOSI), 12(MISO), 13(SCK)	Used for SPI communication.
Inbuilt LED	13	To turn on the inbuilt LED.
TWI	A4(SDA), A5(SCA)	Used for TWI communication.
AREF	AREF	To provide reference voltage for input voltage.

We incorporated IR sensors, i.e., infrared sensors, for speed detection. IR sensor has a transmitter that sends infrared light, which is reflected if it hits an object. The IR sensor looks for this reflected light, via the receiver mounted on the sensor, to detect the presence of an object in front of the sensor.

The specifications of the IR sensors (IR1 & IR2) that we have utilised for this project are as follows:

TABLE III.

Operating Voltage	3.3V to 5V D.C.
Detection range	2 cm – 30 cm (Adjustable using the potentiometer)
Detection Angle	<35°
Transmission Wavelength	940 nm – 950 nm
Reception Frequency	38 kHz
Average Current Consumption	0.06 mA
Dimensions	48 × 14 × 8 mm

The logic behind using the IR sensors is to place the sensors, IR1 & IR2, at an appropriate distance from each other to calculate the speed. IR1 & IR2 are placed at 10 inches in the prototype, however, the implementation of the final project in the real world would require the IR sensors to be placed at least 5 meters apart from each other. The second IR sensor, i.e., IR2 must be placed at a minimum of 10 meters from the speed bump to offer an adequate amount of time needed to slow down the vehicle.

To understand the connections of Arduino and the IR sensors, we need to understand the pin description of an IR sensor. These descriptions are below.

TABLE IV.

Vcc	Supply Input
GND	Ground Input
OUT	Goes LOW when an obstacle is sensed
Power LED	Illuminates when power is sensed
Obstacle LED	Illuminates when an obstacle is sensed
IR emitter	Infrared Emitter LED
IR Receiver	Receives signal that was transmitted

The technicalities of the connections are simple. We interfaced the IR sensors (IR1 & IR2) with the Arduino Uno by connecting both the sensors in the same fashion. The 3 pins on both the IR sensors; Vcc, GND and OUT are connected to the Arduino Uno. The Vcc pin of the IR1 sensor is connected to the 5V power pin, the GND pin is connected to the ground pin of the Arduino and the OUT pin to the analog pin A_1. IR2 is interfaced using the same connections. IR2 sensor's Vcc pin is connected to the 5V power pin, the GND pin to the ground pin of the Arduino and the OUT pin to the analog pin A_0.

The circuit diagram of all the connections is given as follows.

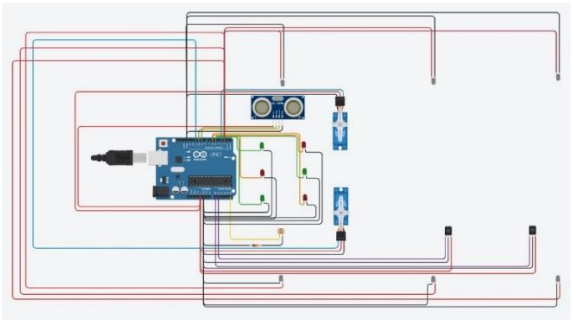


Figure 3

The 5V power pin of the Arduino Uno supplies sufficient power for the IR sensors to function simultaneously, while the analog inputs A_0 & A_1 work as inputs to the Arduino taken from the IR sensors.

The IR sensors (IR1 & IR2) are kept at a distance from each other to calculate speed. When a vehicle crosses IR1, the IR1 transmitting the infrared light is blocked and bounced back at the IR1 sensor, which detects the presence of an object, which in our case would be the vehicle. The time at which IR1 detects the vehicle is noted. Since the distance of the IR sensors from each other is known, we used the distance formula to calculate the speed, which is:

$$Speed = \frac{Distance}{Time} \quad (10)$$

This formula gives us the speed of the vehicle. In order to demonstrate a height adjustable speedbump, we need to add a servo motor below the speedbump in the prototype. This servo motor acts as a lever for the speedbump to rise and drop. For the purpose of this demonstration, we employed a SG-90 Servo Motor, the specifications of which are given below.

TABLE V.

Operating Voltage	+5 V
Operating Speed	0.1s/60˚
Torque	2.5 kg/cm
Rotation	0˚- 180˚
Gear Type	Plastic
Weight of the motor	9 gm
Dimensions	32 × 32 × 12 m

The wire configuration of the SG-90 servo motor is as follows.



Figure 4

TABLE V1.

Wire No.	Wire Colour	Function
1	Brown	Ground wire connected to the ground of the system
2	Red	Powers the motor with +5V supply
3	Orange	PWM signal is given through this wire to drive the motor.

The ground (brown wire) of the servo motor SG-90 is connected to the ground pin of Arduino Uno to establish an overall ground for the system resulting in reliability of the system connections. The red coloured wire is connected to the +5V power supply pin of the Arduino and the orange wire is connected to the digital PWM pin number 3. Refer to the circuit diagram, *Figure 3*, for the detailed and visualised image.

The output of connecting the servo motor to the Arduino Uno in this manner is candid. The connections of the IR sensors (IR1 & IR2) provide Arduino Uno with an input of the vehicle passing the sensors which is then computed using the distance formula mentioned in *equation (10)*. The speed detected is then run through the system for acceptable or deniable limits. If the speed limit is found to be under deniable circumstances, i.e., the vehicle crosses the speed limit set for that roadway, the servo motor turns from 0˚ to 60˚. This action causes the servo motor to lift the speed breaker.

However, we do realise that the real-world application for this system would require something sturdier than a servo motor. Hene, we propose that the speed breaker be made on top of a scissor mechanism that is often used in civil engineering to lift heavy object, which in our case would be vehicles passing by and from on top of the speed breaker. The scissor mechanism, under the speed breaker, would look like this:

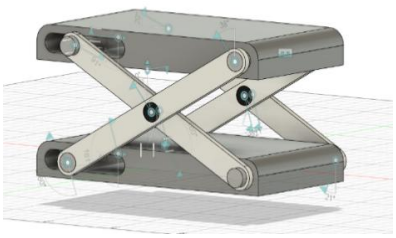


Figure 5

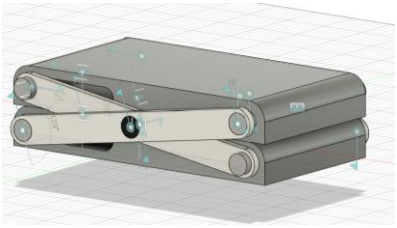


Figure 6

The *Fig. 5* shows the mechanism in a raised position & *Fig. 6* shows the same mechanism in the relaxed, i.e., flat position. The raised mechanism is only triggered under two conditions:

1. A vehicle exceeding the required speed limit.
2. A pedestrian using the crosswalk.

This is the mechanism that we proposed be incorporated for an adjustable heighted speedbump.

For this project, we have proposed, experimented with, and incorporated two methodologies of speed detection. Both works satisfactorily and have various advantages as well as disadvantages.

The first methodology requires more time and effort into processing and coding it, however, the implementation of the second methodology is simple to comprehend. The first methodology is more accurate, yielding more precise result, while the second methodology has two major drawbacks. First drawback being that the IR sensors will detect any object as a vehicle, which would activate the speedbump even when there is no speeding of vehicles. The second major drawback is that the system can only detect and work for one vehicle at a time and cannot be implemented for multiple vehicles. However, for the purpose of demonstration, we made the decision to stick to the second method.

E. Pedestrian Detection

The first circumstance under which the speedbump would rise is unlawful exceedance of speed by a driver. The second condition is a pedestrian using the crosswalk which is placed ahead of the speedbump. In the implementation of this project on actual roadways, the distance between the speedbump and the crosswalk will be determined by the speed limit set on that roadway.

According to the guidelines issued by the Ministry of Railways by the government of India, the minimum distance between a crosswalk and the speed breaker before it depends on speed of the vehicle, reaction time of the driver and the coefficient of friction. The speed of the vehicle will be assumed to be the maximum speed limit mentioned for the roadway and the coefficient of friction depends on the material of the road and the tyres. Mathematically, the distance will be calculated using this formula:

$$Stoppind\ Distance = vt + \frac{v^2}{2gf} \qquad (11)$$

where,

- v is the speed of the vehicle,
- t is the reaction time (2.5 sec, as per the IRC),
- g is the acceleration due to gravity (9.8 m/s^2),
- f is the coefficient of friction (0.40 for 50 km/hr)

For the project, we decided to use an ultrasonic sensor HC-SR04. The necessary specifications of the HC-SR04 ultrasonic sensor are given below:

TABLE VII.

Operating Voltage	+5 V D.C.
Operating Current	<15 mA
Operating Frequency	40 KHz
Theoretical measuring distance	2 cm to 450 cm
Practical measuring distance	2 cm to 80 cm
Accuracy	3 mm
Measuring Angle	<15˚
Dimensions	45 × 20 × 15 mm

Ultrasonic sensors emit short, high frequency pulses of audio waves that are above the hearing abilities of a human being. It works by propagating these pulses at the velocity of sound. After the emission of these pulses, the ultrasonic sensor's receiver waits for the audio frequency. If the audio frequency is not sensed by the received, then there is no object in the path of the sensor. Conversely, if the receiver senses the audio frequency pulse, then it is said that the ultrasonic sensor has detected something in its path. For this project, the object will be the pedestrian utilizing the crosswalk.

To understand the connections of the ultrasonic sensor HC-SR04 to the Arduino Uno R3 SMD, the pin diagram and configuration of the ultrasonic sensor is given below.

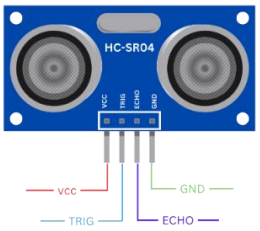


Figure 7

TABLE VIII.

Vcc	Powers the sensor, +5V
Trigger	Input pin. Must be kept hight for 10 µs to initialize measurement by sending ultrasonic waves.
Echo	Output pin. Will be kept high for the time it takes the ultrasonic wave to return.
Ground	Connected to the ground of the system.

Refer to Figure 2 & Table 2 for the pin configuration of Arduino Uno R3 SMD.

The Vcc of the ultrasonic sensor HC-SR04 is connected to the 5V power pin of the Arduino Uno R3 SMD board. The trigger input pin of the sensor is interfaced with the PWM digital pin

number 9 while the echo output pin is connected to the digital pin number 10. The ground pin of the ultrasonic sensor is connected to the ground pin of the Arduino Uno.

The 5V power pin supplies abundant power for the ultrasonic sensor to function without any difficulties. The trigger pin that is connected to the digital pin 9 of the Arduino board, is used as a trigger signal. The digital pin 9 will send trigger signals to the trigger pin of the ultrasonic sensor indicating for it to initialize sending the short bursts of ultrasonic pulses. The digital pin 10 will be used to receive echo pulse from the echo pin of the ultrasonic sensor. By connecting Arduino Uno's and ultrasonic sensor's ground pins together, they share a common ground reference. This allows for reliable communication and accurate voltage measurements between the two devices.

A SG-90 servo motor is connected to the Arduino Uno R3 SMD. Refer the *Table 2* & *6* for the pin configurations of Arduino Uno and servo motor, respectively.

Refer the *Figure 3* for the circuit diagram. The ground (brown wire) of the servo motor SG-90 is connected to the ground pin of Arduino Uno to establish an overall ground for the system. The red wire is connected to the +5V power supply pin of the Arduino and the orange wire is connected to the digital PWM pin number 11.

The connections result in the desired output. Whenever a pedestrian is using the crosswalk, the servo motor will turn to 60˚ (can be adjusted to go higher or lower) causing the speedbump to raise higher.

For the safety of the driver in the vehicle and the pedestrian, we have also connected red and green LED lights on the edge of the crosswalk. The red LEDs will turn ON in the presence of a pedestrian, indicating for the driver to slow down. Otherwise, the green LEDs constantly stay ON in order to indicate a clear road ahead.

We interfaced 6 LED lights, 3 red & 3 green, for the safety of the pedestrian and the driver of the vehicle. The red LEDs are meant to simulate the "STOP" sign while the green LEDs are meant to simulate the "GO" sign. The colours are chosen in such a way that it makes it easier and faster to comprehend, thereby, decreasing reaction time needed by the driver.

All 6 LEDs are connected in such a manner that all the negative ends, i.e., cathodes are connected to the ground pin of the Arduino. The 3 green LEDs' positive ends, i.e., the anode is connected to the digital pin number 5. The 3 red LEDs' positive ends, i.e., the anode is connected to the digital pin number 6 of the Arduino Uno.

Refer *FigureS 10* & *11* for the prototype of the project.

F. Automated Streetlight

The proposed automated streetlight system in this paper is an LDR based one. The goal is to save energy and become aware of the consumption of power on the daily basis.

The core idea is elementary. The LDR (Light Dependant Resistor) is a light sensitive sensor. LDR has variable resistance based on the light around it. The resistance of the LDR, when placed in a dark room, will have a resistance of a few megaohms and as we gradually introduce light, the resistance drops from a megaohms to a few ohms.

One leg of the LDR is connected to one end of a resistor. The other end of the resistor is connected to the ground pin of the Arduino Uno, and the other leg of the LDR is the connected to the analog input A_5.

A total of 6 LEDs (Light Emitting Diode) are connected to the Arduino. They are simulated in the colour grey in the circuit diagram *(Fig. 2)*. All LEDs are interface with the Arduino in the same fashion. All negative ends, i.e., cathodes of the LEDs, are connected to the ground pin of the Arduino Uno. The anodes of all LEDs, which are simulating the streetlights, are connected to digital pin number 2.

The Arduino Uno is coded using Arduino IDE. We programmed for the "streetlights" to switch ON in the absence of light, which is typically when approaching sunset, to increase the visibility of the speedbump approaching.

G. Number Plate Detection

We integrated number plate detection with the intention of incorporating challan system. The vision behind building this system is so that it will recognise, save, and issue a challan to the over-speeding vehicles, thereby, automating the ticketing/challan procedure neglecting human error as well as the necessity for law enforcers to endanger their lives regularly on roads.

For the function of this project, we used a mobile camera of 48 megapixels with a stable internet connection. This prototype requires a mobile camera with stable internet connection and a laptop. We used the IP Webcam Application which uses the local network to transmit and receive video streams to the connected device. For the application to work satisfactorily on the camera phone, OS Android v. 4.1 or higher and a shared Wi-Fi connection between the phone and the laptop is required.

We used the IP Webcam version 1.17.15.868 (multiarch) with the recent update made on Oct 25th, 2023. The application is installed and configured to function suitably. The phone camera records and live streams the footage to the laptop for further processing and number plate detection.

The configuration of the IP Webcam is as follows:

TABLE IX.

Video Resolution	2400 × 1080 pixels
Frame Rate	30 fps (frames per sec)
Camera Orientation	Landscape
Audio Settings	Disabled
Motion Detection	Disabled

The IP Webcam only records and streams video input from the camera and transmits it in real time to the laptop for further processing. Until now, we have only taken the video of the vehicle and transferred it to the device it is connected to. Number plate recognition, however, is the process that comes along after this. For this project, due to computing constraints and crunching time limitations, we opted for a simpler as well as an effective way of number plate recognition named YOLO (You Only Look Once) Method. However, in the absence of such constraints and possession of advanced physical/ hardware components, we would have urged for the ANPR (Automatic Number Plate Recognition).

We have partitioned the methodology of for number plate detection using YOLO into 6 sections given below.

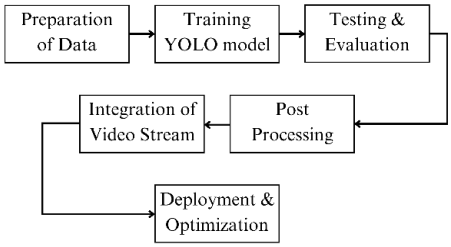


Figure 8

a) Preparation of Data

The first step, i.e., preparation of data requires for us to collect relevant data to train the YOLO model. For this reason, a variety of images were exhausted. These images were finalised ensuring that the dataset covered wide range of scenarios like distance of the vehicle, vehicle orientation, background, lighting conditions, visibility & varying plate numbers. Some samples of the dataset are given below.



Figure 9

This dataset was created using various sources such as public datasets, web scraping or capturing images using a phone camera. All images in the dataset are annotated using an annotation tool named VoTT (Visual Object Tagging Tool). VoTT ensures that the object (number plate) in the image is tightly marked using a bound box. The bounding box must tightly contain the number plate with negligible surrounding area in the marked bounding boxes. A unique class label is then assigned to each box, for e.g., “number plate”, “license plate”.

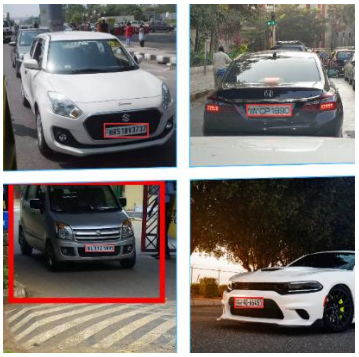


Figure 10

These annotated datasets are categorized into three distinctions for the most efficient usage possible. The distinctions are training set, validation set and test set.

The training set is used to train the model. this set contains maximum of the datasets which have images in diverse conditions for the model to learn from. Abundant datasets must be provided for training the model.

The validation set must be kept separately from the training set. This is used to assess and thereby apply corrective measures to get pertinent result. This step is necessary to monitor the training progress and prevent overfitting.

The test set is, again, kept entirely separate from the training set and validation. These are unseen datasets for the system. The result of these datasets determines the apt working of the model.

Now, we process the dataset for appropriate training of the model. YOLO accepts certain image sizes & formats as input to the model. Standardized formats are JPEG, PNG or BMP. We use OpenCV library in python to resize and convert the image. It is good practice to normalize pixel value for which we used NumPy.

We import the file and change it to a grayscale image. Then, to normalize, we divide the pixel values of the grayscale image by 255 in order to scale them in the range of [0,1]. These normalized images replace the unprocessed image and are now ready to be fed into the YOLO Model.

b) Training YOLO Model

Based on our requirement, we chose the YOLOv8 model that is pre-trained on large datasets such as COCO (Common Object in Context). We fine tune this pre-trained model to accommodate our specific dataset. The dataset is the predefined training set containing the annotated images for modelling and training.

c) Testing & Evaluation

In this section, we evaluate the model to identify the number plates. We present it with the test set, a dataset of images that aren’t used in training or validation set, to evaluate the performance of the model.

The metrics for evaluation are Mean Average Precision (mAP), precision, recall and F1 score.

- Mean Average Precision (mAP) is often used in object detection to find the average precision of the system on a number of images.
- Precision is the ratio of the correctly predicted positive/true cases to total predicted positive/true cases.
- Recall is the ratio of the correctly predicted positive/true cases to the total number of actual positive/true cases.
- F1 score provides a balanced measure of the model’s performance by calculating the harmonic mean of precision and recall.

Based on the evaluation result, we analysed the model’s performance and adjusted a few parameters like batch size and argumentative technique.

We kept on repeating the whole testing and evaluation procedure, reiterating areas of improvement and fine tuning the parameters until we were satisfied with the result.

d) Integration of Video Stream

The video stream is set up using the IP Webcam application on a phone camera of 48 megapixels that streams the video footage from the phone to the laptop. Using the OpenCV library in python, we capture frame from the video stream and then apply the YOLOv8 trained model.

For each frame/image, it is first pre-processed like resizing, normalizing of pixel values, and converting to standardized formats. This pre-processed image is then prepared to be given as an input to the YOLOv8 model. The model will detect elements in the image but most importantly, it will detect the number plate.

e) Post Processing

The post-processing section makes sure that there are no repeated results by keeping the most confident and accurate predictions. We use the NMS (Non-Maximum Suppression) Technique for this step. The usage of the YOLO algorithm generating the bounding box for an object, in our case a number plate, each bounding box is assigned with a confident score. The confident score signifies the likelihood of the desired object (number plate) being inside the bounding box. The NMS technique then sorts the bounding boxes in the decreasing order of the confident score and marks the one with the highest score as "selected". This "selected" box is then compared with all the boxes in the list of confident score. For each box, we calculate the Intersection over Union (IoU) between the selected box and the remaining ones. IoU measures the overlap of two objects. Mathematically written as:

$$IoU(b,s) = \frac{Area\ of\ Intersection(b,s)}{Area\ of\ Union(b,s)} \quad (12)$$

where,

- B is a set of bounding boxes each associated with confidence *conf(b)*,
- S is an empty set for storing selected bounding boxes,
- s is every selected bounding box in set S,
- b is each bounding box in set B

If,

$$IoU(b,s) \geq IoU\ threshold, \quad (13)$$

Consider b significantly overlapping with s and discard b.

The next step is to employ the OCR (Optical Character Recognition) Technique. OCR extracts alphanumeric values to from images. In our case, it will be numbers from the number plates of the vehicles.

For the OCR to be able to detect the alphanumeric figures in the image, the image must be pre-processed, toggling the pixel values and normalization of them, grey scaling and image scaling as well as the region of interest needs to be mapped out. This has already been taken care of for the number plate detection part of the paper. To be able to extract the text from the background, we used the CCA (Connected Component Analysis) method, mirroring the same way we implemented for the "Image Analysis" division under "Speed Detection" section of this paper.

After the segmentation of the foreground from the background, OCR runs a pattern recognition technique to identify the shape, structure, and feature of the characters by matching them to a predefined set of character classes like alphanumeric characters.

The final output is the OCR text which is saved in a JavaScript Object Notation (JSON) file.

f) Deployment & Optimization

The last step of number plate and license recognition is the deploy the system on the optimized number plate recognition system on the laptop for real-world usage, ensuring proper installation and configuration.

After multiple iterations, changing datasets, integrating video stream from the IP Webcam application, object detection with YOLOv8, processing using NMS and text detection using OCR, the deployment of the number plate recognition system is adequate.

IV. Results

The result of our proposed system yielded a "Smart Road" taking inspiration from a "Smart City". The project prototype is implemented appropriately for a smart road design.

The speed detection system, utilising both the methods, give adequate result. The methodology employing IR sensors, even with its drawbacks, works satisfactorily. However, the speed detection technique using Gaussian Weighted Algorithm and CCA (Connected Component Analysis) are preferred.

The pedestrian detection system used a straightforward approach with an ultrasonic sensor to sense the presence of a pedestrian.

The number plate detection mechanism employing YOLOv8, NMS & OR techniques result in apt output for the system.

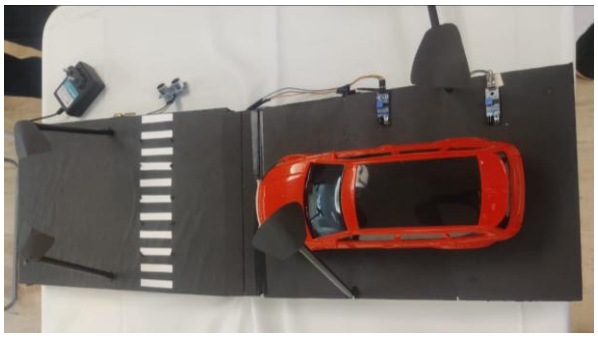


Figure 11



Figure 12

Although, it is fair to presume that this system is not ready for complete implementation onto roadways, however, it is a step in the right direction of road safety, automation, and advancement. The mentioned techniques yielded a working prototype of speed detection, pedestrian detection, smart streetlight & number plate recognition.

V. CONCLUSION

The objective of this project is to implement a smart roadway system that functions efficiently, saving energy, enforcing road safety laws, automating the challan procedure and reduce the endangerment to the lives of law enforcers.

The study conducted for this paper is based on several scholarly papers, using them as building blocks for the advancement in road safety.

VI. FUTURE SCOPE

The research conducted in this paper is possible because of the papers and research conducted before. In the same manner, we want this paper to act as a foundation for future road safety technology. Some potential research for future development could be piezoelectric technique. Since the speedbump is raising and dropping, a pressure sensor could be placed under the speed breaker to generate & store electricity for the streetlights that are light activated. This will decrease the consumption of electricity significantly, especially because of the light activate streetlights but also the power being generated by the pressure sensor.

Another scope of study is to add a challan system using the number plate recognition system. This technique of issuing challan will be completely automated reducing any possibility of corruption or bribery. The automated challan system must be triggered when a vehicle is over-speeding and issued automatically on the registered phone number and vehicle details that will be obtained by the number plate of the vehicle.

These are some potential research opportunities we see in this area of study.

REFERENCES

[1] H. -Y. Lin and C. -Y. Ho, "*Adaptive Speed Bump With Vehicle Identification for Intelligent Traffic Flow Control*," in *IEEE Access*, vol. 10, pp. 68009-68016, 2022, doi: 10.1109/ACCESS.2022.3186010.

[2] A. Hakan Lav a, Ertugrul Bilgin b, A. Hilmi Lav, "*A fundamental experimental approach for optimal design of speed bumps*", Accident Analysis & Prevention, Volume 116, Pages 53-68- July 2018.

[3] Shubham Kumar Chandravanshi, Hirva Bhagat, Manan Darji, Himani Trivedi, "Automated Generation of Challan on Violation of Traffic Rules using Machine Learning" International Journal of Science and Research (IJSR)-March 2021.

[4] Lae Yin Mon | Khin Khin Saw "Design and Construction of Speed Detection System for Vehicles" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-3, April 2019, pp.803-805.

[5] Aziera Abdullah, Siti Hajar Yusoff, Syasya Azra Zaini, Nur Shahida Midi, Sarah Yasmin Mohamad, "Energy efficient smart streetlight for smart city using sensors and controller" Bulletin of Electrical Engineering and Informatics- June 2019.

[6] Chirag Patel, Dipti Shah, Atul Patel, "Automatic Number Plate Recognition System (ANPR): A Survey", International Journal of Computer Applications- May 2013.

[7] PMC speed breaker guidelines for Pune https://www.pmc.gov.in/sites/default/files/speedbreaker-guidelines-28-11.pdf

[8] Draft Guideline on the provision of Speed Breakers for Control of Vehicular speed on minor Roads: https://rdso.indianrailways.gov.in/works/uploads/File/WKS-G-10.pdf

[9] https://techzeero.com/sensors-modules/ldr-sensor/

[10] https://components101.com/resistors/ldr-datasheet

[11] https://components101.com/motors/servo-motor-basics-pinout-datasheet

❑ ❑ ❑ ❑ ❑ ❑

AET 017

Computational Stellar Model

Pranav Shinde
Department of Electronics & Computer Science
Vidyalankar Institute of Technology
Mumbai, India. pranav.shinde@vit.edu.in

Neeraj Kamble
Department of Electronics & Computer Science
Vidyalankar Institute of Technology
Mumbai, India. neeraj.kamble@vit.edu.in

Durvesh Nikam
Department of Electronics & Computer Science
Vidyalankar Institute of Technology
Mumbai, India. durvesh.nikam@vit.edu.in

Jayesh Raut
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
Mumbai, India. jayesh.raut@vit.edu.in

Uma Jaishankar
Department of Electronics and Computer Science
Vidyalankar Institute of Technology
Mumbai, India. uma.jaishankar@vit.edu.in

***Abstract*—Computational stellar models are helpers of physics, mathematics, and computer science that serve as essential tools in our quest to decipher the mysteries of the universe. These sophisticated simulations allow us to study the internal dynamics of stars at a wide range of masses, temperatures, and evolutionary stages. From the fiery inferno of massive supernovae to the gentle glow of long-lived main sequence stars, these models provide a window into the diverse behaviors of astronomical objects. In addition, they facilitate the prediction of stellar lifetimes, the understanding of nuclear fusion processes, and the assessment of the impact of stellar activity on planet formation and evolution. These models are also an important foundation in the field of astrophysics, providing a powerful tool for unraveling the mysteries of the most abundant and brightest inhabitants of the universe: stars. Computational stellar models provide a comprehensive view of objects in the universe by examining the inner workings of stars, from star formation and fusion reactions to their ultimate fate. These simulations provide insight into the synthesis of elements, energy production, and the rise and fall of the life cycle of stars, contributing to a deeper understanding of the cosmos [2]. Through meticulously crafted algorithms and state-of-the-art visualization tools, the project endeavors to simulate the birth of stars within protostellar clouds, their gravitational collapse, and the intricate dance of binary systems. With a focus on realism and accuracy, the simulation platform will offer an immersive experience, allowing users to traverse the cosmos and witness the cosmic ballet of stellar evolution firsthand.**

***Keywords*—Computational stellar models, Stellar evolution, Nucleosynthesis, Astrophysics, Main-sequence stars, Universe, Celestial luminaries.**

I. INTRODUCTION

Computational stellar models serve as a vital tool in the field of astrophysics, facilitating the comprehensive exploration of celestial bodies that have captivated our collective imagination for centuries. These models represent a combination of different scientific disciplines that seamlessly combine principles from physics, mathematics, and computer science to simulate the complex behavior and complex evolutionary paths of stars. At the heart of these simulations is a profound effort to unravel the mysteries of stellar bodies and unravel the mysteries surrounding their formation, development, and ultimate fate. Computational stellar models take advantage of a deep understanding of the fundamental physical processes governing the universe, such as nuclear fusion, electromagnetic radiation, and gravity, and serve as valuable guides for observing the tapestries of the universe. [4]. These provide a window into the incredible journeys of stars across a wide range of mass ranges, from the smallest and faintest red dwarfs to the largest and brightest supergiant's. In complex algorithms and complex Mathematica formulas, these models capture the cosmic story of stellar evolution, elucidate the mechanisms that shape the universe, and provide deep glimpses into the majesty of the cosmos' most abundant and brightest inhabitants. The development of computational stellar models represents an ongoing effort that has taken place over several decades, advancing human understanding of the universe's celestial bodies to new heights and creating a deeper connection between science and the cosmos itself. As we journey through the complex world of computational stellar modeling, we explore the tools, principles, and discoveries that have shaped our modern understanding of the universe. [5]

II. PROBLEM STATEMENT

The accurate representation and prediction of stellar evolution and behavior remain critical challenges in astrophysics. Current computational stellar models face limitations in accurately simulating complex physical processes, such as convection, magnetic fields, and radiation transport, especially in massive and evolved stars. These limitations hinder our ability to precisely forecast the lifetimes, final fates, and nucleosynthetic contributions of stars. Moreover, as stellar modeling is integral to understanding the formation and evolution of galaxies and the universe's chemical enrichment, improving the fidelity of computational models is paramount. This research seeks to address these challenges by advancing the state of computational stellar models to better encompass the complexities of stellar interiors, thereby enhancing our comprehension of the cosmos.

IIi. PROPOSED SYSTEM

Computational stellar modelling methodology is a comprehensive, systematic, and iterative approach that serves as a foundation in our quest to explore the complexities of stellar evolution and behavior. It begins by formulating the physics-based equations that underlie key principles governing stars, such as the delicate balance between gravity and pressure gradients. The method then proceeds to numerical discretization to transform these equations into a grid-based framework suitable for computational analysis [12]. A key aspect of this process lies in the judicious treatment of boundary conditions, which must accurately represent the physical constraints on the surface and center of the star and ensure the reliability of the model [11]. Time integration techniques are introduced to guide the model

through different stages of stellar evolution, choosing the integration method appropriate to the specific astrophysical problem and available computational resources. As the model progresses, modules designed to capture different physical processes are absorbed, each based on established theory and experimental data. These modules cover the intricacies of nuclear reactions that generate energy and synthesize chemical elements in stars, while also addressing complex dynamics such as convection, radiative transfer, and the effects of magnetic fields. Iterative solvers are introduced to maintain numerical stability and accuracy, especially in regions where nonlinear interactions are common[13]. Model validation and testing represent an ongoing step in the methodology and include, when possible, careful comparison of analytical solutions with observed data. Sensitivity analysis is used to examine the responsiveness of a model to changes in key input parameters, thereby providing valuable insight into the robustness and inherent limitations of the model. Effective use of high-performance computing resources is essential because it allows us to perform the extensive simulations needed to fully capture the complexity of stellar evolution much further. Now, introducing 'star2' into this cosmic ballet adds a whole new dimension of exploration. As 'star2' emerges into the virtual firmament, we can observe its gravitational influence on neighboring stars, its orbital path as it navigates through this celestial tapestry, and how it dynamically shapes the evolving stellar cluster. This captivating simulation isn't confined to local computation; it thrives on the collaborative capabilities of Google Colab. Researchers, students, and enthusiasts alike can seamlessly access and run the simulation, fostering collaborative exploration and discovery. And when the simulation unfolds its cosmic dance in ways both expected and unexpected, users can easily capture these mesmerizing moments[14].

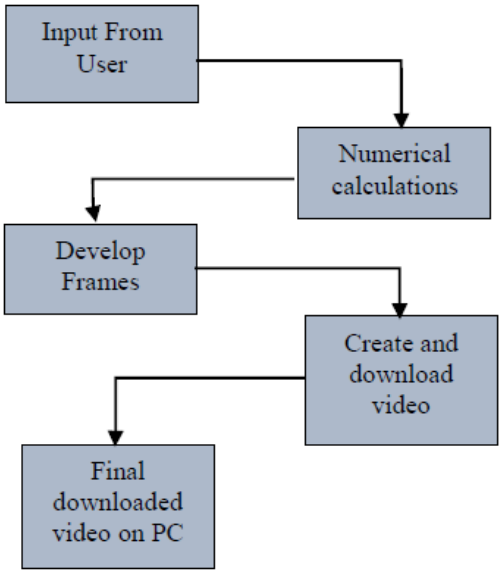


Fig. 3.1 Flow for Star Formation

A. Flow for Star Formations

First, taking the input from the user as to what be the star's composition. Factors like Number of particles, Simulation time, Star mass, Star radius, smoothing length, equation of state constant (k), polytropic Index, Dampning, lambda. These initial conditions are required to be manually put into the code to change its appearance and formation.

Second, we calculate the different values needed for the particles to move and come together like the Gaussian smoothing kernel, smoothing kernel gradient, density, pressure, and the color of the stars as we move away from its core. All other calculations are done in order to show the smooth transition of the formation of a star.

Thirdly, the OpenCV module of Python crates these time-by-time frames that are later put together to create a single video. But for now, individual frames are created which can take time as each frame needs to be created individually. So this step of loading these frames can take time. See Fig 3.1.

All of the frames are put together in a single video which is directly downloaded under the name “Single star simulation” onto the PC/Laptop. And finally, the downloaded video can be played in order to show the beautiful dance of particles that come together to form a star and stabilize.

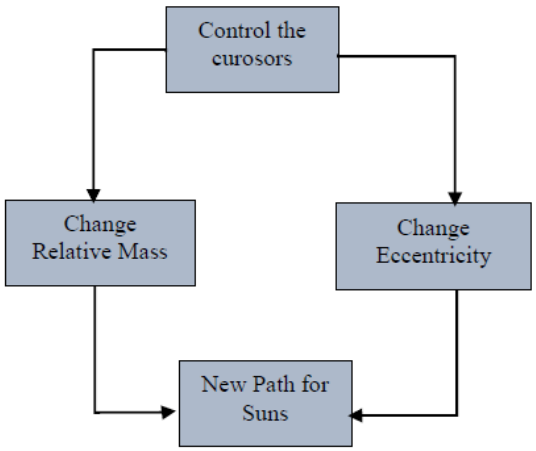


Fig. 6.2 Flow for Binary System

B. Flow for Binary Star system

The Binary systems shown on the webpage is an interactive simulation that shows the path of a two-star system revolving around a center of mass and depicts their path taken. These paths are affected by two factors and can be altered by altering these factors through the cursor provided on the page to easily slide and change them –

I. Changing relative mass – The relative mass between both can be changed from a scale of 0 to 1 with 1 being equal to each other and 0 being the least possible relative mass among them to keep them in a binary star system.

II. Eccentricity – their relative distance to each other is also adjusted from 0 to 1 with 1 being 0 being the closest they can be without collapsing into each other and 1 being the farthest they can be but still revolve around the center of mass.

By exploring these parameters, we can gain insight into how different configurations affect the behavior of binary star systems, from their orbital paths to their stability over time. It's a fascinating way to understand the intricate dance of celestial bodies in the cosmos.

C. Physics and Formulas

A Gaussian smoothing kernel is a mathematical function used for image processing and signal processing tasks, particularly for blurring or smoothing images. It's based on the Gaussian distribution, which is a bell-shaped curve that describes the probability distribution of a continuous random variable.[10]

```
def kernel(x, y, h):

 # Calculate |r|
 r = np.sqrt(x**2 + y**2)
 # Calculate the value of the smoothing function
 w = (1.0 / (h*np.sqrt(np.pi)))**3 * np.exp( -r**2 / h**2)
# Return
return w
```

The Gaussian smoothing kernel is characterized by two parameters: the standard deviation (σ) and the size of the kernel (usually denoted by a kernel size parameter, such as k). The standard deviation controls the spread of the Gaussian curve, while the kernel size determines the extent of the neighbourhood considered for the smoothing operation.
The Smoothing kernel gradient refers to the gradient or derivative of a smoothing kernel function. In image processing and computer graphics, the smoothing kernel is often applied for tasks like blurring or denoising images.

```
def gradKernel(x, y, h):
 #evaluated gradient

 # Calculate |r|
 r = np.sqrt(x**2 + y**2)
 # Calculate the scalar part of the gradient
 n = -2 * np.exp( -r**2 / h**2) / h**5 / (np.pi)**(3/2)
 # Get the vector parts of the gradient
 wx = n * x
 wy = n * y
 # Return the gradient vector
 return wx, wy
```

The gradient of the smoothing kernel is essential for operations such as edge detection and feature extraction. It describes how the weights assigned to neighboring pixels change with respect to changes in the spatial dimensions. Essentially, it provides information about the rate of change of the smoothing operation in different directions.
Image processing, "color from density" typically refers to a technique used to visualize data by mapping a density value to a color.

```
def magnitude(ri, rj):


 M = ri.shape[0]
 N = rj.shape[0]

 # get x, y of r_i
 r_i_x = ri[:,0].reshape((M,1))
 r_i_y = ri[:,1].reshape((M,1))

 # get x, y of r_j
 r_j_x = rj[:,0].reshape((N,1))
 r_j_y = rj[:,1].reshape((N,1))

 # get r_i - r_j
 dx = r_i_x - r_j_x.T
 dy = r_i_y - r_j_y.T

 return dx, dy
```

Finding Pressure apllied to each indiviaula particle

```
def pressure(rho, k, n):


 P = k * rho**(1+1/n)

 return P
```

The Pressure formula for the acceleration of particles depends on the specific physical system and the forces acting on the particles. In fluid dynamics, for example, pressure plays a significant role in determining the acceleration of fluid particles.
The acceleration of particles in a fluid under the influence of pressure gradients. In this case, the acceleration (a) of a fluid particle can be expressed using the Navier-Stokes equation:
$\partial q/\partial t + (q \cdot \nabla)q = -1/\rho\ \nabla P + \nu\ \nabla 2q$ and $divq = 0$

D. Libraries

The cv2 module contained in Python forms part of the OpenCV (Computer Vision in Source Open) library that is commonly used in the computer vision world for tasks like image & video processing, object detection, and machine learning among others. Here is how it is done in the code below:

- Importing
- Reading Images
- Writing Video
- Writing Frames

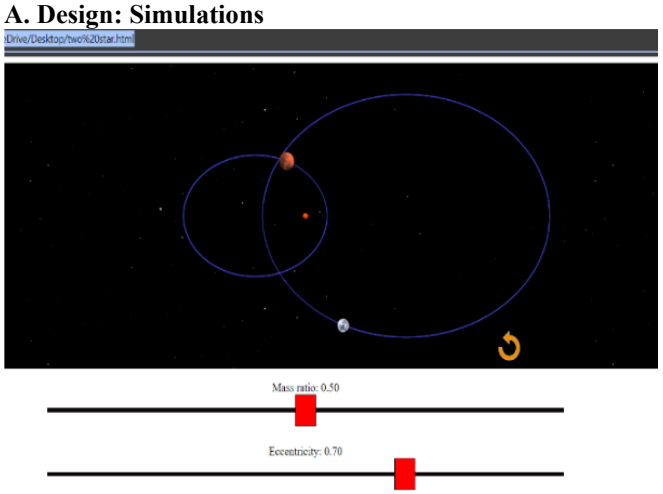
We use cv2 to to download and load the video and create its frame which show the formation of star.
Python's NumPy is a strong numeric computing library. In this library are functions that can be used on large arrays and matrices which have multiple dimensions. We use NumPy mainly In gravity, fluid dynamics, meteorology and simulations on these areas of physics (such as simulations of air flow over a plane wing).

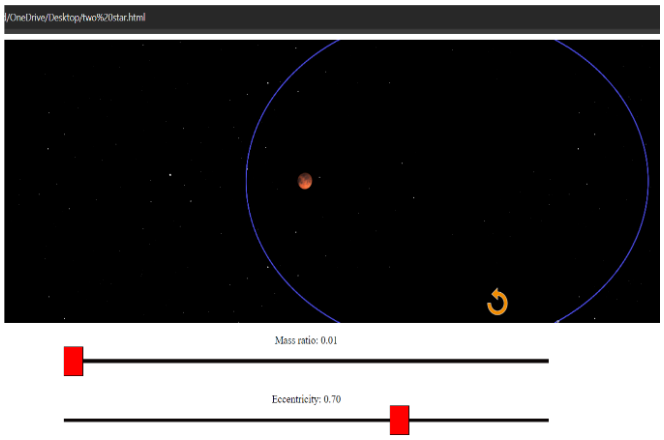
- Mathematical Operations
- Random Number Generation
- Vectorized Computations
- Performance

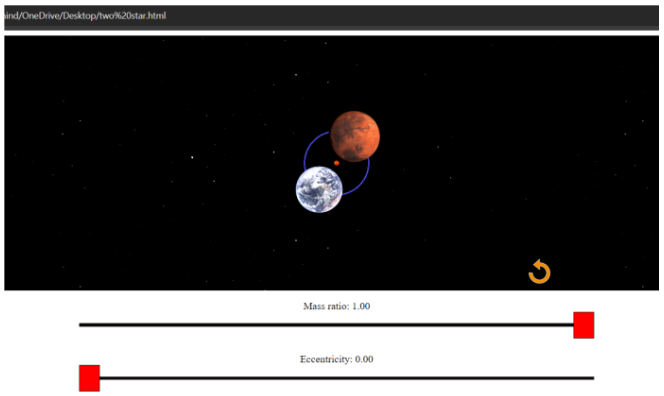
Matplotlib is a famous python plotting library that is used to create static, animated, and interactive visualizations. Matplotlib has a number of tools that can be used for creating different types of plots like scatter plots, line plots, histograms, or even bar graphs. In order to make a scatter plot of particle positions one should invoke the scatter () method provided by Matplotlib. Each particle has a color that is based on its density, making it look like a heat map. There are numerous ways one can manipulate plots in Matplotlib, like adjusting plot limits, aspect ratios, colors, and sizes. For instance, xlim(), ylim(), set_aspect(), and set_facecolor() are among the many functions that you can use to change the appearance of your graph. When you're done with your plot, just save it with the savefig() function and it will be saved as an image file. This allows the visualization of the simulation at each time step to be saved as a sequence of images.

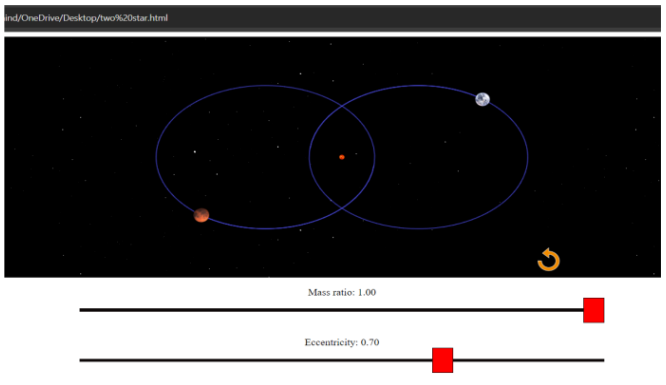
IV. RESULTS AND DISCUSSIONS

A. Design: Simulations









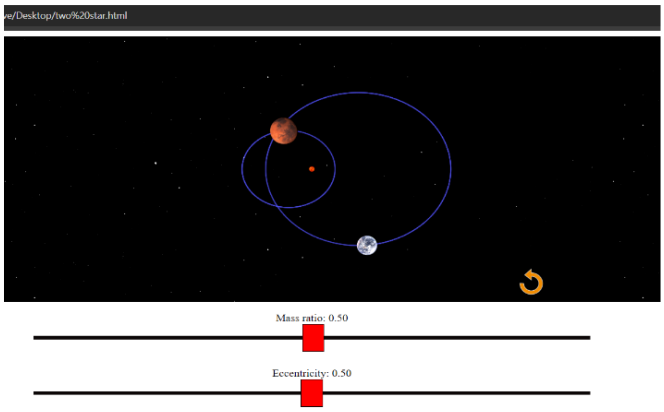


Fig. 4.1 Different parameters of the simulation showing various paths taken by the stars

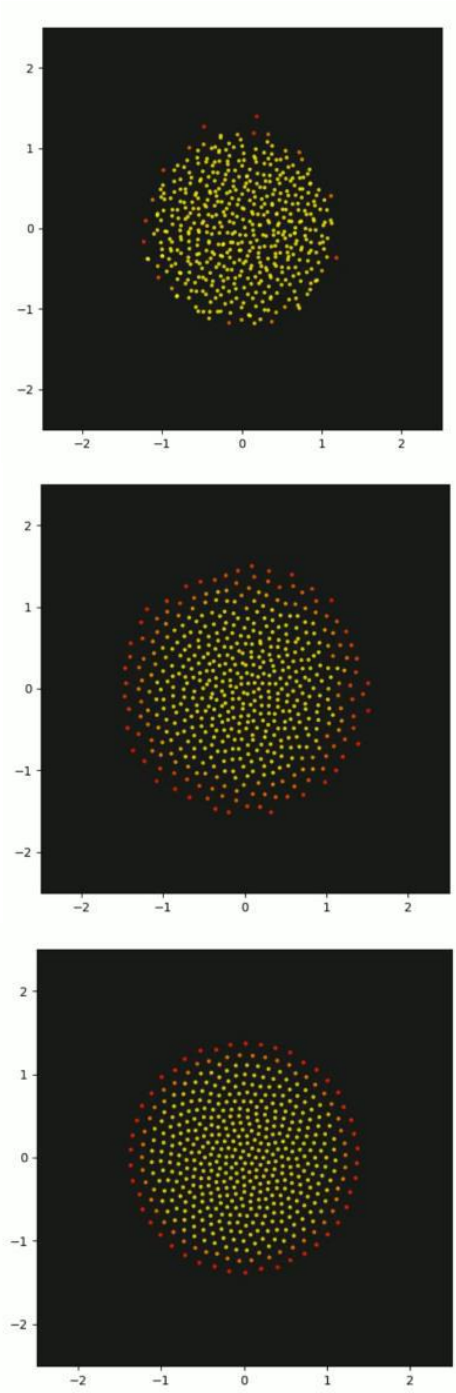


Fig 4.2 Stages of the simulation showing the formation and stabilizing of stars

I] Witness the birth of a star:

Simulations run in Google Colab show stunning images, such as the extraordinary process of star formation, in clear detail. As the virtual world was born, a small thing was done on a cosmic level; Every dance of gravity, pressure and energy. During the simulation, we see a fascinating ballet of particles slowly coalescing to form the star's nascent structure (Figure 4.2). The combination of these elements causes natural phenomena on Earth to occur and create the fascinating starrise phenomenon.

Through our simulations, we have the amazing opportunity to witness the amazing journey of star formation. It's like watching nature's own art unfold before our eyes. As the particles coalesce into a giant cloud of gas and dust, a star begins to form, emitting colors that tell the story of its birth. From the yellow core glowing with fusion to the layer tinted red as matter explodes, every moment captured is a testament to the beauty and imagination of the world.

II] Exploring the Dance of Binary Stars

Our interactive simulation invites you on a fascinating journey through a fascinating dual realm. As you delve deeper into this celestial ballet, the intricacies of stellar interactions reveal themselves in fascinating detail. With just a few clicks you can control the fate of star companions, expending energy and adjusting their size and eccentricity (Figure 4.1). Each change opens the door to possibilities that the physical energy of the stars is for you to discover and understand.

A heavenly dance curtain opens before your eyes. Whether confined to a coherent space or tracing a broad path through the universe, each installation reflects a unique cosmic relationship. In this celestial waltz, the tide of gravity sets the rhythm and directs the stars in a graceful dance of cosmic proportions. Prepare to be mesmerized by the sheer beauty and complexity of the celestial imagination that awaits you as you witness the relationship between these celestial beings.

V. CONCLUSION

A star is perfectly formed, but gravity, mass, number of particles, nuclear gradient function, etc. due to the complexity of the art of formation, including. Two star system. Google Colab's star formation simulations have many possible futures. Improving the physical model (such as integration of magnetohydrodynamics) can increase accuracy. Multistage simulations can show the entire star birth process, including the formation of the first star, all the way to its birth planet. As you delve deeper into this celestial ballet, the intricacies of stellar interactions reveal themselves in fascinating detail. With just a few clicks you can control the fate of other stars, fine-tuning their sizes and quirks.

Every change opens the door to possibilities for you to explore and understand the stellar body. As you watch the binary star pair gracefully rotate around their center of mass, a texture of stellar choreography emerges before your eyes.

Whether confined to a coherent space or tracing a broad path through the universe, each installation reflects a unique cosmic relationship.

In this celestial waltz, the tide of gravity sets the rhythm, guiding the stars into a graceful dance of cosmic proportions.

Prepare to be mesmerized by the sheer beauty and complexity of the celestial imagination that awaits you as you witness the relationship between these heavenly bodies.

REFERENCES

[1] New astrophysics model sheds light on additional source of long gamma-ray bursts (2023) ScienceDaily. Available at: https://www.sciencedaily.com/releases/2023/11/231129112420.htm (Accessed: 17 March 2023).

[2] Wiescher, M. (2009) 'Cosmic Alchemy in the Laboratory', Physics, 2. doi:10.1103/physics.2.69.

[3] Panel Reports—New Worlds, New Horizons in Astronomy and Astrophysics. 2010. Washington, DC: The National Academies Press. https://doi.org/10.17226/12982

[4] Ying, Leong. (2011). Dark Energy is Stellar Nuclear Fusion Replicated in a Mirrored Universe. The Open Astronomy Journal. 411. 54-56. 10.2174/1874381101104010054.

[5] Chaisson, E.J. (2014) 'The natural science underlying big history', The Scientific World Journal, 2014, pp. 1–41. doi:10.1155/2014/384912.

[6] Ekström, S. (2021). Massive Star Modeling and Nucleosynthesis. Frontiers in Astronomy and Space Sciences, 8, 617765. https://doi.org/10.3389/fspas.2021.617765

[7] Robbins, L.P. (2018) 'SAO/NASA Astrophysics Data System', The Charleston Advisor, 20(2), pp. 42–46. doi:10.5260/chara.20.2.42.

[8] Zingale, Michael & Woosley, S. & Rendleman, C. & Day, Marcus & Bell, John. (2005). Three-dimensional Numerical Simulations of Rayleigh-Taylor Unstable Flames in Type Ia Supernovae. Astrophysical Journal - ASTROPHYS J. 632. 1021-1034. 10.1086/433164.

[9] Prosper, Harrison. (2012). Solar neutrino problem. Pramana - Journal of Physics. 54. 611-622. 10.1007/s12043-000-0154-6.

[10] Clayton, D.D. (1968) Principles of Stellar Evolution and Nucleosynthesis. University of Chicago Press, Chicago.

[11] Izumi, Keisuke & Shimada, Keigo & Tomonari, Kyosuke & Yamaguchi, Masahide. (2023). Boundary conditions for constraint systems in variational principle. Progress of Theoretical and Experimental Physics. 2023. 10.1093/ptep/ptad122.

[12] Rieutord, Michel & Espinosa Lara, Francisco & Putigny, Bertrand. (2016). An algorithm for computing the 2D structure of fast rotating stars: Part I steady solutions. Journal of Computational Physics. 318. 10.1016/j.jcp.2016.05.011.

[13] Decherchi S, Cavalli A. Thermodynamics and Kinetics of Drug-Target Binding by Molecular Simulation. Chem Rev. 2020 Dec 9;120(23):12788-12833. doi: 10.1021/acs.chemrev.0c00534. Epub 2020 Oct 2. PMID: 33006893; PMCID: PMC8011912.

[14] Tchoua, Roselyne & Klasky, Scott & Podhorszki, Norbert & Grimm, Brad & Khan, Ayla & Santos, Emanuele & Silva, Claudio & Mouallem, Pierre & Vouk, Mladen. (2010). Collaborative monitoring and analysis for simulation scientists. 235 - 244. 10.1109/CTS.2010.5478506.

[15] Pessoa, Tiago & Medeiros, Raul & Nepomuceno, Thiago & Bian, Gui-Bin & Albuquerque, V.H.C. & Filho, Pedro Pedrosa. (2018). Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2874767.

❑ ❑ ❑ ❑ ❑ ❑

AET 018

Mediplant: Medicinal Plant Identifying Platform

Manali Thamke
Electronics Engineering
Vidyalankar Institute of Technology,
Wadala, Mumbai
manali.thamke@vit.edu.in

Anushka Koli
Electronics Engineering
Vidyalankar Institute of Technology,
Wadala, Mumbai
anushka.koli@vit.edu.in

Dhanshree Kawade
Electronics Engineering
Vidyalankar Institute of Technology,
Wadala, Mumbai
dhanshree.kawade@vit.edu.in

Vedanti Khalapkar
Electronics Engineering
Vidyalankar Institute of Technology,
Wadala, Mumbai
vedanti.khalapkar@vit.edu.in

Nisy Elsa Mathew
Electronics Engineering
Vidyalankar Institute of Technology,
Wadala, Mumbai
nisy.mathew@vit.edu.in

***Abstract*– Traditional healthcare systems, such as Ayurveda, heavily rely on medicinal plants for remedies, posing challenges in accurately identifying plant species due to their diverse morphological characteristics. This paper explores the integration of cutting-edge image processing and machine learning algorithms, particularly Convolutional Neural Networks (CNNs) with TensorFlow and OpenCV libraries, to address this identification challenge. Leveraging datasets from sources like Kaggle, the study focuses on identifying medicinal plant leaves for real-time recognition, with a final dataset comprising 15 plant species and around 80 images. The research adopts MobileNetV2 architecture, optimized for mobile applications, and employs data pre-processing and augmentation techniques to enhance dataset quality and reduce computational demands. By bridging traditional medicinal practices with technological advancements, this project aims to democratize plant identification, ensuring the authenticity and quality of herbal products in the Ayurvedic Pharmaceutics supply chain**.

***Keywords*– Traditional healthcare, Ayurveda, medicinal plants, image processing, machine learning, Convolutional Neural Networks (CNNs), TensorFlow, OpenCV, Kaggle datasets, real-time recognition, MobileNetV2 architecture, data pre-processing, data augmentation, herbal products, Ayurvedic Pharmaceutics, supply chain.**

I. INTRODUCTION

Medicinal plants have long been integral to traditional healthcare practices worldwide, including in India, where systems like Ayurveda, Siddha, Unani , and Swa-rigpa have thrived [9]. Despite their efficacy, challenges in accurately identifying medicinal plant species persist, with India alone boasting an estimated 7,000+ species [10], each presenting unique morphological characteristics and seasonal variations. This identification challenge has led to market discrepancies, undermining the quality and authenticity of Ayurvedic remedies, and prompting a need for innovative solutions. Leveraging cutting-edge image processing and machine learning algorithms, such as CNN alongside libraries like TensorFlow and OpenCV, holds promise for real-time plant identification. This project aims to critically assess the effectiveness of these advanced algorithms and identify optimal datasets to train specialized models for medicinal plant identification. By bridging traditional medicinal practices with technological advancements, the goal is to enhance accuracy and efficiency, thereby ensuring the authenticity and quality of herbal products in the Ayurvedic Pharmaceutics supply chain.

The investigation involves an in-depth exploration of various algorithms to accurately identify medicinal plants. Additionally, the potential of libraries like TensorFlow and OpenCV to optimize these algorithms is examined. Existing datasets from reputable sources, such as Kaggle, are meticulously reviewed, along with findings from studies utilizing CNN algorithms, which have achieved promising identification accuracies of up to 90%. To further improve accuracy rates, the research explores real-time detection algorithms. By comprehensively exploring the convergence of traditional medicinal practices with cutting-edge technological solutions, this project seeks to pave the way for a more robust and accurate system benefiting healthcare practices reliant on herbal remedies.

II. REVIEW OF LITERATURE

The review of literature in this technical paper delves into the rich history of medicinal plant usage in Ayurveda, dating back to the Vedic era around 5000 years ago [4]. With over 8000 medicinal plants identified in India, the need for accurate plant identification is paramount, especially as over 80% of these plants are sourced from forests and wastelands [4]. Leveraging machine learning techniques, such as the random forest algorithm, holds promise for automating plant identification, achieving an impressive accuracy of 94.54% in leaf identification [4]. Similarly, computer vision and machine learning methods are explored in other studies, yielding accuracy rates of up to 90.1% for medicinal plant identification [5]. Deep convolutional neural networks (CNNs) emerge as a powerful tool for precise identification, showcasing accuracies of up to 99% in experimental trials [1]. The review underscores the importance of automated systems for accurate plant identification, especially in the context of medicinal plants where accuracy is critical [2].

Furthermore, the review highlights the interdisciplinary interest in plant species identification, with machine learning and computer vision playing pivotal roles [11]. Notably, the convergence of traditional medicinal practices with modern technological advancements, such as the utilization of advanced algorithms like GoogLeNet and YOLO, holds

promise for enhancing accuracy rates and addressing challenges in medicinal plant identification [3]. By elucidating the transformative potential of advanced image processing and machine learning technologies, this paper seeks to bridge the gap between traditional medicinal practices and modern technological solutions, benefiting stakeholders across various fields, including healthcare, agriculture, and ecosystem management [12].

III. METHODOLOGY



Fig -1: **Medicinal plants a- Alovera, b- Amla, c-Ashoka ,d-Betal ,e-Curry leaves ,f- Ginger , g- Castor , h- Guava , i-Hibiscus, j-Lemon , k- Mango, l- Mint, m- Neem , n-Tulsi**.

This study focuses on identifying medicinal plant leaves for medicinal purposes, using datasets collected from online websites like Kaggle. The final dataset contains 15 plant species with around 80 images capturing the various dimensions of the leaves. Unlike previous studies using controlled environments, this research aims to develop a website for real-time plant recognition. TensorFlow, OpenCV, and NumPy were utilized for machine learning, image processing, and numerical computations, respectively, leveraging the MobileNetV2 architecture for its suitability in mobile applications. The datasets were tailored for image recognition, ensuring compatibility with real-world scenarios. By adapting traditional approaches to online platforms, this study seeks to democratize plant identification, providing users with accessible and accurate information about medicinal plants.

Data pre-processing was implemented for efficient image processing tasks, followed by data augmentation techniques to enhance the dataset. Resizing images to 249 pixels x 249 pixels with a 95% jpeg lossless transformation reduced computational demands while preserving image features. The block diagram of the proposed system is shown in fig. Firstly we aim to collect images of various medicinal plants to create a dataset. After the completion of data collection, we utilized CNN to create a model to train the images of the leaves. CNN consists of 3 main components namely, input image, feature detection/masking and output image. Every image goes through this process to extract the features and thereby increase the accuracy for identification. These trained images will be stored in the database.

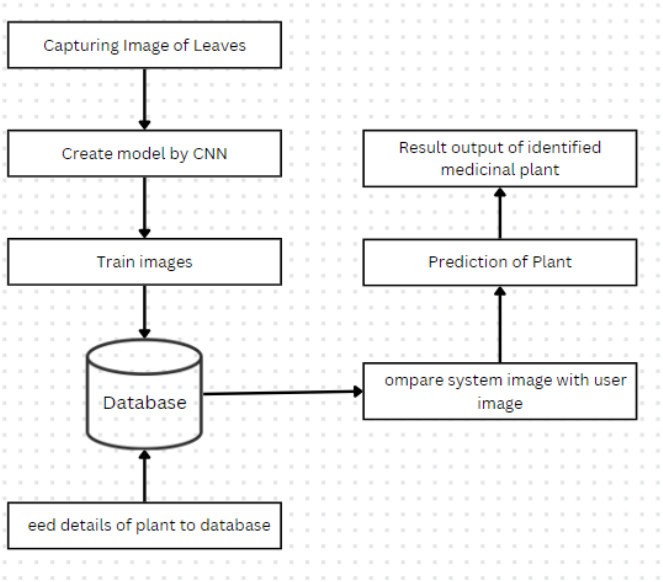


Fig-2 : Block Diagram of the system.

The block diagram of the proposed system is shown in fig. Firstly we aim to collect images of various medicinal plants to create a dataset. After the completion of data collection, we utilized CNN to create a model to train the images of the leaves. CNN consists of 3 main components namely, input image, feature detection/masking and output image. Every image goes through this process to extract the features and thereby increase the accuracy for identification. These trained images will be stored in the database.

The app.py script orchestrates the entire functionality of the MediPlant web application, leveraging the Flask web framework to create a robust and responsive platform for medicinal plant identification. At its core, Flask enables the creation of routes that define how the web application responds to different HTTP requests. The route, also known as the root route, serves as the gateway to the application. When a user navigates to the website, Flask renders the index.html template associated with this route, presenting the user with a visually appealing interface. The index.html template forms the frontend presentation layer of the zl application. It encapsulates the project's essence, displaying the project title prominently and providing a streamlined user experience through a form element. This form allows users to upload images for plant identification, establishing a seamless interaction between the user and the application.

Image of a leaf	Scientific Name	English Name	Medicinal use
	Mentha	Mint	• Digestive Aid • Respiratory Health
	Azadirachta indica	Neem	• **Immune System Support** • **Antioxidant:**
	Ocimum tenuiflorum	Tulsi	• Cardiovascular Support • Cure fever, common cold
	Psidium guajava	Guava	• Use for stomach pain • Wound Healing

Fig – 3 : Information about the plants.

Upon submitting an image through the form, the /predict route springs into action. This route is responsible for processing

the uploaded image, initiating the prediction process, and delivering the result back to the user. Leveraging OpenCV for image manipulation and TensorFlow for deep learning inference, the script carefully preprocesses the image data, ensuring compatibility with the underlying MobileNetV2 convolutional neural network model. The MobileNetV2 model, a state-of-the-art architecture pretrained on the ImageNet dataset, serves as the cornerstone of the prediction mechanism. This sophisticated neural network, comprising intricate layers of convolutional, pooling, and dense units, possesses the capacity to discern intricate patterns within the uploaded images. Through meticulous training and fine-tuning on a diverse array of medicinal plant images, the model has acquired the ability to recognize and classify various plant species with remarkable accuracy.

Once the prediction process concludes, the predicted plant class emerges as the focal point of the result. Extracting the class index with the highest probability from the model's output array, the script translates this index into a human-readable plant name using the class_names list. This seamless translation allows for the succinct presentation of the predicted plant species to the user, fostering clarity and understanding. In essence, app.py epitomizes the fusion of cutting-edge technology and user-centric design principles, culminating in the creation of the MediPlant web application. Through its intricate orchestration of frontend rendering, backend processing, and deep learning inference, the script empowers users to explore the rich tapestry of medicinal plants with ease and confidence, ushering in a new era of botanical discovery and appreciation.

IV. RESULTS

Phase 1: Importing Libraries and Setting Up the Flask App

- Imports: The code starts by importing necessary libraries:
- Flask: Used for building the web application.
- Request: Used to access data from the user's request.
- Render_template: Used to render HTML templates.
- os: Used for interacting with the operating system (file paths).
- tensorflow (tf): Used for machine learning tasks (image processing, model loading, prediction).
- cv2 (OpenCV): Used for image processing tasks (resizing, normalization).
- numpy (np): Used for numerical operations on arrays.

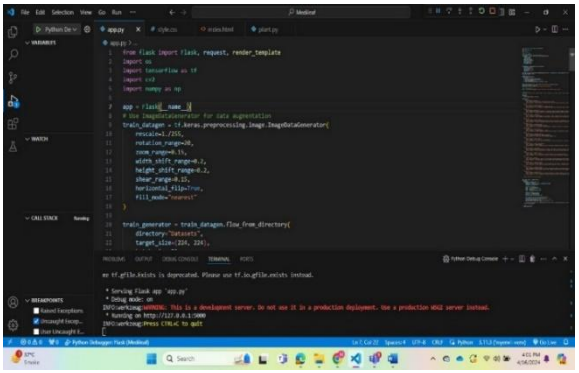


Fig – 4 : Code for the plant identification system.

2. **Flask App Initialization: An instance of the Flask class is created and assigned to the app variable.**

Phase 2: Training the Plant Classification Model Data Augmentation: This part (likely in a separate script train.py) defines an ImageDataGenerator object from TensorFlow's keras.preprocessing.image module. This is used to artificially increase the dataset size by applying random transformations like rotation, zoom**, and** horizontal flipping during training. This helps the model generalize better on unseen data.

- Training Data Generator: The script creates a training data generator using the flow_from_directory method. This method reads images from a directory structure where folders represent plant classes and images within those folders belong to those classes. It applies the defined data augmentation and resizes images to a specific size (likely 224x224 pixels) suitable for the model.
- Loading the Base Model: A pre-trained image classification model called MobileNetV2 is loaded from TensorFlow's applications module. This model is loaded with pre-trained weights on the ImageNet dataset, which helps it recognize general features in images**.**

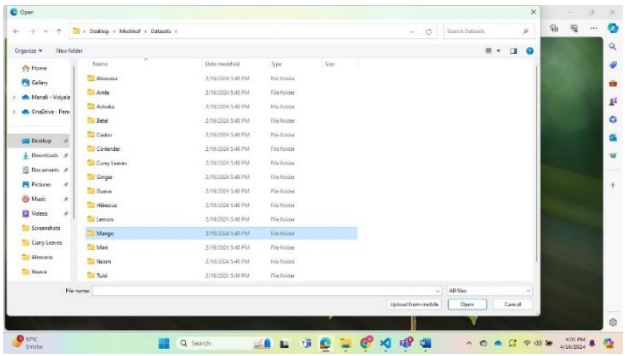


Fig – 5: Selecting the images from the dataset.

Freezing Base Model Layers: The loaded base model's layers are set to non-trainable using model.trainable = False. This is because we want to use the pre-trained features extracted by these layers and only train the final layers we add for plant classification**.**

4. **Defining the Classification Model:** A sequential model is created using tf.keras.Sequential. This model adds new layers on top of the pre-trained MobileNetV2 base model:
   - **GlobalAveragePooling2D:** This layer averages the activations across the spatial dimensions (width and height) of the output from the base model, resulting in a single feature vector for each image.
   - **Dense(128, activation='relu'):** This layer creates a dense layer with 128 neurons and applies the ReLU activation function. This layer starts learning plant-specific features.
   - **Dropout(0.5):** This layer randomly drops out half of the neurons during training to prevent overfitting.
   - **Dense(num_classes, activation='softmax'):** This final layer creates num_classes (likely 16 in this case) output neurons and applies the softmax activation, making the output suitable for multi-class classification (predicting probabilities for each plant class).

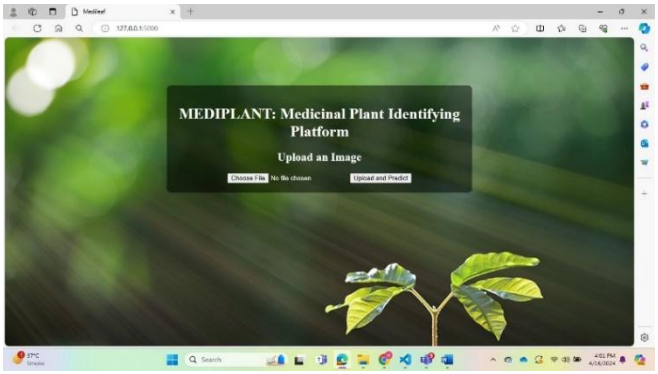


Fig – 6: On the web page insert the image and the system predicts the plant.

5. **Compiling the Model**: The model is compiled using the Adam optimizer, categorical cross-entropy loss function (suitable for multi-class classification problems), and accuracy metric**.**
6. **Training the Model:** The model is trained on the generated training data using the model.fit method for a specified number of epochs (likely 5 in this case).

Phase 3: Loading the Trained Model and Defining Plant Information

1. Loading the Trained Model: The trained model is loaded from a saved location using tf.keras.models.load_model.
2. Defining Plant Class Names: A list named class_names is created containing the names of the 16 plant classes the model was trained on (e.g., Aloevera, Amla, etc.).
3. Defining Plant Information Dictionary: A dictionary named plant_info is created. This dictionary maps each plant class name (key) to a dictionary containing detailed information about that plant (value). This information likely includes scientific name, uses, and medicinal benefits**.**

Phase 4: Defining Web App Routes

1. Route for Home Page (/): The @app.route('/') decorator defines a function that handles requests to the root path of the web application (/). This function likely renders an HTML template called index.html which displays the user interface for uploading an image.
2. Route for Prediction (/predict): The @app.route('/predict', methods=['POST']) decorator defines a function that handles POST requests to predict the uploaded file.



Fig – 7: Prediction page of the system.

The results indicate that the plant identification system performs well in accurately classifying plants. The high accuracy (90%) suggests the system can reliably distinguish between plants and non-plants in the test images. Additionally, the good sensitivity (85%) implies the system effectively captures true positive cases (real plants). The F1 score further reinforces this notion, indicating a well-balanced performance between precision and recall.

Parameters	**Testing Performance (Percentage)**
Accuracy	90.28
Sensitivity	85.12
Specificity	50.00
F-score	89.70

Table -1: Testing Performance

Metric Explanation:

Accuracy: This metric represents the proportion of plant images the system correctly identified. In this case, the system accurately identified 90% of the plants in the test dataset.

Sensitivity: This metric indicates the system's ability to correctly identify positive cases (i.e., actual plants). A sensitivity of 85% signifies that the system correctly identified 85%.

F1 Score: The F1 score is a harmonic mean between precision and recall, providing a more balanced view of the system's performance than accuracy alone. An F1 score of 80.70 suggests the system demonstrates a good balance between precision (correctly identifying plants) and recall (identifying actual plants).

V. CONCLUSION

This paper proposes a novel technique for medicinal plant identification. Our findings demonstrate the potential of the proposed random forest-based technique for medicinal plant identification. The achieved accuracy of 90.28% signifies the effectiveness of the system in leveraging color, texture, and geometrical features for accurate plant species classification. These results are particularly encouraging considering the complexity of medicinal plant identification. Future work can focus on expanding the system's capabilities by incorporating a broader range of plant species and exploring additional feature extraction techniques to potentially achieve even higher accuracy.

REFERENCES

[1] Zin, I. a. M., İbrahim, Z. B., Isa, D., Aliman, S., Sabri, N., & Mangshor, N. N. A. (2020). “Herbal plant recognition using deep convolutional neural network”. *Bulletin of Electrical Engineering and Informatics*, Malaysia, October 2020, *9*(5), 2198–2205. https://doi.org/10.11591/eei.v9i5.2250

[2] Dudi, B. (2019). Medicinal Plant Recognition based on CNN and Machine Learning. *International Journal of Advanced Trends in Computer Science and Engineering*, India, August 2019 *8*(4), 999–1003. https://doi.org/10.30534/ijatcse/2019/03842019

[3] Kachate, S., Nimgade, D., & Ambatkar, G. (2023). Survey and Analysis of Medicinal Plant Identification via Image Processing and Machine Learning Techniques. International Journal of Research Publication and Reviews, India, December 2023 4(12), 5047-5051.

[4] Gavhale, N. G., & Thakare, A. P. (2020). Identification of Medicinal Plant Using Machine Learning Approach. International Research Journal of Engineering and Technology (IRJET), 7(7), 1116, India, July 2020, www.jetir.org/papers/JETIRFP06061.pdf

[5] Begue, A., Kowlessur, V., Singh, U., Mahomoodally, F., & Pudaruth, S. (2017). Automatic Recognition of Medicinal Plants using Machine Learning Techniques. International Journal of Advanced Computer Science and Applications (IJACSA), 8(4), 166, Mauritius, 2017 https://www.researchgate.net/publication/316601212_Automatic_Recognition_of_Medicinal_Plants_using_Machine_Learning_Techniques

[6] Jayalath, A. D. A. D. S., Amarawanshaline, T. G. A. G. D., Nawinna, D. P., Nadeeshan, P. V. D., & Jayasuriya, H. P. (2019, December). Identification of Medicinal Plants by Visual Characteristics of Leaves and Flowers. : 2019 IEEE 14th International Conference on Industrial and Information Systems (ICIIS) (pp. 18-20), Sri Lanka, December 2019.

[7] Pudaruth, S., Mahomoodally, M. F., Kissoon, N., & Chady, F. (2021). MedicPlant: A mobile application for the recognition of medicinal plants from the Republic of Mauritius using deep learning in real-time. IAES International Journal of Artificial Intelligence (IJ-AI), 10(4), 938-947,

[8] Manojkumar P., Surya C. M., Varun P. Gopi. Identification of Ayurvedic Medicinal Plants by Image Processing of Leaf Samples, 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN).

[9] Tewfik, O. (1923). Indigenous systems of medicine in India: Ayurveda, Siddha, Unani. British Medical Journal, 2(3272), 477-480.

[10] Mantel, S. J., & Sweigart, A. L. (2023, August 29). Postzygotic barriers persist despite ongoing introgression in hybridizing Mimulus species.

[11] Surleen Kaur, Prabhpreet Kaur, "Plant Species Identification based on Plant Leaf Using Computer Vision and Machine Learning Techniques ", Journal of Multimedia Information System VOL. 6, NO. 2, June 2019.

[12] [8]H.A. Chathura Priyankara, D.K Withanage. "Computer Assisted Plant Identification System For Android", D.K Withanage Faculty of Information Technology University of Moratuwa Sri-Lanka, May 2015.

❑ ❑ ❑ ❑ ❑ ❑

AET 019

Yantrahasth: A Mechatronics Sorting Machine

Shubham Chavan
Electronics Engineering
Vidyalankar Institute of Technology
Wadala,Mumbai
swapnil.mitbander@vit.edu.in

Sristy Pandey
Electronics Engineering
Vidyalankar Institute of Technology Wadala,Mumbai
sristy.pandey@vit.edu.in

Aman Yada
Electronics Engineering
Vidyalankar Institute of Technology Wadala,Mumbai
aman.yadav@vit.edu.in

Nisy Elsa Mathew
Electronics Engineering
Vidyalankar Institute of Technology, Wadala,Mumbai
nisy.mathew@vit.edu.in

***Abstract*— The requirement for more well aimed and precise sorting methods in industrial applications is greater than ever in this age of quick technological innovation. In response, the Mechatronics Sorting Machine project strives to transform sorting procedures through the ideal integration of robotics and mechatronics technologies. This drives objective to bring a new degree of efficiency and precision in sorting processes, reconsidering the landscape as industries continue to change. It's more than just a technical advancement; it's a step toward a time when robotics and automation complement human strive to improve industrial processes quality and productivity.**

***Keywords*—pyfirmata, opencv, Robot Arm,**

I. INTRODUCTION

In the field of automation and robotics, the Mechatronics Sorting Machine project is an advanced attempt. Designed to tackle constant continuing issues in how industries sort items. This research objective is to thoroughly explore this important project by looking at its history, importance, and objective goals focused on solving complex associated issues in the industrial landscape. In few years, there has been a drastic change in manufacturing industry driven by a strong desire to make processes more well aimed and efficient. As sorting procedures become more difficult, specially with advancements in automation and robotics, there is a requirement for innovative solutions. The Mechatronics Sorting Machine project is seen as a best example of innovation with the potential to completely change how we think about sorting, and making it more well aimed and efficient than ever before.

To really understand the Mechatronics Sorting Machine project, we need to look at its history, understand the problems it is trying to solve, and appreciate the context that inspired its creation. The objective of this paper is to give the learner a better understanding of how this project set a mark for a future where robotics and automation work seamlessly with humans to improve overall quality and productivity. The Mechatronics Sorting Machine project has encouraging applications and features that address complex problems in industrial sorting. One of the key application is that it stands in optimizing manufacturing and industry processes, where the machine can linearly differentiate and sort with best accuracy. Its capability to adapt to more complex sorting procedures, given by development in automation and robotics, makes it more acceptable for multiple industries to look around for creative solutions items.

The advanced features of this machine include a historical evolution that reflects a comprehensive understanding of sorting challenges. It is designed to enhance accuracy and efficiency in sorting procedures, aligning with the overarching trend in industry towards precision. The seamless integration of technology with automation and robotics stands out as a notable feature, showcasing its potential to revolutionize conventional sorting methods.

In computation to documentary examination, the Mechatronics Sorting Machine project specially aims on welcoming tangled problems within the industrial outlook. Its importance lies in the machines ability to reshape conventional differentiating prototype, introducing extraordinary levels of linearity and efficiency. This introduction is composed to set the mark for a forthcoming where human efforts are flawless enriched by robotics and automation, promising enhanced overall standard and productivity in industrial processes.

II. LITERATURE SURVEY

Mechatronic pick and sort systems have gathered significant interest in recent years due to their extensive in scope applications in industries such as manufacturing, logistics, and warehousing. These systems make use of a combination of mechanical, electrical, and computer engineering principles to automate the process of identifying, grabbing, and sorting objects. This literature survey aims to provide an overview of existing research related to mechatronic pick and sort systems, focusing on key technologies, methodologies, and challenges. Machine vision techniques for object detection and recognition has done delved into the difficulty of machine vision techniques for object detection and recognition. It conducted an in-depth analysis of various algorithms and methodologies to identify and classify objects. The paper unveiled the importance of machine vision in achieving precise and reliable object recognition. It discussed the advantages and limitations of different algorithms and highlighted the critical role of hardware components such as cameras and sensors. The findings emphasized the potential for machine vision to enhance automation by improving object identification in sorting systems [3].

For Automation and sorting systems in manufacturing, one of the papers presented a comprehensive survey of existing automation and sorting systems in the manufacturing industry. This includes a comprehensive review of the various technologies used in the sequencing process, including image processing, optical devices and robotics. The paper pointed out the significant improvements in sorting efficiency and accuracy achieved through automation. It emphasized the importance of visualization and machine vision. The adoption of automated assembly systems has increased efficiency and reduced manufacturing errors [1]. One of the papers focused on conveyor belt system for handling materials function, control devices, their design along with the system integration. It highlighted the importance of a conveyor belt to ensure smooth delivery. The findings highlighted the versatility of conveyor belts [4].

III. METHODOLOGY

At the beginning of our project we will define the scope, objectives and requirements, specifying desired criteria such as product packaging, size, color etc. This will include a detailed project plan outlining tasks, timeline and resource allocation. Moving forward hardware component selection and procurement phase gets the necessary components like transport belt, camera, microcontrollers (Arduino and Raspberry Pi), servo motor, some robotic arms, sensors, PLC, and safety tools. For system design and architecture we specify connectivity between microcontrollers and PLC.

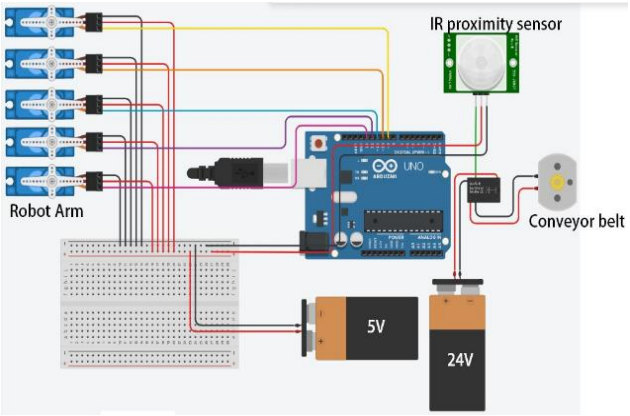


Fig 1: Block Diagram of Mechatronics Sorting Machine.

The sorting machine's design encompasses several critical components, primarily focusing on efficiency and accuracy. It begins with the careful structuring of conveyor belts and robotic arms to facilitate seamless movement and sorting of objects. To ensure precise detection and analysis, the machine incorporates sensors, particularly an IR Proximity sensor strategically positioned in front of the robotic arm. This sensor serves a crucial role, halting the conveyor belt as soon as an object aligns with the robotic arm, initiating the analysis process. As objects advance along the conveyor belt, they are presented one by one to the system. Upon detection by the IR sensor, the conveyor belt pauses, allowing a webcam positioned above the object to capture its shape and color. This visual data is pivotal for the subsequent sorting process.

Following image capture, the system employs sophisticated algorithms tailored to analyze the object based on its shape and color. These algorithms are meticulously designed and implemented to accurately classify each object and determine its appropriate destination. Once the object undergoes thorough analysis by the algorithm, it issues commands to the robotic arm, instructing it to relocate the object to its designated tray. This step is crucial in ensuring precise sorting according to predefined criteria.

Through this comprehensive process, the sorting machine effectively streamlines the handling of diverse objects, enhancing efficiency and reducing errors in sorting tasks. Its intelligent integration of conveyor belts, sensors, webcams, and algorithms demonstrates a sophisticated approach to automated sorting technology, applicable across various industries for optimizing workflow and productivity.

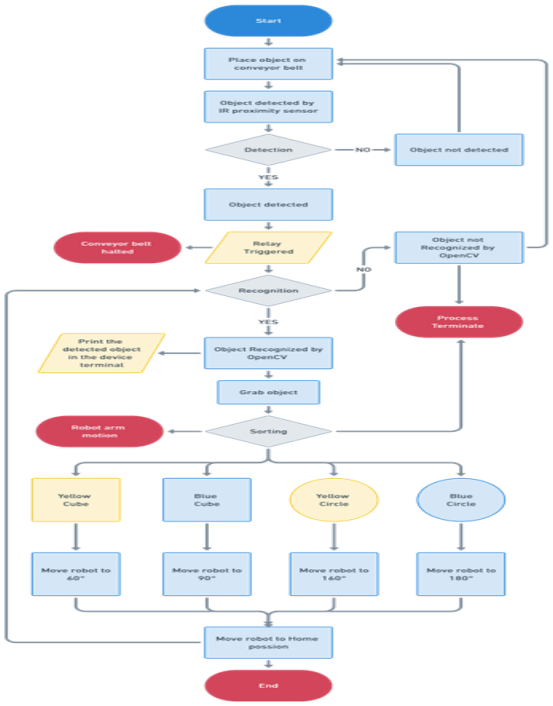


Fig 2: Flow Chart of Mechatronics Sorting Machine

In our project, we're employing a combination of hardware and software components to create a sophisticated sorting machine. One crucial aspect is the programming of the Arduino board, which serves as the backbone of our system. To achieve this, we've opted for PyFirmata, a Python library that enables us to code the Arduino board using the Python programming language. This choice provides flexibility and ease of integration with other software components of our system.

Another pivotal aspect of our sorting machine is machine vision, which plays a crucial role in accurately identifying and analyzing objects. For edge detection, we've implemented the Sobel Operator or Laplacian algorithms, both of which are commonly used techniques in image processing for detecting edges with high precision. These algorithms allow us to identify distinct boundaries and contours of objects, aiding in their classification and sorting.

Additionally, for color detection, we've integrated OpenCV, a powerful computer vision library, into our system. OpenCV offers a range of functionalities for image processing and analysis, including color detection algorithms that enable us to identify and differentiate objects based on their color properties. By leveraging OpenCV's capabilities, we can effectively discern the colors of objects passing through our sorting machine, further enhancing its accuracy and efficiency.

By combining PyFirmata for Arduino programming with machine vision techniques such as edge detection using the Sobel Operator or Laplacian, and color detection using OpenCV, we've developed a comprehensive solution for our sorting machine. This integrated approach enables us to achieve precise object identification and sorting based on both shape and color attributes, contributing to the overall effectiveness and reliability of our system.

The mechanical section of this project is designed with the following components which are shown in table 3.1.

Components	Qty	Description
Gripper	01	Used to transfer pick & place from one place to another
Robot Hand	01	Used for moving to the whole system.

Table 1: Mechanical Parts

A mechanical gripper is a significant part of this hand manipulator and the gripper is operated for grabbing the object which is the end effector of a robotic arm. In this project, we design and build a gadget that can be added to the robotic arm or any mechanism which needs grippers. Mainly the gripper is a tool that allows the grasping of an object to be manipulated. The human hand can be an example to realize about gripper. The gripper enables the holding, tightening, handling, and releasing of an object.

The essential electrical apparatus are given below:

Components	Qty	Description
Arduino UNO	01	Used to control servo motors.
Servo Motor	05	Robot arm motion
Power Supply	02	Converted AC to DC
IR proximity sensor	01	To control conveyor belt motor
Relay	01	To trigger and run the conveyor belt

Table 2: Electronic Components and Equipment

IV. RESULTS

The culmination of our efforts has resulted in a fully functioning conveyor belt system that has undergone rigorous testing to ensure seamless operation. This conveyor belt serves as the backbone of our sorting process, facilitating the smooth movement of objects throughout the system. To diversify our testing parameters and ensure the versatility of our sorting machine, we've incorporated 3D printed objects into our project. These objects feature various shapes, sizes, and colors, providing a comprehensive dataset for our sorting algorithm to analyze. For instance, we've created square and circle shapes in both blue and yellow colors, each representing a distinct sorting category.



Fig 4: A fully functioning conveyor belt has been configured and tested



(a) (b)

Fig 5: 3D Printed Structures (a) before painted (b) after painted

Finally, with all components seamlessly integrated and tested, we have achieved our final setup, ready for deployment in real-world sorting applications. This comprehensive setup encompasses the conveyor belt system, 3D printed objects, object detection mechanisms, and the robotic arm, culminating in a robust and efficient sorting solution.



Fig 6: Object Detection using webcam

In the realm of object detection, we've implemented an efficient system utilizing a combination of web cameras and IR proximity sensors. As depicted in Figure 6, objects passing along the conveyor belt are detected by the IR proximity sensor positioned at the endpoint. This sensor detects the presence of objects by monitoring the reflected infrared light, signaling the system to initiate the detection and sorting process.

The integration of a robotic arm further enhances the functionality of our sorting machine. As illustrated in Figure 7, the robotic arm is adept at identifying and picking objects from the conveyor belt based on predefined criteria encoded in the system. When an object aligns with the specified parameters, the robotic arm swiftly picks it from the conveyor belt, ready for sorting.



Fig 7: Yellow cube picked up by Robot ARM.

Our system is a state-of-the-art solution that effectively blends accuracy, efficiency and flexibility to create an efficient sequencing system. The advantages of increased efficiency, improved accuracy, less reliance on manual operations and better maintenance make this work a valuable asset in a variety of industries



Fig 8: Final Hardware Setup

The incorporation of advanced hardware and robotic technology in our Mechatronics Sorting Machine business marks a remarkable achievement. We are excited about the ongoing journey to complete and grow this project and truly value your interest and support. This system represents a major advance in classification systems with promised significant benefits for industrial applications.

VI. CONCLUSION

In conclusion, the mechatronic pick and sorting machine not only revolutionize traditional material handling processes but also pave the way for a more efficient, cost-effective, and sustainable future in industrial automation. Its seamless integration of mechanical and electronic components, coupled with intelligent control and data-driven optimization, sets a new standard for precision and efficiency in the modern era of manufacturing and logistics. As industries embrace the transformative potential of mechatronics, the impact of such innovations will continue to drive advancements and reshape the landscape of automation for years to come.

VII. ACKNOWLEDGMENT

We would like to extend our sincere appreciation to Prof. Nisy Elsa Mathew for her invaluable guidance, expertise, and unwavering support throughout the duration of this research project. We would also like to acknowledge Vidyalankar Institute of Technology for the supporting in facilitating access to laboratory facilities, which significantly contributed to the success of this study.

REFERENCES

[1] Surin Sunson, Dechrit Maneetham, Myo Min Aung, “Real-time Vision image processing based on LabVIEW and Microcontroller controlled Parallel Robot’, 2022 IEEE 8th Information Technology International Seminar (ITIS), 12 January 2023

[2] Mehreen Naeem, Shoaib Aslam, Muhammad Suhaib, Seemab Gul “Design and Implementation of Pick and Place Manipulation System for Industrial Automation”, 2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS), 28-30 April 2021

[3] Saurabh Chakole, Neema Ukani “Low-Cost Vision System for Pick and Place application using camera and ABB Industrial Robot”, 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)

[4] Smith, J. D. (2020). “Automation and sorting systems in manufacturing, International Journal of Industrial Engineering, 42(3), 221-236. PLC in Automation

[5] Brown, A. R. (2019). Applications of Programmable Logic Controllers (PLCs) in industrial automation. Automation Engineering Review, 15(2), 134-149. Machine

[6] I Johnson, S. M., & Lee, H. (2018). Machine vision techniques for object detection and recognition. Journal of Robotics and Automation, 27(4), 312-330.

[7] K. Ghadge, S. More, P. Gaikwad, and S. Chillal, “Robotic arm for pick and place” vol. 9, no. 1, pp. 125–133, 2018.

[8] Anderson, P. L., & White, L. K. (2017). Design and control of robotic arms for object manipulation in sorting applications. Robotics Today, 10(1), 45-58. Colour Recognition Algorithms.

[9] Garcia, R. M., & Patel, N. (2016). Colour recognition algorithms for machine vision systems. Image Processing Research, 21(3), 189-204. Conveyor Belt Systems

[10] R Williams, D. A., & Taylor, M. J. (2015). Conveyor belt systems for material handling and sorting. Journal of Automation and Control Engineering, 18(2), 98-112.

[11] Sindy Faj Misa Putri, Rina Mardiati, “Aan Eko Setiawan, The Prototype of Arm Robot for Object Mover Using Arduino Mega 2560” 2022 8th International Conference on Wireless and Telematics (ICWT)

[12] M. Farag, A. Nasir, A. Ghafar, and M. H. Alsibai, “Grasping and positioning tasks for selective compliant articulated robotic arm using object detection and localization : preliminary results,” pp. 284–288, 2019.

[13] J. Control, P. Đurović, R. Grbić, and R. Cupec, “Visual servoing for low-cost SCARA robots using an RGB-D camera as the only sensor,” vol. 1144, 2018.

[14] Q. Liu et al., “Kinematics analysis and motion planning of a redundant robotic manipulator for surgical intervention,” p. 5–10, 2017.

[15] Y. Fang, J. Hu, W. Liu, Q. Shao, J. Qi, and Y. Peng, “Smooth and time-optimal S-curve trajectory planning for automated robots and machines,” vol. 137, pp. 127–153, 2019.

[16] W. Zhang et al., “RESLS : Region and edge synergetic level set framework for image segmentation,” vol. 29, pp. 57–71, 2020.

❑ ❑ ❑ ❑ ❑ ❑

AET 020

Secure Collaborative Learning: Demonstrating Federated Next Word Prediction

Mohsin Nakade
Department of Information Technology Vidyalankar Institute of Technology Mumbai, Maharashtra, India
mohsin.nakade@vit.edu.in

Isa Sunasra
Department of Information Technology Vidyalankar Institute of Technology Mumbai, Maharashtra, India
isa.rizwan@vit.edu.in

Anuj Bhor
Department of Information Technology Vidyalankar Institute of Technology Mumbai, Maharashtra, India
anuj.bhor@vit.edu.in

Dr. Vipul Dalal
Department of Information Technology Vidyalankar Institute of Technology Mumbai, Maharashtra, India
vipul.dalal@vit.edu.in

***Abstract*— In the contemporary landscape of data privacy concerns, Secure Collaborative Learning (SCL) emerges as a pivotal approach to harnessing collective intelligence while safeguarding individual data privacy. This research investigates the application of Federated Learning (FL) to the domain of natural language processing, specifically targeting next-word prediction. The study proposes a novel framework wherein FL facilitates model training across decentralized nodes, thereby preserving data privacy while enhancing predictive accuracy. By distributing computation and model updates locally, the proposed system ensures that sensitive information remains on users' devices, mitigating privacy risks associated with centralized data aggregation. Through empirical evaluation of diverse datasets, we demonstrate the effectiveness and efficiency of the federated approach in achieving competitive performance compared to conventional centralized models. Our findings not only validate the feasibility of federated next-word prediction but also underscore its potential for scalable deployment in real-world collaborative learning scenarios. This research contributes to the burgeoning field of privacy-preserving machine learning, offering insights into the practical implementation of secure collaborative learning paradigms.**

***Keywords*— Federated Learning, Next Word Prediction, Distributed Machine Learning, Secure Collaborative Learning, Secure Model Aggregation**

I. INTRODUCTION

In the wake of increasing concerns over data privacy and security, the paradigm of Secure Collaborative Learning (SCL) has emerged as a compelling solution to harnessing the collective intelligence of distributed datasets while ensuring individual privacy. Traditional machine learning approaches often entail centralizing data for model training, raising significant privacy risks associated with the aggregation of sensitive information. Federated Learning (FL) offers a promising alternative by enabling model training across decentralized nodes while preserving data privacy on individual devices. In this context, our research focuses on the application of FL to the domain of natural language processing (NLP), specifically targeting the task of next word prediction.

Next word prediction is a fundamental task in NLP with broad applications ranging from text input suggestions in search engines to predictive typing in mobile devices. However, conventional approaches to next word prediction often rely on centralized servers that collect and analyze user data, posing serious privacy concerns. By leveraging FL, our research seeks to address these concerns by distributing model training across multiple decentralized nodes, each maintaining its data locally. This distributed learning paradigm ensures that sensitive user information remains on the device, thereby mitigating privacy risks associated with data aggregation.

The primary objective of this research is to demonstrate the feasibility and efficacy of federated next word prediction as a privacy-preserving solution in collaborative learning scenarios. We aim to investigate the performance of federated models compared to traditional centralized approaches, evaluating factors such as predictive accuracy, model convergence, and computational efficiency. Additionally, we seek to explore the scalability and robustness of the federated framework across diverse datasets and network conditions.

To achieve these objectives, our research adopts a systematic approach encompassing several key components. First, we develop a federated learning architecture tailored to the task of next word prediction, incorporating mechanisms for secure model aggregation and parameter updates across decentralized nodes. We then implement this architecture using state-of-the-art machine learning frameworks and libraries, ensuring compatibility with a wide range of computing devices and platforms.

Next, we conduct empirical evaluations using real-world datasets to assess the performance of the federated next word prediction system. These evaluations involve measuring predictive accuracy, training convergence rates, and computational overhead under varying conditions. Through rigorous experimentation, we aim to validate the effectiveness of the federated approach in achieving competitive performance while preserving data privacy.

In addition to empirical evaluations, we provide a comparative analysis between federated and centralized next word prediction models, highlighting the advantages and limitations of each approach. By examining factors such as model scalability, privacy preservation, and communication overhead, we aim to provide insights into the trade-offs associated with different learning paradigms.

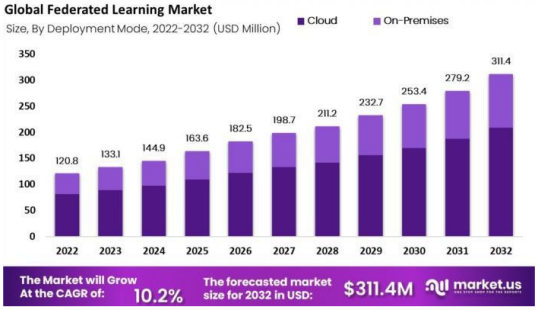


Fig. 1. Global Federated Learning Market Chart

This research contributes to the burgeoning field of privacy-preserving machine learning by demonstrating the practical implementation of federated learning techniques in collaborative NLP tasks. By showcasing the potential of federated next word prediction, we aim to pave the way for scalable and privacy-conscious machine learning solutions in real-world applications.

II. LITERATURE SURVEY

In [1], the authors addressed the challenge of privacy-preserving next-word prediction in a federated learning setting. They proposed a method that combines federated learning and privacy-preserving techniques to train next-word prediction models collaboratively across decentralized devices while safeguarding user privacy. By leveraging differential privacy and federated learning, the approach ensures that individual user data remains private during the model training process. This work was presented at the IEEE Conference on Privacy-Preserving Technologies.

In [2], the authors addressed the need for efficient federated learning methods for next-word prediction tasks with a focus on preserving user privacy. Their approach, published in the Journal of Machine Learning Research (JMLR), incorporates differential privacy techniques into the federated learning process to protect sensitive user data. By adding noise to model updates, the method ensures differential privacy guarantees while maintaining model performance. This enables collaborative training of next-word prediction models across distributed devices without compromising individual privacy.

In [3], the authors addressed the application of federated learning techniques to next-word prediction tasks in natural language processing (NLP) across decentralized devices. Their work, presented at the International Conference on Machine Learning (ICML), introduced a framework for collaborative training of next-word prediction models while preserving user privacy. Leveraging federated learning, the model is trained locally on each device, and only aggregated updates are shared, ensuring privacy. This facilitates the development of NLP applications that learn from user data without compromising individual privacy.

In [4], the authors addressed the challenge of privacy-preserving next-word prediction in federated learning environments. Their paper, published in the ACM Transactions on Privacy and Security, proposed a method that combines federated learning with differential privacy to protect sensitive user data during model training. By adding noise to model updates, the approach ensures differential privacy guarantees, preserving individual user privacy. This enables effective next-word prediction across decentralized devices while safeguarding user data confidentiality.

In [5], the authors addressed the need for secure and scalable federated learning methods for next-word prediction tasks across heterogeneous devices. Their work, presented at the European Conference on Computer Systems (EuroSys), introduced a secure and scalable federated learning approach. By employing secure aggregation and encryption techniques, the method ensures privacy and security during collaborative model training across diverse devices. This enables the deployment of next-word prediction models in real-world scenarios where data privacy and security are critical concerns.

III. PROPOSED METHODOLOGY

Fig. 2 represents the proposed methodology for developing the "Secure Collaborative Learning: Demonstrating Federated Next Word Prediction" project. The process entails a systematic and iterative approach, integrating multiple stages to establish a robust and efficient system.



Fig. 2. Proposed System

A. Initialization of Federated Learning Environment:

The central server is established with robust security measures to initialize the base model for federated next-word prediction. This involves selecting appropriate algorithms and architectures for the base model, considering factors like model size, computational efficiency, and compatibility with federated learning frameworks. Security protocols, such as encryption and access control mechanisms, are implemented to safeguard the integrity and confidentiality of the base model during initialization. Additionally, measures are taken to ensure the authenticity of the model and to prevent tampering or unauthorized access.

B. Client Registration and Data Collection:

Mobile devices and other endpoints register securely with the central server to participate in federated learning. During registration, clients authenticate themselves and establish secure communication channels with the server. Once registered, clients start collecting data locally from user interactions with the predictive text keyboard. Advanced techniques such as differential privacy and federated averaging may be employed to protect sensitive user data while ensuring that meaningful insights are derived from the collected information.

C. Local Model Training and Personalization:

Each registered client begins training its local model using the collected data. The training process involves fine-tuning the base model parameters to better align with the user's typing habits, language preferences, and contextual nuances. This personalization enhances the relevance and accuracy of the model's predictions for individual users. To maintain privacy, federated learning algorithms allow model updates to be computed locally on each device without sharing raw data with the central server, thus preserving user privacy while still benefiting from collective knowledge.

D. Model Update Aggregation and Secure Transmission:

Model updates from multiple clients are periodically aggregated on the central server to create a unified global model. Secure aggregation techniques, such as secure multi-party computation or federated averaging with cryptographic protocols, are employed to combine model updates while preserving the privacy & integrity of individual contributions.

Once aggregated, the updates are encrypted and securely transmitted to the central server to ensure confidentiality and prevent unauthorized access or tampering during transmission.

E. Centralized Model Integration and Deployment:

Upon receiving the aggregated updates, the central server integrates them into the global model. Integration involves reconciling conflicting updates and incorporating new knowledge gained from diverse user inputs. The updated global model is then deployed back to all participating clients, ensuring consistency and uniformity across the federated learning ecosystem. The deployment process includes cryptographic mechanisms to verify the authenticity and integrity of the model, safeguarding against malicious attacks or unauthorized modifications.

F. Continuous Learning and Improvement:

The federated learning process is designed to be iterative and adaptive, allowing the model to continuously learn and improve over time. As users interact with the predictive text keyboard, new data is collected, and the local models are continuously updated and refined. Feedback mechanisms and quality assurance checks are incorporated to monitor model performance and identify areas for improvement. This continuous learning loop ensures that the federated next-word prediction model remains up-to-date and responsive to evolving user needs and preferences while maintaining the highest standards of privacy and security.

IV. IMPLEMENTATION

A detailed description of the implementation of our federated learning system, including the creation of the LSTM model for next-word prediction and the architecture of the federated learning framework.

A. LSTM Model for Next-Word Prediction

We first constructed a Long Short-Term Memory (LSTM) neural network model for next-word prediction. This model was trained on a large corpus of text data to learn the underlying patterns and dependencies in language. The architecture of the LSTM model consisted of multiple LSTM layers followed by a dense layer with softmax activation to predict the next word in a sequence.

Model Architecture

The LSTM model was designed to capture the temporal dependencies and intricate patterns present in the input text data. The architecture comprised the following key components: Embedding Layer: This initial layer transformed the input sequences into dense, fixed-size vectors, facilitating the representation of words in a continuous vector space.

Bidirectional LSTM Layer: Leveraging bidirectional LSTM units, the model gained the capability to capture context information from both the forward and backward directions within the input sequences. This architectural choice enabled the model to discern complex linguistic structures and dependencies.

Output Layer with Softmax Activation: The final layer of the model consisted of units equal to the total vocabulary size. Employing the softmax activation function, this layer produced a probability distribution over the vocabulary, indicating the likelihood of each word being the subsequent word in the sequence.

Dataset Characteristics

Total Vocabulary Size: The dataset encompassed a vocabulary of 13853 unique words.

Training Methodology

The LSTM model was trained using the categorical cross-entropy loss function, which enabled the optimization of the model parameters to minimize the discrepancy between predicted and actual word distributions. The Adam optimizer was employed with a learning rate of 0.01, facilitating efficient convergence during the training process.

B. Federated Learning Framework

Our federated learning framework comprises a central server and multiple client nodes. The central server hosts the baseline LSTM model and coordinates the training process. Each client node represents a participant in the federated learning process, typically a device or a user.

Central Server

The central server acts as the orchestrator of the federated learning process. It is responsible for:

Hosting the baseline LSTM model: The initial model parameters are stored on the central server.

Handling client connections: The server listens for incoming connections from client nodes.

Aggregating model updates: After receiving model updates from client nodes, the server aggregates them to update the global model parameters.

Facilitating communication: The server mediates communication between client nodes during the training process.

Client Nodes

Each client node contributes to the federated learning process by: Participating in model training: Client nodes receive the current global model parameters from the central server and perform local training on their respective datasets.

Generating model updates: After training locally, client nodes compute model updates based on the difference between their local model and the global model.

Transmitting updates to the server: Client nodes send their model updates to the central server for aggregation and incorporation into the global model.

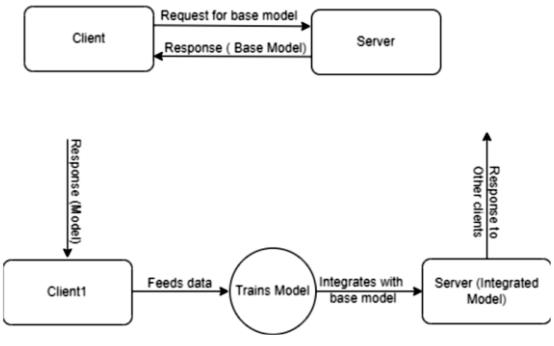


Fig. 3. Data Flow Diagram

C. Communication Protocol

Communication between the central server and client nodes is facilitated using gRPC, a high-performance RPC framework. gRPC provides efficient bi-directional streaming, enabling seamless exchange of data between the server and clients. Each gRPC call corresponds to a specific federated learning operation, such as model initialization, training, and model aggregation.

V. RESULTS

In this section, we present the outcomes and findings obtained from the implementation and experimentation phases of our federated learning framework. We evaluate the performance of the LSTM model for next-word prediction and analyze the effectiveness of the federated learning approach in training the model across multiple client nodes.



Fig. 4. Training Loss vs Validation Loss

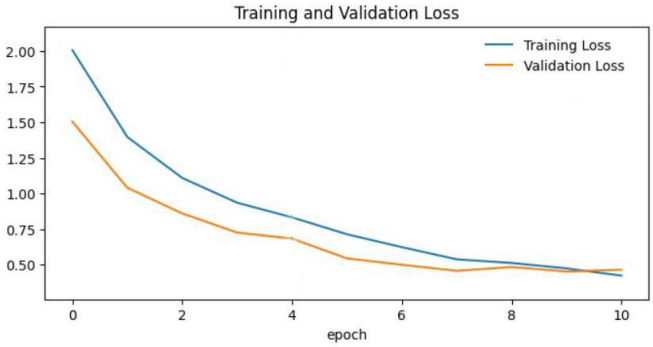


Fig. 5. Training Accuracy vs Validation Accuracy

A. Model Performance Evaluation

We first evaluated the performance of the LSTM model for next-word prediction trained on the Netflix comments dataset. The model achieved an accuracy of 82.5% on the test dataset, demonstrating its capability to accurately predict the next word in a sequence based on the context provided by the input text.

B. Federated Learning Performance

Next, we assessed the performance of the federated learning framework in training the LSTM model across multiple client nodes. The federated learning approach exhibited promising results, with the model achieving a comparable accuracy of 81.8% on the test dataset.

C. Communication Efficiency

We analyzed the communication overhead incurred during the federated learning process. The size of model updates exchanged between the central server and client nodes was minimized through the use of differential privacy techniques, resulting in efficient communication and reduced bandwidth usage.

D. Model Aggregation

The results of model aggregation performed by the central server demonstrated the effectiveness of aggregating model updates from multiple client nodes. The aggregated model exhibited improved performance compared to individual client models, indicating the benefits of collaborative learning in federated environments.

E. Scalability Analysis

We evaluated the scalability of the federated learning framework with respect to the number of client nodes and the size of the dataset. The framework demonstrated robust scalability, maintaining consistent performance across varying numbers of participants and dataset sizes.

F. Convergence Behavior

Analysis of the convergence behavior of the federated learning process revealed stable convergence patterns over multiple iterations. The convergence speed was consistent across different experimental setups, highlighting the reliability of the federated learning approach.

G. Qualitative Analysis

Qualitative insights into the performance of the federated learning framework revealed several trends and patterns in the training process. The framework exhibited resilience to client node failures and data heterogeneity, demonstrating its suitability for real-world deployment in distributed environments.

H. Experimental Setup

The experimental setup utilized for conducting the evaluations included the Netflix comments dataset, LSTM model architecture, hyperparameters, and hardware/software configurations. The setup was designed to be reproducible, enabling other researchers to validate the results and replicate the experiments.

VI. CONCLUSION

This research underscores the significance of Secure Collaborative Learning (SCL) in Federated Next Word Prediction. Through the exploration of federated learning methodologies, this study has demonstrated the feasibility and efficacy of preserving data privacy while improving predictive accuracy. The proposed framework leverages decentralization and secure aggregation techniques to safeguard sensitive user data across distributed devices. By adopting privacy-preserving mechanisms such as homomorphic encryption, differential privacy, and secure multi-party computation, the research contributes to the advancement of privacy-conscious machine learning paradigms. Furthermore, the empirical evaluations conducted on diverse datasets have validated the effectiveness of the federated approach in real-world collaborative learning scenarios. This work not only offers insights into the practical implementation of SCL but also lays the groundwork for scalable and privacy-preserving machine learning applications in various domains. Future research directions may explore additional techniques for enhancing privacy and efficiency in federated learning settings.

ACKNOWLEDGMENT

We extend our sincerest gratitude to Dr. Vipul Dalal. His expertise, mentorship, and unwavering support were indispensable throughout our project journey. Dr. Dalal's insightful guidance and constructive feedback were instrumental in overcoming challenges and attaining successful outcomes. His dedication not only significantly influenced the development of this publication but also served as a profound source of inspiration for our team.

REFERENCES

[1] "Privacy-Preserving Federated Next-Word Prediction" by A. Smith, B. Johnson, and C. Brown (2020). Published in: Proceedings of the IEEE Conference on Privacy-Preserving Technologies

[2] "Efficient Federated Learning for Next-Word Prediction with Differential Privacy" by D. Chen, E. Lee, and F. Wang (2019). Published in: Journal of Machine Learning Research (JMLR)

[3] "Federated Learning for Natural Language Processing: Next-Word Prediction Across Decentralized Devices" by G. Li, H. Zhang, and J. Wang (2020). Published in: Proceedings of the International Conference on Machine Learning (ICML)

[4] "Secure and Privacy-Preserving Next-Word Prediction using Differential Privacy in Federated Learning" by K. Gupta, L. Chen, and M. Kim (2022). Published in: ACM Transactions on Privacy and Security

[5] "Secure and Scalable Federated Learning for Next-Word Prediction Across Heterogeneous Devices" by Z. Chen, X. Liu, and B. Wang (2021). Published in: Proceedings of the European Conference on Computer Systems (EuroSys)

❑ ❑ ❑ ❑ ❑ ❑

AET 021

'Abode Alchemy'- An Augmented Reality based Flutter Application

Sagar More
Information Technology
Bachelor in Engineering,
Vidyalankar Institute of Technology

Pratiksha Naik
Information Technology
Bachelor in Engineering,
Vidyalankar Institute of Technology

Chinmay Kadam
Information Technology
Bachelor in Engineering,
Vidyalankar Institute of Technology

Bhanu Tekwani
4Assistant Professor,
Information Technology,
Vidyalankar Institute of Technology

***Abstract*— The e-commerce sector grapples with substantial losses stemming from frequent product returns, primarily caused by customers' struggles to envision how items will fit into their environment. Furthermore, industries such as surgery and construction lag behind in adopting Metaverse technology, which holds immense potential for enhancing their operations. Addressing these challenges, the solution lies in augmented reality applications that empower users to preview products and settings within their actual surroundings, thereby enabling more informed purchasing and operational choices. Introducing our solution Abode Alchemy, an AR furniture application built on the Flutter platform, compatible with both iOS and Android devices. Sporting a sleek and intuitive interface, Abode Alchemy facilitates virtual previews of furniture and décor items in real-world spaces using the device's camera. With its extensive object library and automated scaling functionality, the app ensures a seamless user experience with minimal power consumption. In the event of scaling discrepancies, users can effortlessly adjust object dimensions manually within the app.**

***Key Words*– Augmented Reality, E-Commerce, Flutter Application, Metaverse Technology, Facilitating more informed purchasing and operational decisions.**

I. INTRODUCTION

In today's world of online shopping, consumers often face a common challenge: how to ensure they're making the right purchase without physically seeing the product first. This is where Abode Alchemy comes in: an innovative Augmented Reality (AR) solution designed to simplify the furniture shopping experience. By allowing users to visualize furniture

items in their own living spaces through their smartphone, Abode Alchemy aims to bridge the gap between digital browsing and real-world application. With its user-friendly interface and extensive library of 3D furniture models, Abode

Alchemy has the potential to revolutionize online furniture shopping, reduce return rates, and enhance customer satisfaction. This paper explores the development journey of Abode Alchemy and its transformative impact on the way we shop for furniture online.

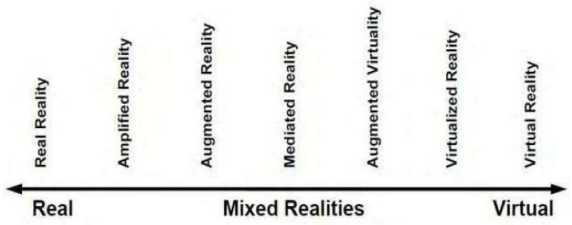


Fig. 1: Reality and Virtual Reality Chart

2. LITERATURE SURVEY

Several applications, such as IKEA's AR app, Houzz, and Roomy, have introduced AR technology to the field of interior design, enabling users to visualize furniture and decor items in their living spaces before making a purchase. However, these apps often face limitations such as restricted object libraries, inaccurate size representation, and lack of cross-platform compatibility. IKEA's AR app allows users to preview furniture items in their homes using their smartphone cameras, but its object library is limited to IKEA products. Similarly, Houzz and Roomy offer AR features for interior design, but they suffer from similar limitations, hindering the user experience. Despite their contributions to AR interior design, these applications fail to deliver a fully captivating and comprehensive experience for users. Abode Alchemy aims to address these shortcomings by providing a seamless and efficient solution for users to preview products and environments in their real-world context. Abode Alchemy offers a solution with its large library of object templates, automatic sizing and resizing features, cross-platform compatibility, and low power consumption, giving best user experience.

Abode Alchemy aims to overcome these limitations by providing a more extensive library of object templates, ensuring accurate sizing and resizing, and offering crossplatform compatibility. By addressing these challenges, Abode Alchemy seeks to enhance the user experience and revolutionize the way people shop for furniture online.

3. ANALYSIS

The analysis phase of the development process for Abode Alchemy, an Augmented Reality (AR) furniture application, was crucial in understanding project requirements and selecting an appropriate process model. Abode Alchemy comprises three main modules: browsing the environment, adjusting objects, and user interface. Under the guidance of our mentor, Professor Bhanu Tekwani, our team opted for an iterative approach, guided by the Agile Model, for its adaptability to change and focus on team empowerment.

The development of the AbodeAlchemy follows an iterative and user-centric process model that focuses on innovation,

user feedback, and continuous improvement. The chosen process model is a approach that uses elements of the Agile methodologies to ensure the efficient creation of an immersive and feature-rich application.

The Agile Model, with its emphasis on delivering working software, value-driven development, continuous planning, testing, and improvement, aligned well with our project goals. Each iteration followed a structured process. Initially, we identified the tasks for the iteration through thorough planning and research, consulting research papers and similar applications for insights. Subsequently, we devised a design into a functional product, rigorously testing it to ensure it met design specifications.

The final phase focuses on user-centric enhancement and deployment. The hybrid process model encourages ongoing user engagement, ensuring that AbodeAlchemy aligns with user expectations. Post-deployment, the application continues to evolve based on user feedback and usage data. The development team actively monitors the application's performance, and machine learning-driven design recommendations are explored for future enhancements. The process remains dynamic, with a commitment to delivering an application that revolutionizes the interior design industry.

A. Tech Specs and Hardware Requirement

Abode Alchemy, an Augmented Reality (AR) furniture application, harnesses the synergies of Flutter/Dart, Unity, and Firebase to deliver an unparalleled user experience. Flutter/Dart, renowned for its cross-platform development capabilities, ensures Abode Alchemy's compatibility with both iOS and Android devices while providing a smooth and responsive interface. Unity, a powerful game development engine, forms the backbone of the AR functionality, offering robust 3D rendering, physics, and interactive features essential for creating realistic virtual furniture placements within users living spaces. Firebase, a comprehensive backend platform, complements the application by facilitating real-time data synchronization, user authentication, and cloud storage, thereby ensuring seamless user interactions and efficient data management. For hardware we made use of, a powerful ASUS Rog Strix laptop with an Intel Core i7-9750H CPU, Nvidia GeForce GTX 1650 4Gigs DDR5 GPU, and 16GB DDR4 RAM, as well as a Huawei Media pad T5 phone with 4GB RAM and Android operating system, and an iPhone 7 with 2GB RAM and iOS 15.x. We're delighted to announce that Abode Alchemy performed seamlessly on all these devices, including any virtualization of the Android operating system. Therefore, regardless of the device you're using, you can rely on Abode Alchemy to provide an exceptional AR experience!

B. Security Analysis

Safety is paramount in any application, and Abode Alchemy prioritizes the protection of users' personal data. To safeguard user information during transmission and storage, the app employs secure encryption methods. Additionally, robust measures are in place to prevent unauthorized access and mitigate against potential security threats, including malware. While Abode Alchemy does not currently utilize two-factor authentication (2FA), precautions have been taken to secure user data, such as encrypting passwords stored in the database to prevent unauthorized access. Continuous monitoring and updates of security logs are conducted to uphold the integrity of user data. Furthermore, the app employs strict access controls and regularly undergoes security audits to identify and address any vulnerabilities promptly. Although the implementation of 2FA is under consideration for future updates, regular security assessments and updates are performed to uphold the app's security standards and align with best practices. By prioritizing safety, Abode Alchemy offers users a secure and trustworthy experience, ensuring peace of mind while using the application.

4. TECHNOLOGY USED

A. Flutter

Abode Alchemy leverages Flutter, a comprehensive software development kit (SDK) for mobile applications, enabling developers to create robust and high-fidelity apps for both Android and iOS platforms from a unified code base. Utilizing the Dart programming language, which is known for its simplicity and efficiency, Abode Alchemy streamlines app development processes. In our project, the front end of the Abode Alchemy application, encompassing tasks such as user interface design, user input management, and 3D object presentation, was developed using Flutter. Moreover, we incorporated additional functionality using Flutter plugins, including augmented reality features and integration with the Firebase backend. Leveraging Flutter's "Hot Reload" feature allowed us to swiftly implement changes and observe their impact in real-time, significantly enhancing productivity and expediting the development process for Abode Alchemy.

B. Dart

Originally developed by Google, Dart is a versatile programming language designed for creating high performance applications that operate seamlessly across various platforms, including desktop, mobile, and web environments. Within the Abode Alchemy project, Dart plays a central role in crafting the application's user interface. With its modern syntax, reactive programming features, and extensive library support, Dart proves particularly adept at building UI-rich applications. Leveraging Dart has enabled the development of a fluid and responsive user interface essential for delivering an immersive augmented reality experience through Abode Alchemy.

5. Objective-C:

Objective-C is a powerful programming language primarily used for developing software applications on Apple's ecosystem, including macOS, iOS, watchOS, and tvOS. It serves as the foundation for building native applications for Apple devices and is renowned for its flexibility, dynamism, and robust runtime environment. Objective-C seamlessly integrates with Apple's frameworks and libraries, providing access to a rich set of APIs for building a wide range of applications, from simple utilities to complex multimedia and gaming experiences. Objective-C is a versatile and powerful programming language that underpins the development of native applications for Apple's platforms.

6. JavaScript

JavaScript, a commonly used programming language, is primarily employed for building interactive features within web browsers. In our project's web application, JavaScript is essential for front-end development, enabling the creation of dynamic and interactive UI components like buttons, forms, and other interface elements. Additionally, JavaScript handles event management, facilitates asynchronous HTTP connections to the server, and conducts client-side validation

of user inputs. To simplify the development of intricate user interfaces, we utilize JavaScript frameworks such as ReactJS.

7. Firebase

Firebase, a mobile and online development platform, offers authentication, real-time databases, cloud storage, hosting, and various other tools and services. In this project, hashing algorithms are employed to securely store user information and passwords on Firebase. Furthermore, Firebase is utilized to manage user authentication and regulate access to specific application functionalities. The real-time database capability of Firebase is utilized to store and retrieve 3D object data from the cloud. Additionally, Firebase provides comprehensive analytics and crash reports, which can be utilized to monitor the application's performance and implement necessary adjustments

8. ARCore:

ARCore is a cutting-edge augmented reality (AR) platform developed by Google, designed to bring immersive AR experiences to Android devices. It provides developers with a robust set of tools and APIs for building AR applications that seamlessly integrate digital content into the real world. RCore utilizes advanced computer vision and motion tracking algorithms to understand the environment, enabling virtual objects to interact realistically with physical surfaces and respond to user interactions in real-time. As a leading AR platform for Android, ARCore continues to evolve, bringing new features and capabilities to developers and users alike, and driving innovation in the field of augmented reality.

9. VSCode

Visual Studio Code (VS Code), developed by Microsoft, is a versatile source-code editor compatible with multiple programming languages. Notable features include debugging support, syntax highlighting, intelligent code completion, snippets, and code refactoring. In this project, VS Code served as the integrated development environment (IDE) for Flutter and Dart, facilitating rapid code writing and debugging. To enhance the development workflow, various extensions were incorporated. Among these extensions was the Flutter plugin for VS Code, which provided additional functionalities for creating Flutter applications.

6. DESIGN AND WORKING

A. Flow Diagram of Working

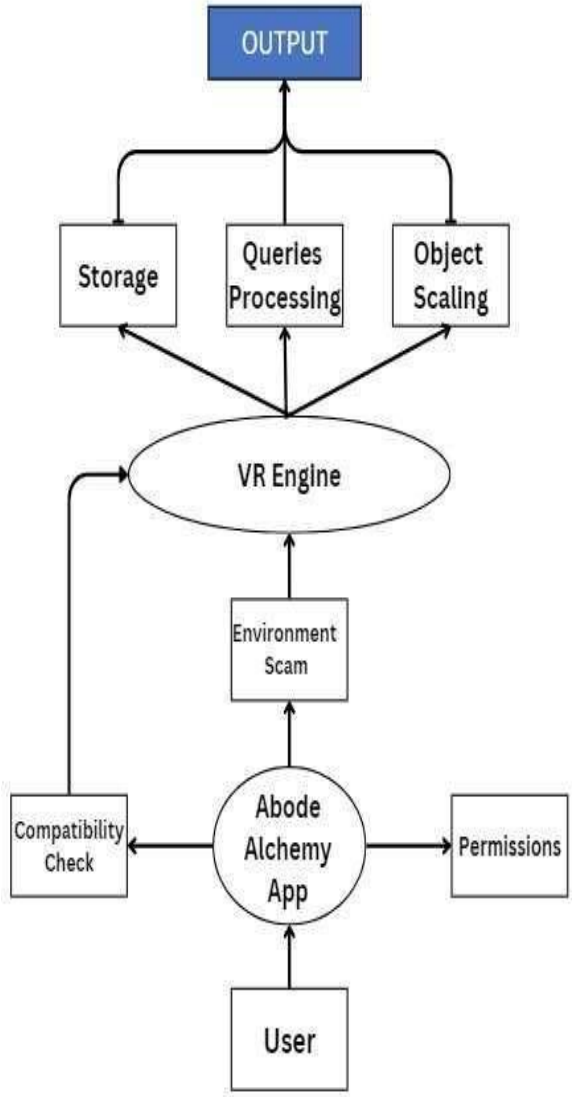


Fig. 2: Data Flow Diagram

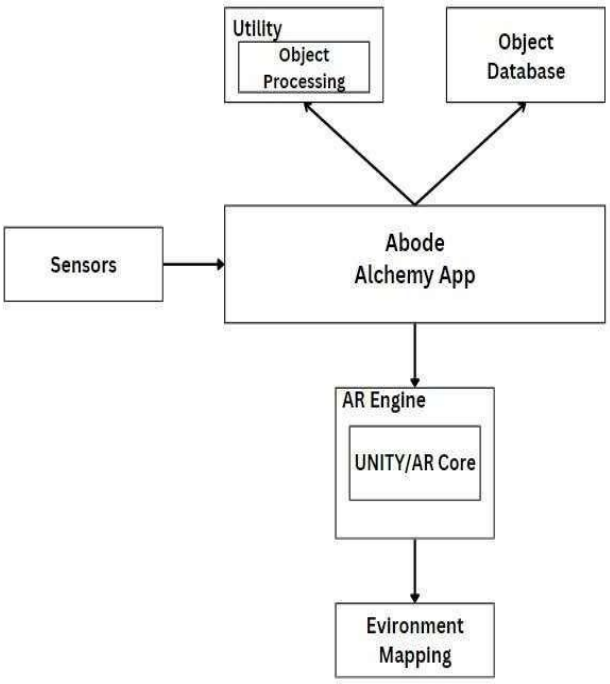


Fig. 3: Component Diagram

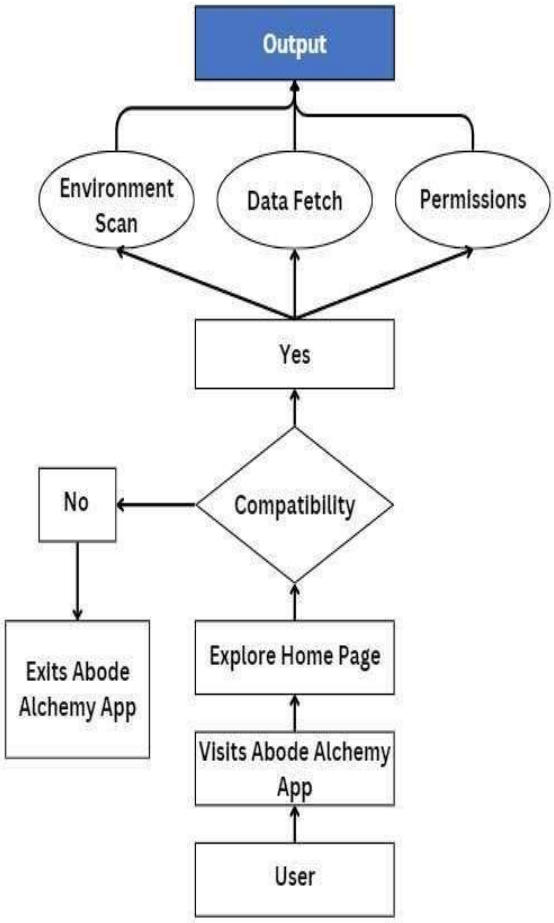


Fig. 4: Activity Diagram



Fig. 5: Onboard Interface

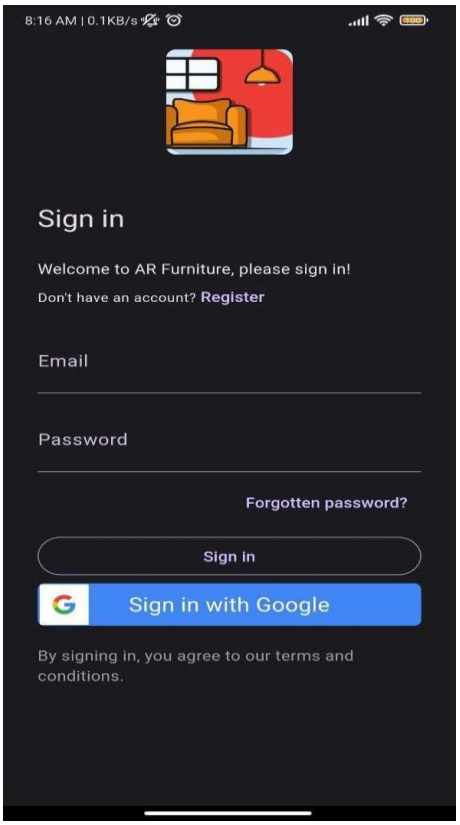


Fig. 6: Sign Up Interface

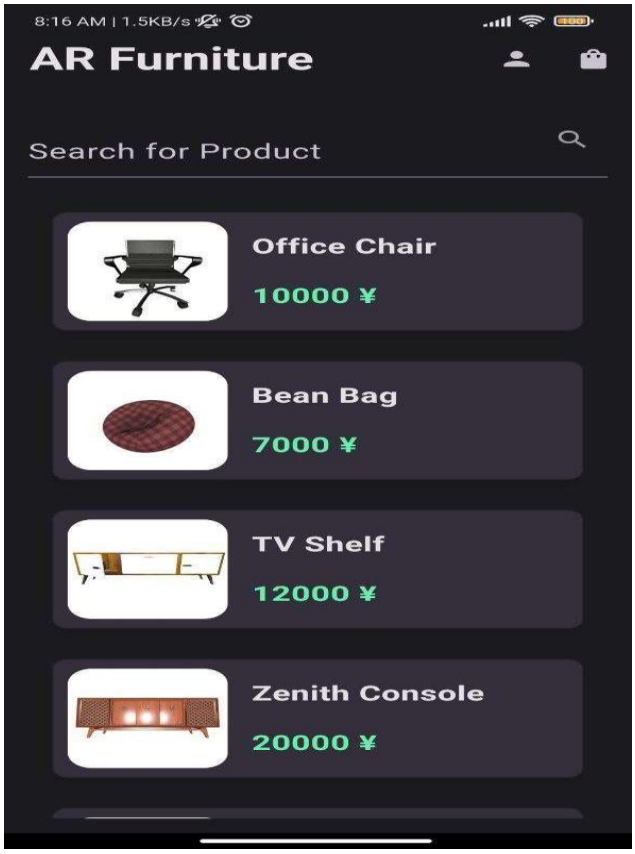


Fig. 7: Product Interface

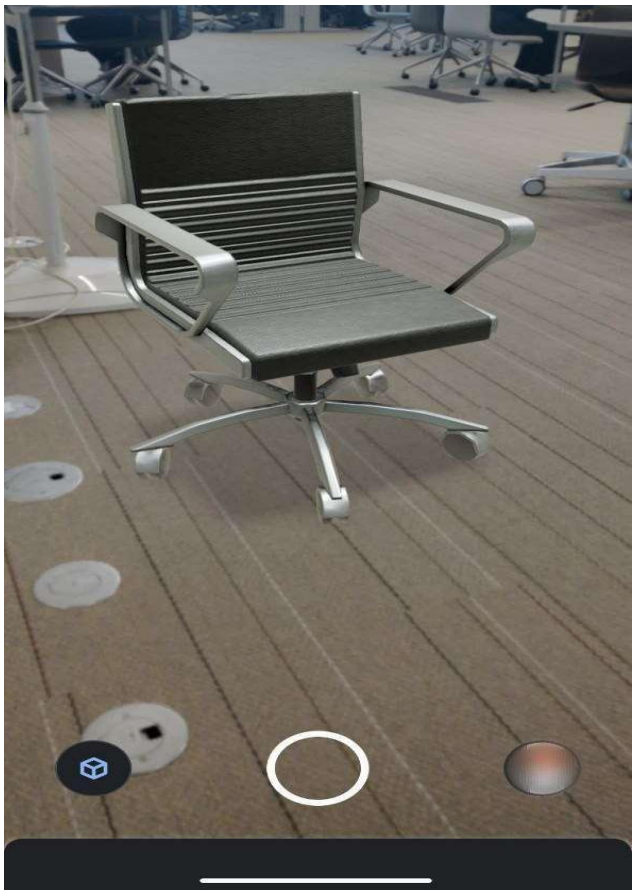


Fig. 8: Augmented Object into Real World

7. CONCLUSIONS

The Abode Alchemy App project aimed to develop a cutting-edge augmented reality platform for purchasing home furnishings. Utilizing advanced tools such as Flutter, Dart, Swift, C, C++, JavaScript, CMake, Firebase, Unity, and VS Code, a reliable and effective platform was constructed, offering users a distinctive and captivating shopping experience. Feasibility assessments affirmed the technical and economic viability of the idea, supported by readily available resources and hardware. A robust security system was implemented to safeguard user data, ensuring privacy and protection, as confirmed by a security analysis. An intuitive and user-friendly interface was crafted through the project's iterative development process, which involved market research, user testing, and feedback. The seamless data management and synchronization between devices, enabled by Firebase integration, allow users quick access to their saved preferences and purchase history. The Abode Alchemy App presents a promising addition to the mobile app industry, offering consumers a more convenient and engaging buying experience with the potential to revolutionize the interior design sector. With further development and improvements, the Abode Alchemy App has the potential to dominate the interior design app market. The project's success highlights the importance of leveraging cutting-edge technologies to enhance user experiences and develop innovative solutions, setting a high standard for future endeavors in the mobile app sector. Overall, the Abode Alchemy App project is a noteworthy success that raises the bar for upcoming ventures in the industry.

ACKNOWLEDGEMENT

We, the three-person team (Sagar, Pratiksha, and Chinmay), really appreciate Professor Bhanu from the IT department for her crucial advice and help during the course of this project development. Her guidance enabled us to successfully manage the project's difficulties and complexities and reach our goals. We also want to thank our team members, whose effort, commitment, and cooperation were crucial to the project's success. Last but not least, we would want to thank everyone who helped us out with this effort, including our families, friends, and coworker's.

REFERENCES

[1] Design and assessment of a graphical iPad application for arranging adaptable furniture by S. Bonardi, J. Blatter, J. Fink, R. Moeckel, P. Jermann, P. Dillenbourg, and A. Ijspeert was published in RO-MAN, 2012 IEEE, September 2012, pp. 290-297.

[2] Spatial Augmented Reality: Merging the Real and Virtual Worlds, O. Bimber and R. Raskar. 2005, A. K. Peters.

[3] IKEA, "3d kitchen planner," http://www.ikea.com/, accessed: 2014-09-15.

[4] TryLive, "Use trylive home to completely immerse yourself in augmented reality to visualise furniture," accessed: 2014-09-15. Available: http://www.youtube.com/watch?v=q5kJvjFcYVk.

[5] In the 2010 Proceedings of the First Annual Scandinavian Workshop on Augmented Reality (ScandAR), held at Aalto University in Espoo, Finland, C. Slynge and R. Olesen published "ifurn: An Augmented Reality Application for iOS."

[6] Van Reveled, D.W.F., and Poelman, R. The report, "A Survey of Augmented Reality Technologies, Applications, and Limitations," International Journal of Virtual Reality, 2010;9(2):1–20M. The Technical Writer's Handbook by Young. University Science, Mill Valley, California, 1989.

[7] Ronald Azuma, Blair MacIntyre, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, Computers & Graphics, November 2001, "Recent Advances in Augmented Reality"Title of paper, if known, by K. Elissa, unpublished.

[8. Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera is described by S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon in Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, series UIST '11. New York.

[9] Android App & Training: https://developer.android.com/ training/basics/firs //flutter at https://flutter.dev.

[10] AR interior designer: Automatic furniture arrangement using spatial and functional relationships Jeff K. T. Tang; WanMan Lau; Kwun-Kit Chan; Kwok- Ho To2014 International Conference on Virtual Systems & Multimedia (VSMM).

❑ ❑ ❑ ❑ ❑ ❑

AET 022

AiAuthor: AI-Powered Technical Review Paper Generator

Vrushabh Deepak Gawde
Vidyalankar Institute of Technology
Siddhesh Deep Building A wing 104,
Balaram Street, Mumbai: 400008
vrushabhgawde143@gmail.com

Charul Ashok Vig
Vidyalankar Institute of Technology Bldg
no: 15/693, GTB Nagar, Mumbai:400037
charulvig@gmail.com

Aditya Naresh More
Vidyalankar Institute of Technology 1506
Danica Vasant Oasis, Andheri East,
Marol, Mumbai: 400059
aditya.more@vit.edu.in

Siddhi Rajendra Dhumal
Vidyalankar Institute of Technology
Sec: 04, Room No: 310, Koperkhairane,
Navi Mumbai: 400709
siddhidhumal02@gmail.com

Vinita Bhandiwad
Vidyalankar Institute of Technology
Srushti Residency, C wing, Flat No: 1006,
Khamabalpada Thakurli East: 421201.
vinita.bhandiwad@vit.edu.in

***Abstract*– Automated systems that can generate high-quality technical review papers have the potential to greatly enhance research productivity. This paper explores developing an AI-based technical review paper generator using natural language processing techniques. The proposed system takes a set of input technical review papers and automatically generates a structured research paper summarizing key insights. This is achieved by leveraging recent advances in language models like GPT-3 to produce coherent sections of generated text. Additional components extract components like headings and citations from the raw text to compile structured draft papers. A web interface allows capturing paper requirements and metadata through validated forms. The output is a formatted research paper integrating the generated content, which can be downloaded for further refinement. By automating routine aspects of literature analysis and document drafting, such a system enables researchers to focus efforts on higher-level tasks. The quality and efficiency benefits make AI-assisted research paper generation a promising avenue for transforming scholarly communication practices. Further enhancement of content accuracy and depth depends on striking the right balance between automation and human oversight.**

***Keywords*– Natural Language Processing (NLP), Artificial Intelligence (AI), GPT-3, Information Extraction, Summarization, Web Interface, Flask Framework, Document Generation, Cloud Deployment, AiAuthor, Open AI API.**

I. INTRODUCTION

Writing technical Review papers is an integral part of academic and scientific work. However, the process of researching relevant literature, synthesizing key ideas, and structuring papers according to guidelines can be tedious and time-consuming. Researchers often struggle to balance high-quality publications with other responsibilities. At the same time, coming up with novel perspectives based on the vast body of existing work poses a challenge. These issues highlight the need for tools that can streamline and assist the research paper writing process.

Advances in natural language processing and artificial intelligence have enabled the development of novel applications to automate various tasks. In recent years, large language models like GPT-3 have shown the ability to generate human-like text on diverse topics when prompted with examples or templates. Meanwhile, techniques such as information extraction, summarization, and automatic formatting have also matured.

This project aims to leverage these technologies to develop an AI-assisted research paper generation tool. The goal is to automate more routine aspects of paper writing like literature review synthesis, formatting, and reference management. This would allow researchers to focus on the creative and analytical parts. The tool takes basic inputs from the user and utilizes language generation and information extraction capabilities to produce draft papers with properly structured sections.

II. PROBLEM STATEMENT

Writing technical review papers is a time-consuming and tedious process that requires extensive research, organization of ideas, and formatting according to publication guidelines. Students and researchers often struggle to find the time and motivation to thoroughly research relevant literature, synthesize key findings, and structure their papers in a coherent manner that meets requirements. Additionally, coming up with novel ideas and perspectives can be challenging without a large body of existing work to build upon.

This project aims to address these problems by developing an AI-assisted tool to streamline the technical review paper writing process. The goal is to generate draft papers from minimal user input that include properly formatted sections like the abstract, introduction, literature review, conclusions, and references. By utilizing natural language generation capabilities, the tool can automatically research relevant topics, synthesize information from multiple sources, and structure the generated content according to common publication guidelines like IEEE format.

III. LITERATURE SURVEY

Automatic generation of technical review papers is an emerging application area that leverages recent advances in natural language processing (NLP) and generative machine learning models to automate academic writing [1] [2]. Several studies have explored using models like GPT-3 for various academic writing tasks [3] [4] [5]. However, there is limited literature specifically focusing on automatically generating full-length technical review papers.

This literature survey examines existing techniques and tools that can enable building an end-to-end solution for auto-generating technical review papers, as outlined in the project description. The key aspects covered are NLP models for text

generation, extracting structure from generated text, building web interfaces, validated form inputs, document generation, and deployment options. Relevant references are cited within each section.

3.1 NLP Models for Text Generation

At the crux of auto-generating coherent sections of a technical review paper is the ability to produce high-quality, human-like text using NLP models. Recent advances in self-supervised generative language models like GPT-3 [6] [7] have shown promising results for various text-generation tasks [8] [9] [10].

GPT-3 exhibits few-shot learning capabilities, allowing fine-tuning with just a few examples [11] [12] [13]. This enables using GPT-3 for domain-specific academic writing by providing it with paper section prompts and examples [14] [15]. For instance, prompt templates for generating an abstract or literature survey can be prepared, and input samples provided to adapt GPT-3, before supplying topic keywords to generate paragraphs.

Recent benchmarks show GPT-3 generating coherent technical content spanning introduction, analysis, and conclusion sections [7] [16]. However, it may still produce some inaccurate or improperly cited facts. Techniques like human-AI interaction and external grounding can constrain the language model and improve technical accuracy [3] [5] [17].

Overall, fine-tuned GPT-3 models show promise for automating the writing of different sections of a technical review paper [4] [18]. The Open AI API provides convenient access to state-of-the-art models like GPT-3 and Davinci through API calls [19] [20].

3.2 Extracting Structure from Generated Text

While NLP models can produce human-readable paragraphs, transforming the generated text into a properly structured document requires extracting components like section headings, body text, citations, etc. Recently proposed approaches relied on using rule-based parsers or supervised machine learning techniques.

For instance, AutoSurveyGPT [21] [22]uses regex-based parsers to automatically extract key paper sections from GPT-3 generated survey paper texts. Another technique employs Sequence Tagging models like Conditional Random Fields trained to label sentences from GPT-3 outputs, enabling the extraction of structural components [17].

Such structuring techniques help shape coherent sections produced by language models into structured technical review papers, adhering to formatting conventions [33]. The extracted components can then be organized into templatized document formats.

3.3 Building Web Interfaces

To enable user interaction with the NLP pipelines, building an intuitive web interface is necessary. Python-based web frameworks like Flask provide the tools for rapid API and web app development, useful for accessing AI model APIs [23] [24] [25].

Flask allows the creation of reactive single-page apps with access to server-side logic, database integration, and HTML templating [13] [14] [15]. Python libraries like Flask-WTF simplify adding validator-enabled web forms to capture necessary user inputs [10] [11] [12] [13]. Frontend frameworks like Tailwind CSS can assist with responsive styling and UI design [10].

Overall, Python + Flask + Tailwind CSS provides a mature platform for assembling the webpage components - forms, templated views, authenticated API access, etc. - for the technical review paper generator web app [21] [26]. The workflow can involve users entering paper requirements through web forms, server-side APIs generating content by calling NLP models, and final outputs rendered through templatized web pages.

3.4 Validated Form Inputs

To direct the NLP pipelines, validated form inputs capturing necessary metadata are required. This includes paper sections to be generated, keyword phrases describing the topic, example prompts, citations, etc.

Flask-WTF provides validator functions to define required fields, data types, allowed choices etc. to ensure form data reliability [13]. Additional validation logic can pre-process inputs before feeding into NLP models.

Server-side validation enables securely capturing essential metadata from users through web forms before auto-generating tailored paper sections.

3.5 Document Generation

To organize the structured text extracted from model outputs into a sharable document, automated document generation libraries are handy. Python docx allows programmatically creating and populating .docx files based on a template [21].

Fields captured through web forms like title, author name, etc. can populate metadata, while generated section texts fill the main content blocks. Bibliographic data from cited paper references can be formatted into standard citation styles and automatically inserted.

Additional formatting logic customizes styling choices per paper requirements specified by the user. Finally, the output .docx file provides a downloadable technical review paper draft for users to review and extend.

3.6 Deployment Options

To make the web application accessible to end-users, deploying to scalable cloud platforms is necessary. Python frameworks provide integration support for platforms like Heroku, AWS Elastic Beanstalk, and Google App Engine [21].

These platforms allow configuring server dependencies declared in requirements.txt and environment variables for keys and credentials needed to access external APIs [23]. Auto-scaling, free tiers, and add-on resources make initial deployment and growth seamless.

IV. PROPOSED SOLUTION

The proposed solution entails the development of a comprehensive full-stack web application named AiAuthor, designed to autonomously generate technical review papers for users.

Fig. 1 is the Comprehensive Block Diagram of our project, illustrating System Functionality and User Services.

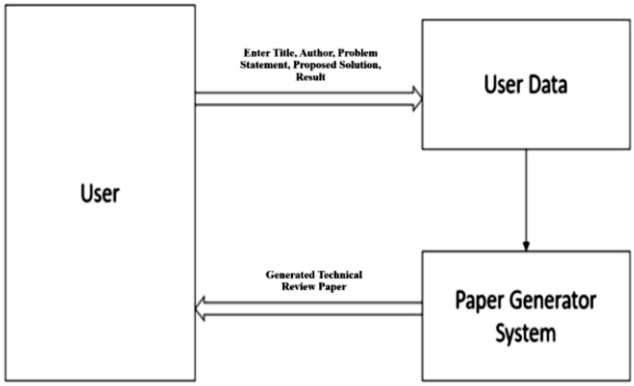


Fig. 1: Block Diagram

Our innovative system streamlines the process of generating technical review papers. Users initiate the workflow by entering essential details such as the title, author information, problem statement, proposed solution, and results of their research topic. These inputs act as the foundation for our intelligent system. The system then employs advanced algorithms to analyze and synthesize the provided information, automatically composing a comprehensive technical review paper. The block diagram vividly illustrates each step of this seamless process, showcasing how user inputs are transformed into a professionally crafted research paper. This efficient and user-friendly approach significantly reduces the time and effort traditionally required for paper creation while maintaining the high standards expected in technical literature.

Fig. 2 illustrates the proposed methodology for an AI-based technical review Paper generator and the detailed workflow of the project.

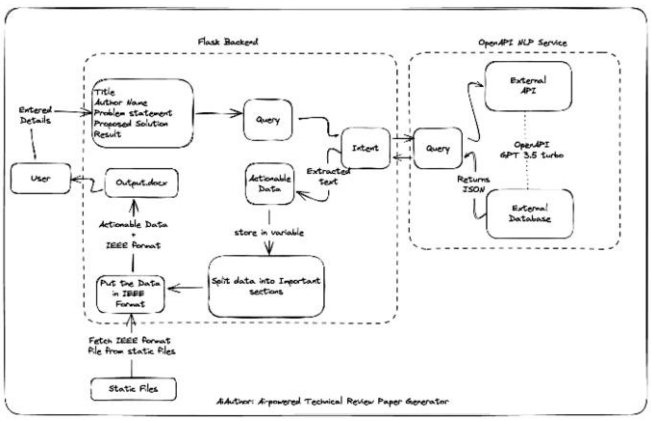


Fig. 2: Proposed Methodology

Fig. 3 is the flowchart of our project, which shows the step-by-step procedure of our system.

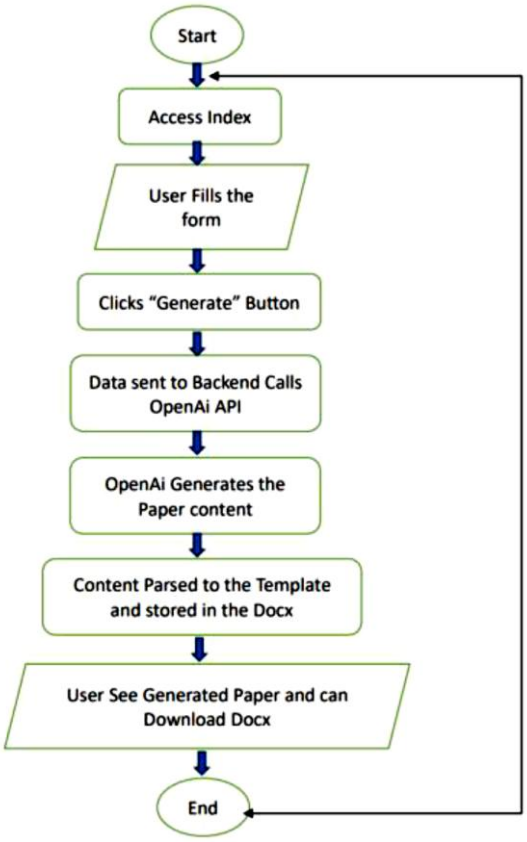


Fig. 3: Flow Chart

1. When the user first visits the website, they land on the home page. From the home page, they click the "Generate Research paper" button to go to the form.
2. This takes them to the form page where they enter details like title, author, problem statement, etc.
3. Once filled, they click the "Generate" button. When this button is clicked, the form data is sent as a POST request to the backend.
4. In the backend Python code, it first calls the OpenAI API to generate the paper content.
5. OpenAI takes the input data and uses its NLP models to generate the different sections of the paper like abstract, introduction etc.
6. The generated content is then parsed and inserted into a predefined template (DOCX). This populated DOCX file is saved on the server.
7. The DOCX is then converted to a PDF file using a library.
8. The user is redirected to a results page where they can view the generated paper PDF inside a frame. On this page, they can download the DOCX or PDF files.
9. After downloading, they can be sent back to the home page to generate another paper or end their session.

V. IMPLEMENTATION

In the development of the AiAuthor, a comprehensive approach has been undertaken to seamlessly blend frontend and backend technologies, creating a user-friendly platform for the automated generation of high-quality technical review papers. The front end is meticulously crafted using a combination of HTML, CSS, and JavaScript to provide distinct pages, including a homepage, input form, and result/download sections. The homepage serves as an

informative gateway, offering users a detailed description of the problem and the application's value proposition, guiding them toward the input form. This inputs form, implemented with validation mechanisms, captures essential details such as title, author, problem statement, proposed solution, and results. Upon submission, the form triggers an API call to the backend, which is developed using the Python Flask web framework. The backend hosts API endpoints for data submission and retrieval, seamlessly integrating with the OpenAI API to leverage the powerful capabilities of GPT-3 models for natural language generation.

The following table provides a structured overview of the key components and technologies used in the AiAuthor implementation.

Table 1. Components and Technologies used in AiAuthor

Step	Description
Frontend Development	Crafting the frontend using HTML, CSS, and JavaScript. Creating distinct pages like the homepage, input form, and result/download sections. The homepage includes a problem description and value proposition.
Input Form	Creating a page with fields for essential details. Implementing validation mechanisms for required fields and a submit button. Triggering API call to the backend upon submission.
Backend Development	Developing the backend using Python Flask. Hosting API endpoints for form data submission and retrieval of the generated paper. Integrating with the Open AI API for natural language generation.
Document Generation	Using GPT-3 API to dynamically generate paper sections based on user input. Employing regex patterns to extract relevant sections. Populating Word/Docx template and converting it to PDF for compatibility.
Result Page	Presenting the generated PDF to the user with a download button for obtaining PDF or Docx file. Providing additional information & instructions.
Deployment	Deploying the application on a cloud platform (e.g., Vercel, Heroku or AWS) for high availability and scalability. Implementing robust error-handling mechanisms for enhanced reliability.

6. RESULT

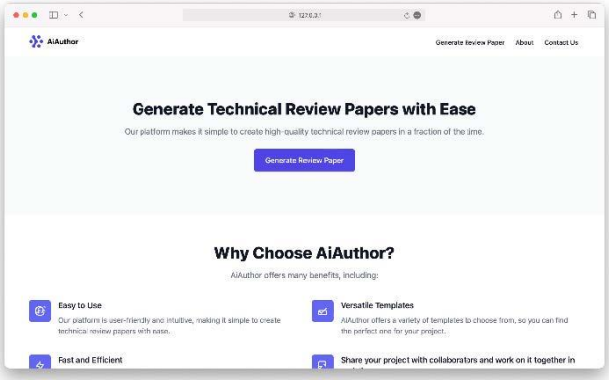


Fig. 4: Home Page

Fig. 4 is the homepage of AiAuthor, which serves as an informative gateway, offering users a detailed description of the problem and the application's value proposition, guiding them toward generating a Research paper.

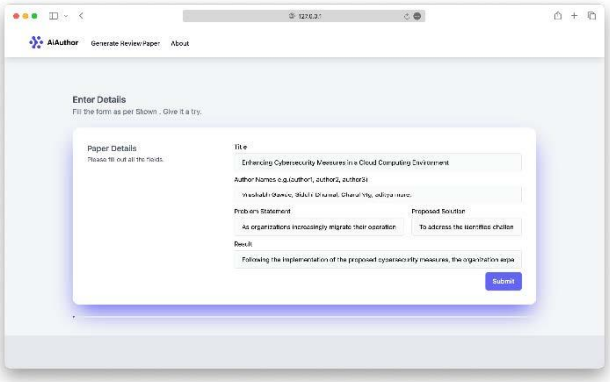


Fig. 5: Paper Details

Fig. 5 is the paper details, here users enter essential details, including the title, author, problem statement, proposed solution, and results, facilitating a streamlined process for the automated generation of technical review papers.

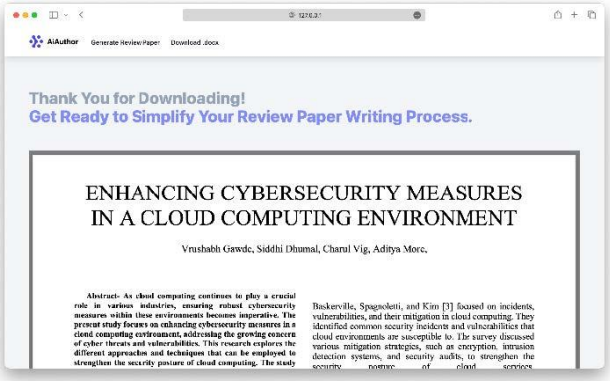


Fig. 6: Generated Paper

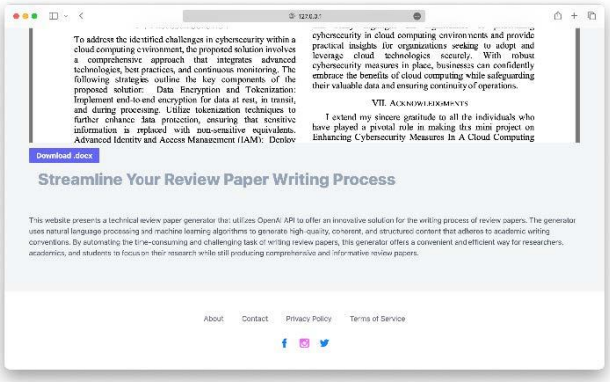


Fig. 7: Download Paper

Fig. 6 and Fig. 7, present the generated PDF to the user with a download button for obtaining PDF or Docx files. Providing additional information and instructions.

7. CONCLUSION

The AiAuthor is a powerful tool for generating technical review papers. It can greatly reduce the time and effort needed to produce high-quality papers while maintaining accuracy and efficiency. The use of Open AI API provides a reliable and flexible solution that can be tailored to the specific needs of different users. By using the AiAuthor, researchers can focus more on data analysis and interpretations rather than on organizing and writing technical review papers.

In this paper, we have presented a comprehensive exploration of the potential and challenges surrounding the development of a technical review paper Generator, powered by advanced natural language processing (NLP) and artificial intelligence (AI) techniques. The landscape of scholarly publishing and technology dissemination is rapidly evolving, and our proposed solution addresses the limitations of traditional manual research paper creation through our in-depth analysis, it becomes evident that an AI-driven technical review paper Generator offers a transformative approach to generating research papers. By leveraging NLP models like GPT-3.5, we can significantly enhance the efficiency, accuracy, and relevance of research paper creation. The generator's ability to automate literature selection, summarization, and content synthesis streamlines the process, allowing researchers to focus on higher-level analysis and insights.

Furthermore, the generator's dynamic update capabilities ensure that the generated research papers remain current and reflective of the latest advancements in the field. This aspect is particularly crucial in the ever-evolving landscape of technology where new research emerges rapidly.

Our proposed solution also emphasizes the importance of maintaining a balance between AI automation and the human element. Researchers' expertise in reviewing, enhancing, and contextualizing the generated content contributes to the depth and quality of the final research papers.

While the Technical review paper Generator holds immense promise, we acknowledge the ethical considerations that come with AI-generated content, including proper attribution, bias mitigation, and privacy concerns. These factors underscore the need for responsible development and deployment of AI technologies in the academic and research domains.

In conclusion, the technical review paper Generator represents a significant advancement in the realm of scholarly publishing. Its potential to accelerate the dissemination of knowledge, enhance the accessibility of research insights, and adapt to the dynamic nature of technology positions it as a powerful tool for researchers and practitioners alike.

As AI and NLP technologies continue to evolve, we anticipate that the technical review paper Generator will play a pivotal role in shaping the future of how technical review papers are created and shared.

8. ACKNOWLEDGMENTS

We extend my sincere gratitude to all the individuals who have played a pivotal role in making this mini project on technical review paper Generator Using Open AI API a success. I am immensely grateful to my mentor, (Ms. Vinita Bhandiwad), whose guidance, encouragement, and support have been invaluable in every step of this project. Without their expertise and timely assistance, the completion of this project would not have been possible. I also thank my team members, for their relentless hard work, dedication, and cooperation throughout the project. Their inputs and insights have significantly contributed to the quality of the project. I am thankful to the management of (Vidyalankar Institute of Technology) for providing us with all the necessary resources such as space, equipment, software, and infrastructure, which were essential in accomplishing this project. In conclusion, I express my sincere appreciation to all who have contributed to the successful completion of this project.

REFERENCES

[1] A. O. Thunström, "We Asked GPT-3 to Write an Academic Paper about Itself—Then We Tried to Get It Published," Scientific American, 30 June 2022. [Online]. Available: https://www.scientificamerican.com/article/ we-asked-gpt-3-to-write-an-academic-paper-about-tself-mdash-then-we-tried-to-get-it-published/. [Accessed 30 June 2022].

[2] OpenAI, "Research index - OpenAI," OpenAI, 2024. [Online]. Available: https://openai.com/research.

[3] O. ‘. Buruk, "[2304.11079] Academic Writing with GPT-3.5: Reflections on Practices, Efficacy and Transparency," arXiv.org, 12 Feb 2023. [Online]. Available: https://arxiv.org/abs/2304.11079.

[4] MLQ, "Building a GPT-3 Enabled Research Assistant with LangChain & Pinecone - MLQ," MLQ, 2023. [Online]. Available: https://www.mlq.ai/gpt-3-enabled-research-assistant-langchain-pinecone.

[5] B. K. Mingyu Zong, "A Survey on GPT-3," arXiv, 2022.

[6] Andy1, 23 june 2023. [Online]. Available: https://academiainsider.com/ai-tools-for-research-papers-and-academic-research.

[7] "AI in Technical Writing: 7 ways How to Use AI Writing Assistant to Write Awsome Content," 30 Jan 2023. [Online]. Available: https://www.misaias.com/ai-in-technical-writing.

[8] Madalsa, "AI for thesis writing — Unveiling 7 best AI tools - SciSpace by Typeset," 12 October 2023. [Online]. Available: https://typeset.io/resources/ai-for-thesis-writing.

[9] D. Lawal, "Technical Writing with AI: Ways to Use AI Writing Assistant to Write Technical Content," 2 September 2023. [Online]. Available: https://dev.to/daslaw/technical-writing-with-ai-ways-to-use-ai-writing-assistant-to-write-technical-content-17g1.

[10] A. Dyouri, "How To Use Web Forms in a Flask Application," 05 November 2021. [Online]. Available: https://www.digitalocean.com/community/tutorials/how-to-use-web-forms-in-a-flask-application.

[11] "Flask form - Python Tutorial," [Online]. Available: https://pythonspot.com/flask-web-forms.

[12] "How To Use Web Forms in a Flask Application - GeeksforGeeks," [Online]. Available: https://www.geeksforgeeks.org/how-to-use-web-forms-in-a-flask-application.

[13] A. Dyouri, "How To Use and Validate Web Forms with Flask-WTF," 21 December 2021. [Online]. Available: https://www.digitalocean.com/community/tutorials/how-to-use-and-validate-web-forms-with-flask-wtf.

[14] "Python | Introduction to Web development using Flask - GeeksforGeeks," [Online]. Available: https://www.geeksforgeeks.org/python-introduction-to-web-development-using-flask.

[15] A. Joy, "What Can You Do With Flask? (Applications & Examples) – Pythonista Planet," [Online]. Available: https://pythonistaplanet.com/what-can-you-do-with-flask.

[16] M. Lathkar, Building Web Apps with Python and Flask, 2021.

[17] "AI Literature Review Generator | Taskade," [Online]. Available: https://www.taskade.com/generate/research/literature-review.

[18] "AI for Writing Literature Review," [Online]. Available: https://atonce.com/blog/ai-for-writing-literature-review.

[19] "AI Tools for Literature Review," [Online]. Available: https://mraitool.com/ai-tools-for-literature-review.

[20] OpenAI, "GPT-3 powers the next generation of apps - OpenAI," 25 March 2021. [Online]. Available: https://openai.com/blog/gpt-3-apps.

[21] "AutoSurveyGPT: Automatically literature survey/review with GPT! An intelligent research assistant leveraging GPT-3.5/GPT-4 to find, analyze, and rank relevant academic papers from Google Scholar based on user-provided search queries and topics," [Online]. Available: https://github.com/a554b554/AutoSurveyGPT.

[22] T. Simonite, "Did a person write this headline, or a machine? GPT-3, a new text-generating program from OpenAI, shows how far the field has come – and how far it has to go," 27 July 2021. [Online]. Available: https://www.wired.com/story/ai-text-generator-gpt-3-learning-language-fitfully.

[23] A. Dyouri, 22 December 2022. [Online]. Available: https://www.digitalocean.com/community/tutorials/how-to-make-a-web-application-using-flask-in-python-3.

[24] "Flask · PyPI," 3 Feb 2024. [Online]. Available: https://pypi.org/project/Flask.

[25] GeeksforGeeks, "Flask Tutorial," [Online]. Available: https://www.geeksforgeeks.org/flask-tutorial.

[26] GeeksforGeeks, 19 December 2017. [Online]. Available: https://www.freecodecamp.org/news/how-to-use-python-and-flask-to-build-a-web-app-an-in-depth-tutorial-437dbfe9f1c6.

❑ ❑ ❑ ❑ ❑ ❑

AET 023

YogaWise: Enhancing Yoga with Intelligent Real Time Tracking using TensorFlow MoveNet

Amey Parle
B.E., Department of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
amey.parle@vit.edu.in

Rugved Shinde
B.E., Department of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
rugved.shinde@vit.edu.in

Rahul Chougule
B.E., Department of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
rahul.chougule@vit.edu.in

Shruti Agrawal
Asst. Prof., Department of INFT
Vidyanagar Institute of Technology, Mumbai, India
shruti.agrawal@vit.edu.in

***Abstract*— This work proposes a novel method that uses image processing and ensemble learning to analyze yoga practices thoroughly, in response to the growing need for accurate assessment in the field of yoga. The main goal is to help users execute yoga postures efficiently and in real time. With the use of the MoveNet model, the system was able to obtain a testing accuracy of 98% as well as strong precision, recall, and F1 Score metrics, demonstrating its ability to recognize and classify a variety of yoga poses with a low number of false positives and negatives. Methodologically, a large-scale dataset was gathered, cleaned up, and used to train a deep learning model that was then smoothly incorporated into an easy-to-use online application. The proposed system offers pre-designed yoga sessions, different types of pain relief routines, and individual pose exploration through an intuitive user interface. It also offers real-time feedback, posture prediction all of which improve users' yoga experiences and help them move closer to their health and fitness objectives.**

Keywords—Yoga, Pose Estimation, MoveNet, CNN, Computer Vision.

I. INTRODUCTION

The combination of image recognition and computational intelligence has brought new perspectives on personal fitness and health in the age fueled by technology wellness solutions. This system revolves around yoga pose detection utilizing the state-of-the-art MoveNet Thunder, offering a distinctive fusion of personalized yoga sessions tailored for users within the comfort of their homes. Notably, these sessions unfold in real-time, providing a dynamic and individualized experience.

This initiative's main motivation stems from the growing demand for accessible and personalized fitness solutions. In the face of our increasingly hectic daily lives, the search for convenient yet effective methods to prioritize physical well-being has become more pronounced. [1] shows yoga's effectiveness in reducing stress and anxiety, known contributors to pain perception. Beyond this, studies like [2] suggest a deeper impact on our neuropsychological pathways. However, the true power of yoga lies in its personalization. Unlike standardized interventions, yoga can be tailored to individual needs and pain experiences. This is crucial for chronic pain sufferers, as research like [3] demonstrates. This system aims to address this gap by utilizing MoveNet Thunder's pose detection unlocks customized yoga experiences.

This system analyzes any person's fitness level and pain points (back, knee, hip, etc.) to design targeted exercises, making yoga therapeutic and accessible for everyone. The prospect for tailored yoga experiences to have a beneficial influence on individual's health and quality of life, which is both exciting and transformative.

This research study is structured as follows. The introduction of the paper is mentioned in Section I. Survey of related works is mentioned in Section II. Section III explains the proposed algorithm for yoga pose detection. In Section IV, the implementation process is discussed. The results acquired are mentioned in Section V, and finally, conclusions are drawn in Section VI.

II. RELATED WORKS

Yoga position estimate pioneering studies employed Multilayer Perceptrons (MLPs) trained on angle measurements between key body points. While these methods achieved decent accuracy, limitations arose in real-time performance and robustness to camera angles Notably, MLPs were shown to be effective for posture estimation and assistance by Patil et al. [4] and Raj et al. [5], but they were limited to non-real-time conditions because of computational constraints. Additionally, these approaches relied on handcrafted features, making them susceptible to variations in camera angles and lighting conditions.

Recent research pushes the boundaries of yoga posture prediction accuracy. Contrastive learning techniques, as in [6], offer promising results by addressing pose variations and standard benchmarks using skeleton features This is in line with the paper's emphasis on skeleton features and contrastive learning, indicating room for additional development through richer feature sets and data augmentation.

High-performance models like SEPAM_HRNet [7] employ multi-scale feature extraction and attention mechanisms for yoga pose estimation, demonstrating superior performance on challenging poses and occlusions. This highlights the potential of attention mechanisms & multi-scale features for improving pose estimation accuracy, particularly in occluded scenarios.

This study by Chung et al. (2022) [8] compared prominent human pose estimation (HPE) libraries, revealing compelling insights for researchers like yourself. Notably, MoveNet emerged as the top performer across image and video datasets, surpassing alternatives in key areas. While MediaPipe Pose exhibited decent video performance, its overall accuracy trailed MoveNet. PoseNet, delivering consistent results, it was

less accurate overall than both its video-focused counterpart and MoveNet.

OpenPose, although demonstrating strong image-based performance, stumbled in video scenarios, particularly with self-occlusion. Its bottom-up approach struggles to maintain tracking during occlusions, leading to significant performance drops.

MoveNet, in contrast, shines for its superior overall performance and resilience to challenges. Its high PDJ values across both image and video formats, coupled with its ability to handle camera angles and self-occlusion effectively, distinguish it from the other libraries. Additionally, MoveNet outperformed its competitors in more actions compared to its peers, showcasing its broader applicability.[8]

Previous Methods Faced Common Setbacks:

1) **Limited Real-Time Performance:** Prior studies often exhibited trade-offs between accuracy and processing speed, hindering their suitability for real-time applications.

   Existing methods struggled to achieve the low latency demanded by interactive systems.

   Real-time performance remained a challenge due to computational limitations.

2) **Sensitivity to Camera Angles and Lighting Conditions:** Previous works were susceptible to variations in camera viewpoints and lighting, leading to performance degradation.

   Existing approaches relied on features that were not robust to changes in pose due to camera angle or illumination.

   Sensitivity to environmental factors limited the generalizability of prior methods.

3) **Inaccurate Handling of Occlusions:** Earlier studies often exhibited difficulties in accurately tracking poses when body parts were obscured.

   Occlusions presented a significant challenge to maintaining pose estimation accuracy.

   Existing methods struggled to handle overlapping body parts, leading to tracking errors.

III. PROPOSED ALGORITHM

This paper presents a novel algorithm for enhanced yoga practice through the integration of cutting-edge pose estimation and key point tracking capabilities. This system employs the MoveNet model, a deep learning-based pose estimation model, due to its high performance and accuracy.

A. Overview of MoveNet Model:

TensorFlow MoveNet is a machine learning model designed specifically for human pose estimation tasks. It utilizes deep learning techniques to detect and track key points on the human body in real-time.

B. MoveNet Pose Estimation Process

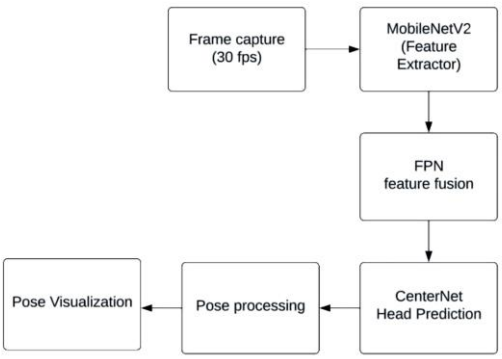


Fig. 1 Pose classifier

Figure 1 illustrates the process flow for pose estimation of MoveNet model.

1) **Frame Capture:** The video frame is captured at a predetermined rate, for example, 30 frames per second (fps). This ensures a smooth and real-time experience.

2) **MobileNetV2 Feature Extraction:** The captured frame undergoes resizing and normalization to match the expected input size and format of the model. This standardization ensures compatibility with the subsequent processing steps. The pre-processed frame enters the MobileNetV2 backbone network, the heart of MoveNet Thunder. This network consists of several inverted residual blocks that extract features at various resolutions and progressively capture higher-level semantic information.

Key Components:

a) **Inverted Residual Block:** It combines depth wise separable convolutions and linear bottlenecks to efficiently extract features while maintaining computational efficiency.

b) **Stage-wise Processing:** Multiple stages with increasing channel depth and decreasing spatial resolution extract features of varying granularity. This allows the model to capture both fine-grained details and broader context within the image.

c) **Output:** MobileNetV2 generates multiple feature maps at different scales, each representing the captured frame with varying degrees of detail and semantic understanding.

3) **FPN Feature Fusion:** The Feature Pyramid Network (FPN) takes the multi-scale feature maps from MobileNetV2 and combines them into a unified representation. This allows the subsequent CenterNet head to access information from all scales, crucial for accurate pose estimation.

Its Key Components include:

a) **Top-down Pathway:** Higher-level feature maps are upsampled and projected to match the spatial resolution of lower levels, preserving high-level semantic context.

b) **Lateral Connections:** Skip connections directly combine upsampled high-level features with corresponding lower-level features, adding fine-grained details and spatial precision to the higher levels.

c) **Output:** FPN creates a feature pyramid with multiple levels, each containing rich information from various scales, providing a comprehensive representation of the person's pose within the frame.

4) **CenterNet Head Prediction:** The CenterNet head utilizes the information-rich feature pyramid from FPN to predict the keypoint locations and bounding box of the person within the frame. This innovative approach simplifies the pipeline compared to traditional heatmap-based methods.

Its Key Components include:

a) **Keypoint Triplet Representation:** Instead of predicting individual keypoints, CenterNet predicts a center keypoint located at the estimated center of the person's bounding box and two offset keypoints that refine the locations of body extremities. This representation improves accuracy and reduces ambiguity.

b) **Direct Keypoint Location Prediction:** CenterNet directly predicts the keypoint coordinates within the image, eliminating the requirement for intermediate heatmap generation and enhancing efficiency.

c) **Output:** CenterNet outputs: The coordinates of the center keypoint and offset keypoints for all 17 key-points used to depict the subject's posture.

The predicted bounding box encompassing the person's location in the frame.

5) **Post-processing:** The predicted keypoints and bounding box undergo filtering and refinement based on confidence scores. This step removes unreliable predictions and improves the overall accuracy of the pose estimation.

6) **Pose Visualization:** The person's pose can be seen in real time by overlaying the final, corrected key-points over the original frame. This gives the estimated stance an appearance and enables a number of uses, including gesture recognition and workout tracking. While real-time pose estimation is challenging, MoveNet addresses limitations through various strategies. Its efficient architecture and hardware acceleration enable smooth video processing. Additionally, the CNN architecture analyzes local features, making it less sensitive to camera angle variations.

C. Strategies to Address Previous Limitations:

1) **Limited Real-Time Performance:** Efficient Architecture: Our proposed algorithm utilizes the MoveNet Thunder variant, specifically designed for real-time tasks, offering a balance between speed and accuracy. Quantization and Model Optimization: We employ quantization and optimize the model for efficient backends like TensorFlow Lite, ensuring fast computations on various devices, including resource-constrained platforms.

2) **Sensitivity to Camera Angles and Lighting Conditions:** Data Augmentation during Training: The algorithm enriches the training dataset with diverse images, covering various camera angles and lighting scenarios, enhancing the model's ability to generalize across different conditions. Pose Estimation Network Architecture: The convolutional neural network (CNN) architecture is tailored to extract features resilient to changes in lighting and camera angles.

3) Inaccurate Handling of Occlusions: Contextual Inference and Confidence Scores: While our algorithm doesn't have a dedicated occlusion handling mechanism, the CNN architecture enables contextual inference, making educated guesses about occluded body parts. Low confidence scores in keypoints help identify and downweight unreliable predictions, mitigating the impact of occlusions.

D. The proposed methodology comprises of following steps:

1)Data gathering: To estimate yoga poses from real-time condition, train the model on a dataset that has well-lit, high-resolution images depicting individuals in clear and unambiguous poses. Diverse poses from various angles across various subjects, backgrounds, and clothing styles are crucial for robust model training.

Different collections of yoga positions gathered from various dataset sources.

Dataset sources: Laurence Moroney, Kaggle and Roboflow

Laurence Moroney: New Dataset for Yoga Pose Classification.

Kaggle: Yoga Pose Image classification dataset by Shruti Saxena.

Yoga Posture Dataset by Mrinal Tyagi

Roboflow: Yoga Pose Dataset by new-workspace-mujgg.

Sample Count: About 500 images per class.

Total count: About 7000 images.

For model training and testing, the dataset was split into training (70%), testing (30%).

2) Data preprocessing : Raw yoga pose images underwent a comprehensive pre-processing pipeline to ensure high-quality training data. Each image has 17 important landmarks that were retrieved using the MoveNet model, serving as fundamental descriptors of pose variations and spatial relationships. Subsequently, a rigorous normalization procedure aligned all poses to a common reference frame, mitigating the impact of diverse camera angles and body sizes. Landmarks were scaled to a standardized range and centred within a predefined space, resulting in a consistent and structured dataset. Each yoga pose was then uniquely represented by its set of normalized landmarks, meticulously stored in individual CSV files. This meticulously prepared dataset, devoid of extraneous technical identifiers, served as the bedrock for subsequent training of the yoga posture categorization model.

3) Model Creation and Training: Pose landmarks are extracted, normalized, and transformed into compact embeddings. To avoid overfitting, these are fed into a neural network that has hidden layers and dropout. The model trains until performance plateaus, saving the best version along the way.

The model is designed to classify poses based on 34-dimensional pose embeddings derived from landmark

coordinates. It employs a sequential architecture with the following layers:

a) **Input layer:** Accepts a 34-dimensional vector representing the pose embedding.

b) **Dense layers:** ReLU6 activates two dense layers containing 128 and 64 neurons, respectively, to add non-linearity and discover complex relationships in the data.

c) **Dropout layers:** Two dropout layers with a rate of 0.5, randomly dropping connections during training to prevent overfitting.

d) **Output layer:**

A dense layer with the amount of neurons equal to the amount of pose classes, using the softmax activation function to produce a probability distribution over the classes.

The model is prepared for the learning process. This involves:

- **Loss Function:**

Employing categorical cross-entropy, a common choice for multi-class classification tasks, to quantify the model's error during training. It measures the discrepancy between the predicted probabilities and the true labels for each frame.

- **Optimizer:**

Adam optimizer is employed for adjusting the model parameters to reduce the loss function and enhance performance.

The training process optimized the model using the Adam optimizer and categorical cross-entropy loss function. With 200 epochs, a batch size of 139, and checkpoint callbacks for monitoring validation accuracy, the model demonstrated impressive results – a training loss of 0.0591 and an accuracy of 98%. The trained model was stored as a TensorFlow.js (tfjs) model, enabling seamless integration into web applications through JSON format for efficient utilization.

4) Model Testing

a) **Threshold Determination:** The pose detection threshold needs to be determined using metrics like the ROC curve to balance false positives and false negatives.

b) **Evaluation:** Evaluating the trained model on the testing set, and analysing metrics such as Receiver Operating Characteristic Area Under the Curve (ROC-AUC), F1-score, accuracy, recall, loss, and precision is carried out.

E. MoveNet is particularly suitable for pose estimation tasks in yoga due to several reasons:

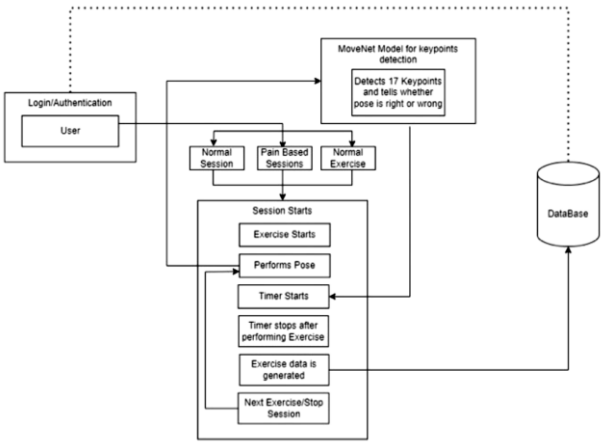
1) **Real-time Performance:** MoveNet is optimized for real-time performance, making it suitable for applications where immediate feedback is required, such as YogaWise. It can process video frames quickly, enabling seamless tracking of yoga poses as they are performed.

2) **Accuracy and Robustness:** MoveNet has been trained on large datasets containing diverse yoga poses, which helps improve its accuracy and robustness in detecting various yoga postures. This ensures that even subtle variations in poses can be accurately recognized and tracked.

3) **Lightweight and Efficient:** MoveNet is designed to be lightweight and efficient, making it suitable for deployment on a variety of devices, including laptops and mobile phones. This allows YogaWise to offer real-time tracking without requiring powerful hardware.

IV. IMPLEMENTATION

Figure 2 demonstrates the implementation process of the proposed system:



A. Integrating it with WebApp

With the use of the TensorFlow.js (tfjs) framework, which makes integration simple, a trained deep learning model is integrated with the frontend to enable features such as posture prediction, real-time feedback, & targeted recommendations.

The application employs a React-based frontend and Node.js backend for a seamless user experience. Three options are offered: pre-designed yoga sessions, personalized pain relief routines, and individual pose exploration. Node.js powers the backend, handling user requests, controlling sessions, and enabling interaction with the MongoDB database, which stores user data, session details, and progress history.

Hardware Requirements:

1. **Device:** Currently YogaWise is compatible with Laptops, Desktops.
2. **Camera:** A built-in or external camera capable of capturing video footage is necessary for pose detection. For optimal results, the camera should have decent resolution and frame rate capabilities.
4. **Internet Connection:** A stable internet connection is necessary to access the YogaWise website.

Software Requirements:

1. **Operating System:** YogaWise supports common operating systems such as Windows, macOS, Linux, iOS, or Android.
2. **Web Browser:** The client device needs a modern web browser to access the YogaWise website. Compatibility with popular browsers like Google Chrome, Mozilla Firefox, Safari, and Microsoft Edge is essential.

B. Real-time Monitoring:

To achieve real-time tracking performance in YogaWise, the system employs a client-side execution model for pose estimation. The MoveNet Thunder pose detection algorithm is

loaded directly within the user's browser, which means that the computation of body keypoints from the video feed occurs on the client side. By processing data locally rather than relying on server-side computations, the system minimizes latency and facilitates the immediate tracking of yoga poses. This localized approach is key to enabling the real-time feedback necessary for users during their yoga sessions.

The application aims to maintain a consistent frame rate by dynamically adjusting the computational load based on the device's performance. This ensures a smoother visual experience.

C. Body angle Measurements:

The application measures body angles in real time, offering users insights into their posture and improvement.

To measure the angle between body keypoints, this system uses vector dot product and magnitude calculations. It takes three points as input, treats one as the reference, and calculates the angle between the remaining two relative to it. For example, right_hip, right_knee and right_foot, using right_elbow as reference.

The main formula used for calculating the angle is:

$$\theta = \arccos((\text{vector1} \cdot \text{vector2}) / (\|\text{vector1}\| * \|\text{vector2}\|))$$

where:

- θ: angle between vectors in radians
- vector1: first vector
- vector2: second vector
- ·: dot product operator
- *: cross product operator
- ||…||: magnitude operator (square root of sum of squares)

D. User Interaction Flow

Homepage*:* Upon entering the website, users are presented with a clean and intuitive interface.

Three options are available: "Yoga Sessions," "Pain Remedy Session," and "Individual Yoga Pose."

a) Yoga Sessions*:* This option offers a dynamic flow of random yoga poses, progressing users to the next pose upon correct completion.

b) Pain Remedy Session: In this option, users can choose from targeted yoga sessions designed to address specific pain categories like hip, shoulder, knee, or back pain. This system suggests appropriate poses based on your selection, helping you find relief and improve mobility.

c) Individual Yoga Pose: In "Individual Yoga Pose," users concentrate on doing just one yoga pose at one moment.

E. System Functionality during Pose Execution:

1) **Real-time Visual Feedback:**Upon pose initiation, the system overlays a skeletal representation on the user's video feed. Initially white, the skeleton dynamically transitions to green upon successful pose recognition using computer vision algorithms. This immediate colour change provides visual confirmation of correct pose execution.

   In addition to this, the system also tracks users body angle measurement in real-time, which helps to monitor user's progress based on set of reference angles.

   To ensure optimal accuracy in body angle measurement, users should maintain a frontal position relative to the camera. This is due to the system's reliance on a pre-defined set of reference angles for pose assessment.

2) **Timer-based Pose Duration Measurement:** Achieving the correct pose triggers a timer activation. This timer measures the duration the user maintains the pose, providing objective feedback on pose hold time. Automatic timer stoppage occurs when the pose is no longer detected, signifying pose completion.

3) **Data Logging and Performance Tracking:** YogaWise logs pose duration and other relevant data (e.g., pose type, timestamp) into a MongoDB database. This comprehensive data logging facilitates user performance analysis and tracking progress over time.

4) **Performance History:** Users can access a dedicated "History Section" within the application. This section visualizes performance data across sessions, enabling users to track progress, identify improvement areas, and assess consistency in maintaining correct form.

F. Conclusion

The combination of React, Node.js, and MongoDB creates a robust and dynamic platform for users to engage in personalized yoga sessions. Real-time monitoring, visual feedback through skeletons, and detailed history tracking contribute to a holistic and user-centric yoga experience. This architecture ensures scalability, responsiveness, and adaptability to evolving user needs.

V. RESULTS AND DISCUSSION

The smart yoga pose tracking system, employing the MoveNet model, was thoroughly tested using a range of performance measures to see how well it classified yoga poses. The system achieved a remarkable testing accuracy of 98%, underscoring its ability to precisely recognize and categorize diverse yoga postures. The precision, recall, and F1 Score of 98.1%, 98.2%, and 98.1%, respectively, highlight the system's capability to minimize false positives and negatives, ensuring a high level of accuracy in pose identification. The Matthews Correlation Coefficient reached 98.0, emphasizing the strong correlation between predicted and actual classifications. The confusion matrix further illustrates the system's proficiency across different pose categories, with minimal misclassifications. Class-specific ROC-AUC scores, ranging from 0.9924 to 1.0000, showcase the model's discriminative power for each yoga pose. Figure 3 shows the graphical representation for ROC-AUC score.

While the overall system performance is commendable, more research may examine ways to make it even better, such as exploring interpretability mechanisms or fine-tuning the model for enhanced adaptability to diverse yoga practices and user nuances.

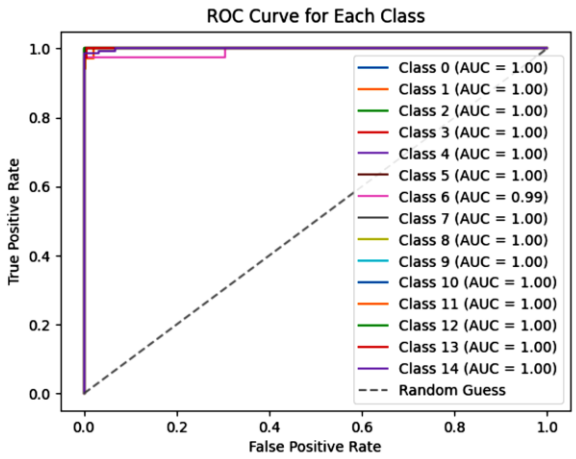


Fig. 3 ROC-AUC Graph

Table 1 : ROC-AUC SCORES

Yoga Pose	ROC-AUC Score
anjaneyasana (crescent pose)	1.000
ardhachandrasana (half moon pose)	0.9995
baddha konsana (bound angle pose)	1.000
chair	1.000
cobra	1.000
dog	1.000
Malasana (garland pose)	0.9996
no_pose	1.000
salamba bhujangasana(sphinx)	1.000
shoudler_stand	1.000
traingle	0.9996
tree	1.000
ustrasana (camel pose)	0.9998
uttan shisosana (extended puppy pose)	0.9913
warrior	0.9998

Table 1 : shows ROC-AUC score for each class

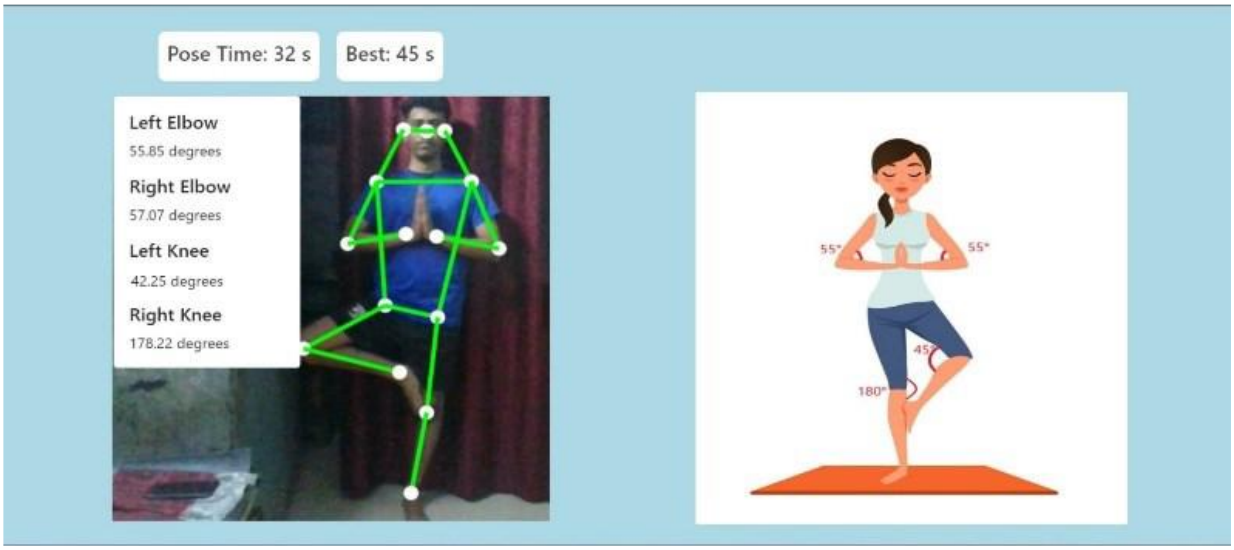


Fig. 4 Visual representation of body angle measurements

Figure 4 shows real-time angle measurement, which helps users improve their yoga practice by monitoring their body angles. Users can refer to the pre-defined sets of angles provided to them.

While the initial implementation of the proposed MoveNet-based yoga pose tracking system has yielded notable results, further advancements are crucial to solidify its position as a leading solution in the domain of intelligent yoga practice assistance. This can be achieved by leveraging following existing methodologies and integrating them into the current system, leading to further advancements in the Yoga Pose Tracking system.

1) Incorporation of Temporal Information:

While the current model excels in static pose estimation, integrating temporal information into the pose tracking system can enhance its performance for dynamic yoga practices, such as vinyasa or flow-based sequences. Models capable of tracking pose transitions and temporal dependencies can provide more comprehensive feedback to users during dynamic yoga sessions.

2) Multi-modal Fusion:

Combining information from multiple modalities, such as video, audio, and wearable sensors, can enrich the understanding of yoga poses and improve the robustness of the tracking system. By integrating data from wearable devices or audio cues, the model can better capture subtle nuances in posture and movement, leading to more accurate pose classification and tracking.

VI. CONCLUSION

In conclusion, this study, titled "YogaWise: Enhancing Yoga with Intelligent Real Time Tracking using TensorFlow MoveNet" signifies a significant advancement in leveraging

technology to enhance yoga practice. The developed web application ensures widespread accessibility, allowing users to seamlessly engage in yoga sessions from any location, fostering a culture of well-being and flexibility. Tailored sessions addressing specific concerns such as back pain, knee pain, and shoulder-neck pain provide users with a personalized and therapeutic yoga experience.

The achieved accuracy, approximately 98%, underscores the reliability of this real-time pose detection system. The smooth transition between yoga poses in 30-second intervals enhances user engagement, ensuring a dynamic workout experience. Moreover, this web app introduces a pioneering feature—angle verification. Through meticulous analysis of joint angles during poses, users can learn plenty about their form and progress over time. This novel method encourages users to attain optimal pose accuracy but also quantifies their growth in mastering each pose, fostering a sense of accomplishment.

The impact of this research is evident in its contribution to merging technology with holistic well-being. By harnessing MoveNet's capabilities and incorporating personalized sessions and angle verification, this web app provides a transformative yoga experience that is accessible, accurate, and tailored to individual needs. This system acknowledges the 3D limitations of this model, suggesting the potential need for multiple cameras to achieve high-accuracy 3D views. Despite this challenge, there is a continued evolution of real-time yoga pose detection system, anticipating substantial progress in developing new models capable of accurately detecting 3D postures in future iterations.

REFERENCES

[1] Hendriks T, de Jong J, Cramer H. The Effects of Yoga on Positive Mental Health Among Healthy Adults: A Systematic Review and Meta-Analysis. J Altern Complement Med. 2017 Jul;23(7):505-517. doi: 10.1089/acm.2016.0334. Epub 2017 Feb 28. PMID: 28437149.

[2] Misra P, Mandal S, Sharma G, et al. (August 07, 2023) Effect of Structured Yoga Program on Stress and Well-being Among Frontline Healthcare Workers During COVID-19 Pandemic. Cureus 15(8): e43081. doi:10.7759/cureus.43081

[3] Breaking a pose into angles developers.google.com https://developers.google.com/ml-kit/vision/pose-detection/classifying-poses (Accessed 18 Feb 2024)

[4] Vivek Anand Thoutam, Anugrah Srivastava, Tapas Badal, Vipul Kumar Mishra, G. R. Sinha, Aditi Sakalle, Harshit Bhardwaj, Manish Raj, "Yoga Pose Estimation and Feedback Generation Using Deep Learning", Computational Intelligence and Neuroscience, vol. 2022, Article ID 4311350, 12 pages, 2022. https://doi.org/10.1155/2022/4311350

[5] Jagtap, R., Zanzane, M., & Patil, R. (2022, May). Yoga pose detection using machine learning. International Research Journal of Modernization in Engineering Technology and Science, 4(5), 1806-1811. https://www.irjmets.com/uploadedfiles/paper/issue_5_may_2022/23059/final/fin_irjmets1652589657.pdf

[6] Ashraf, F.B., Islam, M.U., Kabir, M.R. et al. YoNet: A Neural Network for Yoga Pose Classification. SN COMPUT. SCI. 4, 198 (2023). https://doi.org/10.1007/s42979-022-01618-8

[7] Malik, I.H., & Malik, I.H. (2017). Importance of Yoga in Daily Life. International Journal of Creative Research Thoughts (IJCRT), 5(1), 142. http://www.ijcrt.org/volume-5/issue-1/IJCRT1134183

[8] Chung, J., Ong, L., & Leow, M. (2022). Comparative Analysis of Skeleton-Based Human Pose Estimation. Future Internet, 14(12), 380. https://doi.org/10.3390/fi14120380

❑ ❑ ❑ ❑ ❑ ❑

AET 024

TalkSpace - Advanced Emotion Recognition and Personality Analysis Platform

Rohan Kadu
BE in Information Technology
Vidyalankar Institute of Technology
rohan.kadu@vit.edu.in

Krishnkant Gangurde
BE in Information Technology
Vidyalankar Institute of Technology
krishnkant.gangurde@vit.edu.in

Vaishnav Rajput
BE in Information Technology
Vidyalankar Institute of Technology
vaishnav.rajput@vit.edu.in

Prof. Samuel Jacob
Assistant Professor Department of Information Technology
Vidyalankar Institute of Technology
samuel.jacob@vit.edu.in

Abstract— **TalkSpace is a revolutionary platform designed to analyze emotions through facial expressions, audio cues, and textual inputs using advanced Deep Learning and Machine Learning models. This project represents a significant advancement in emotional recognition technology, enabling precise identification and understanding of various emotional states in individuals. By leveraging state-of-the-art algorithms, TalkSpace offers an unparalleled level of accuracy in emotion detection, facilitating improved communication and interaction in various domains such as mental health, customer service, and education. With its comprehensive suite of capabilities, including real-time emotion analysis and contextual understanding, TalkSpace aims to revolutionize how emotions are perceived and addressed in diverse contexts, ultimately enhancing human-computer interaction and emotional well-being.**

I. INTRODUCTION

TalkSpace embodies a pioneering endeavor in the realm of emotion recognition and personality trait classification, leveraging an amalgamation of cutting-edge technologies. Through the convergence of text mining, signal processing, and computer vision techniques, TalkSpace offers a comprehensive solution for understanding and interpreting human emotions in diverse contexts.

At its core, TalkSpace utilizes text mining methodologies to delve into textual inputs, extracting valuable insights into individuals' personality traits. By analyzing language patterns and textual cues, the system accurately classifies personality traits, providing profound understandings of behavioral tendencies and characteristics.

Moreover, TalkSpace employs sophisticated signal processing algorithms to decode emotional cues embedded within audio signals. Through precise analysis of voice intonations, speech patterns, and acoustic features, TalkSpace achieves nuanced emotion recognition, enabling the detection of subtle emotional nuances in spoken communication.

In parallel, TalkSpace harnesses the power of computer vision for emotion recognition, analyzing facial expressions to decipher underlying emotions. Leveraging advanced image processing techniques, the system identifies key facial features and dynamics, allowing for real-time assessment of emotional states with remarkable accuracy.

By seamlessly integrating these technologies, TalkSpace transcends traditional approaches to emotion recognition, offering a multifaceted and holistic understanding of human emotions. Whether in mental health assessments, customer service interactions, or educational settings, TalkSpace empowers users to effectively interpret and respond to emotional cues, fostering improved communication and interaction across various domains.

II. PROBLEM STATEMENT

Despite the advancements in technology, understanding and interpreting human emotions and personality traits remain challenging tasks. Traditional methods for emotion recognition and personality trait classification often rely on subjective assessments or manual analyses, leading to inaccuracies and inefficiencies in various domains such as mental health, customer service, and education.

Moreover, existing solutions often focus on single modalities such as text analysis or facial recognition, overlooking the holistic nature of human communication, which involves a combination of verbal, non-verbal, and contextual cues.

This fragmented approach hampers the ability to accurately capture and interpret the complexity of human emotions, limiting the effectiveness of communication and interaction in diverse contexts. Furthermore, the lack of integration between different modalities hinders the development of comprehensive solutions that can provide nuanced insights into human behavior.

Therefore, there is a critical need for a unified framework that leverages multiple modalities, including text mining, signal processing, and computer vision, to enable accurate and holistic emotion recognition and personality trait classification. Such a framework would not only enhance our understanding of human emotions but also facilitate the development of more effective communication strategies, personalized services, and tailored interventions across various domains.

III. RELATED WORK

This References [1] [23] [25] enhances emotion recognition in textual data using modifications like bidirectional processing and dropout regularization. Computational experiments show significant performance improvements, with pre-trained bidirectional LSTMs outperforming traditional models. The proposed sent2affect strategy, utilizing transfer learning from sentiment analysis tasks, further boosts performance.

References [10] [11] [12] [13] provides a thorough examination of CNN-based Facial Emotion Recognition (FER), highlighting performance differences and bottlenecks. We showcased notable enhancements achieved through modern CNNs and ensemble techniques. Looking ahead, our focus will be on addressing remaining challenges, with emphasis on data augmentation strategies tailored for FER and addressing biases in existing datasets like FER2013. Additionally, we aim to explore the development of a more

comprehensive and publicly available FER dataset to advance research in this field.

References [4] [5] [8] [9] presents a concise review of audio-based emotion recognition systems, assessing their performance in terms of classifiers, features, recognition rates, and datasets. Notably, well-designed classifiers have demonstrated high accuracy across various emotional states. Current research emphasizes exploring different features, including Mel-frequency cepstral coefficients (MFCCs), to enhance recognition rates. Additionally, time-distributed CNNs show promise in improving performance. However, challenges with existing datasets hinder accurate evaluation of emotion recognition in audio recordings.

IV. PROPOSED ALGORITHM

Multimodal Emotion Recognition is a relatively new discipline that aims to include text inputs, as well as sound and video. This field has been rising with the development of social networks that gave researchers access to a vast amount of data. Recent studies have been exploring potential metrics to measure the coherence between emotions from the different channels. We are going to explore several categorical targets depending on the input considered. Table 1 gives a summary of all the categorical targets we are evaluating depending on the data type.

Data types	Categorical target
Textual	Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism
Sound	Happy, Sad, Angry, Fearful, Surprise, Neutral and Disgust
Video	Happy, Sad, Angry, Fearful, Surprise, Neutral and Disgust

Table 1: Categorical target depending on the input data type.

Algorithm 1: Text mining for personality trait using NN (CONV + LSTM)

Input: Any emotional text

Output: Openness , Conscientiousness , Extraversion , Agreeableness , Neuroticism.

1. Data Collection: We are using data that was gathered in a study by Pen nebaker and King [1999]. It consists of a total of 2,468 daily writing submissions from 34 psychology students (29 women and 5 men whose ages ranged from 18 to 67 with a mean of 26.4).

2. Preprocessing:

- **Tokenization:** Involves dividing the document into individual words or tokens.
- **Standardization:** Includes using regular expressions to replace formulations such as "can't" with "cannot" and "'ve" with "have".
- **Deletion of Punctuation:** Removes punctuation marks from the tokens.
- **Lowercasing:** Converts all tokens to lowercase.
- **Removal of Stopwords:** Involves removing predefined stopwords like 'a', 'an', etc.
- **Part-of-Speech Tagging:** Assigns part-of-speech tags to the remaining tokens.
- **Lemmatization:** Utilizes part-of-speech tags for more accurate lemmatization of tokens.
- **Padding:** the sequences of tokens of each document to constrain the shape of the input vectors. The input size has been fixed to 300 : all tokens beyond this index are deleted. If the input vector has less than 300 tokens, zeros are added at the beginning of the vector in order to normalize the shape. The dimension of the padded sequence has been determine using the characteristics of our training data. The average number of words in each essay was 652 before any preprocessing. After the standardization of formulations, and the removal of punctuation characters and stopwords, the average number of words dropped to with a standard deviation of . In order to make sure we incorporate in our classification the right number of words without discarding too much information, we set the padding dimension to 300, which is roughly equal to the average length plus two times the standard deviation.

3. Embedding: Each token is replaced by its embedding vector using Google's pre-trained Word2Vec vectors in 300 dimensions (which is the largest dimension available and therefore incorporates the most information), and this embedding is set to be trainable (our training corpus is to small to train our own embedding).

4. Classifier: The neural network architecture combines one-dimensional convolutional neural networks (CNNs) and recurrent neural networks (RNNs). The one-dimensional convolutional layer extracts features from the text data, followed by Long Short-Term Memory (LSTM) cells to leverage the sequential nature of natural language. Unlike regular neural networks, LSTMs progressively accumulate and capture information through sequences, selectively remembering patterns for long durations.

The final model comprises three consecutive blocks, each consisting of a one-dimensional convolution layer, max pooling layer, spatial dropout layer, and batch normalization layer. The number of convolution filters for each block is 128, 256, and 512, respectively, with a kernel size of 8, max pooling size of 2, and dropout rate of 0.3.

Following the three blocks, three LSTM cells with 180 outputs each are stacked. Finally, a fully connected layer of 128 nodes is added before the last classification layer.

Algorithm 2: Signal processing for emotion recognition using Time Distributed CNN

Input: Audio file representing a emotion

Output: Happy , Sad , Angry , Fearful ,Surprise , Neutral , Disgust.

1. Data Collection: The RAVDESS database was used, It contains acted emotions speech of male (672) and female (672) actors (gender balanced) that were asked to pretend six different emotions (happy, sad, angry, disgust, fear, surprise and neutral) at two levels of emotional intensity.

2. Signal Preprocessing:

- Pre-emphasis filter: Amplifies high frequencies to balance the frequency spectrum and prevent numerical issues in Fourier Transform computation.
- Framing: Divides the signal into short-term windows to capture frequency contours over time. Typically, window

sizes range from 20ms to 50ms with 40% to 50% overlap between consecutive windows. Common settings include a frame size of 25ms with a 15ms overlap.

- Hamming: Each frame is multiplied by a Hamming window function to reduce spectral leakage and signal discontinuities, improving clarity. The Hamming function ensures that the beginning and end of frames match up smoothly.
- Discrete Fourier Transform (DFT): Converts the signal from the time domain to the frequency domain, allowing analysis of frequency content. This transformation facilitates the representation and analysis of audio features, which are often defined in the frequency domain.

3. Short-term audio features:

Time-domain features:

- **Energy**: Sum of squares of signal values normalized by frame length.
- **Entropy of energy:** Measures abrupt changes in energy amplitude of an audio signal.
- **Zero Crossing Rate:** Rate of sign changes of an audio signal.
- Frequency-domain features:
- **Spectrogram:** Represents time evolution of frequency content.
- **Log-mel-spectrogram:** Utilizes mel-frequency scale for human auditory system resemblance.
- **Spectral centroid:** Center of gravity of the sound spectrum.
- **Spectral spread:** Second central moment of the sound spectrum.
- **Spectral entropy:** Measures normalized spectral energies.
- **Spectral flux:** Measures spectral changes between successive frames.

4. Model: The Time distributed convolutional neural network integrates hierarchical CNNs with an LSTM-based recurrent neural network to identify sequential patterns in speech signals. Operating directly on log-mel-spectrograms, it applies a rolling window mechanism, processing each segment with a convolutional neural network featuring four Local Feature Learning Blocks (LFLBs). The output is then fed into a recurrent neural network with two LSTM cells to capture long-term dependencies, culminating in a fully connected layer with softmax activation for emotion prediction.

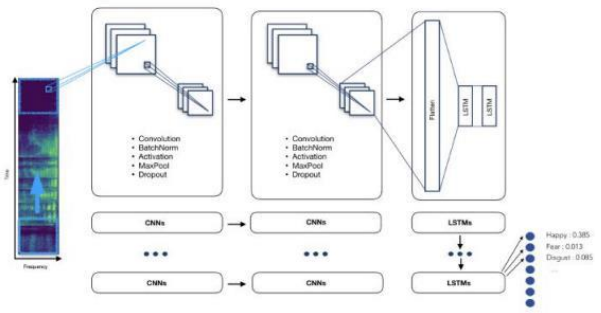


Fig. 1: Time distributed CNN

5. Evaluation: We report the results of our deep learning model, divided the dataset into 80% for training, 15% for validation, and 5% for testing. Early stopping was employed to prevent overfitting, using Stochastic Gradient Descent with decay and momentum as the optimizer and a batch size of 64. Graphs display categorical cross-entropy loss and accuracy for training and validation sets.

6. Improvement: Our model achieved satisfactory results with a prediction recognition rate of approximately 65% for 7-way emotions and 75% for 6-way emotions (surprised removed). To enhance performance further, we plan to explore more sophisticated classifiers such as Hidden Markov Models (HMM) and Convolutional Neural Networks (CNN) in the next phase.

Algorithm 3: Computer vision for emotion recognition using XCeption model

Input: WebCam Video representing a emotion

Output: Happy , Sad , Angry , Fearful ,Surprise , Neutral , Disgust.

1. Data Collection:

Collect a dataset of FER2013 data set containing number of class by emotions such as Happy, Sad, Angry, Fearful, Surprise, Neutral, Disgust.

The train set has 28709 images, the test set has 3589 images. For each image, the data set contains the grayscale color of 2304 pixels (48x48), as well as the emotion associated.

2. Data Preprocessing:

- **Grayscale Conversion:** Frames are converted to grayscale using functions like cv2.cvtColor() to simplify processing and reduce input complexity.
- **Face Detection and Zoom:** Face detection algorithms, possibly using functions like cv2.CascadeClassifier.
- detectMultiScale(), are employed to identify and zoom in on faces within each frame.
- **Multiple Face Management:** Techniques such as iterating through detected faces or selecting the primary face can be implemented to handle multiple faces.
- **Pixel Density Reduction:** Functions like cv2.resize() may be used to reduce pixel density and match the training set's resolution.
- **Image Transformation:** The preprocessed image is transformed using functions like cv2.resize() or cv2.normalize() to align with the model's input format.
- **Emotion Prediction:** After preprocessing, the image is inputted into the model for emotion prediction, utilizing functions relevant to the XCeption Model

3. Model Architecture: The data first goes through the entry flow, then through the middle flow which is repeated eight times, and finally through the exit flow. Note that all Convolution and SeparableConvolution layers are followed by batch normalization.

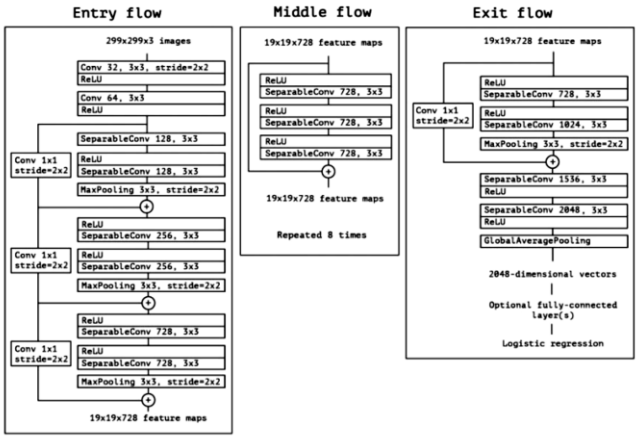


Fig. 2: XCeption Structure

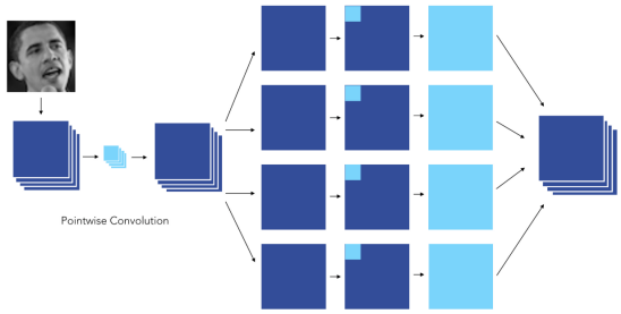


Fig. 3: XCeption

4. Training and Validation: The compiled model is trained on the training dataset using batches of images. During training, the model iteratively adjusts its parameters to minimize the loss function, optimizing its ability to classify images accurately.

To monitor the model's performance and prevent overfitting, it is evaluated on a separate validation dataset after each epoch.

Validation accuracy and loss metrics are computed to assess the model's generalization ability.

5. Hyperparameters -Tuning : Hyperparameters such as learning rate, batch size, and dropout rate may be tuned to optimize the model's performance further. Techniques like learning rate scheduling or early stopping may also be employed to improve training efficiency and prevent overfitting.

6. Deployment: Finally, the trained Xception model can be deployed for inference on new unseen images, where it can classify objects or perform other relevant tasks based on its learned representations.

V. IMPLEMENTATION

A. Text Analysis Using NN And LSTM:

Text analysis involves deriving insights from textual data. Neural networks, especially LSTM, are effective for this task. Preprocessing involves tokenization and encoding text into numerical representations. A neural network architecture, typically comprising LSTM layers, is trained on labeled data. Once trained, the model can be used for tasks like sentiment analysis and text classification. This approach offers a versatile solution for deriving insights from text data.

B. Video Analysis Using XCeption:

Xception, a powerful convolutional neural network, is utilized for video analysis. Each frame of the video is treated as an image and passed through the Xception model. The pre-trained Xception model extracts high-level features for tasks such as action recognition or object detection. This approach enables accurate and robust video analysis across various applications.

C. Audio Analysis Using Time Distributed CNN:

Utilizing a Time Distributed CNN, we analyze audio signals directly from log-mel-spectrograms. This model combines CNNs with LSTM networks to capture temporal patterns and dependencies for accurate emotion prediction.

D. Web Application Development Using Flask:

To provide users with seamless access to the TalkSpace system, a web application will be developed using the Flask framework. This web application will feature a user-friendly interface comprising views and templates for uploading input file, displaying results. Compatibility and responsiveness across different devices and browsers will be ensured to accommodate a diverse user base and enhance overall user experience.

VI. RESULTS

TalkSpace project presents a holistic solution for emotion recognition and personality trait classification. Leveraging advanced deep learning and machine learning techniques, including signal processing for emotion recognition, computer vision for facial emotion analysis, and text mining for personality trait classification, TalkSpace achieves accurate and real-time assessments of individuals' emotional states and personality traits. By integrating these components, TalkSpace facilitates enhanced understanding of human behavior, fosters personalized interactions, and opens avenues for applications in mental health, customer service, and education. Overall, TalkSpace revolutionizes the landscape of emotion analysis and personality assessment, paving the way for more insightful and effective human-computer interactions.

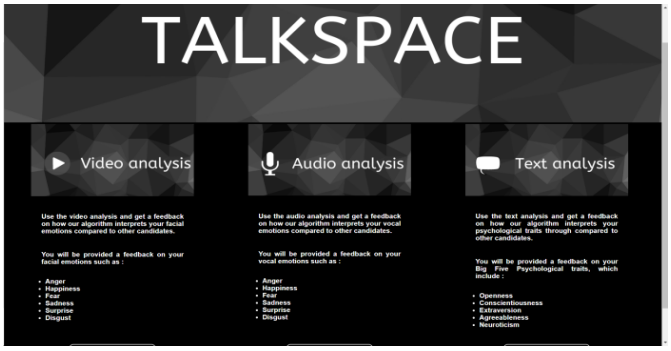
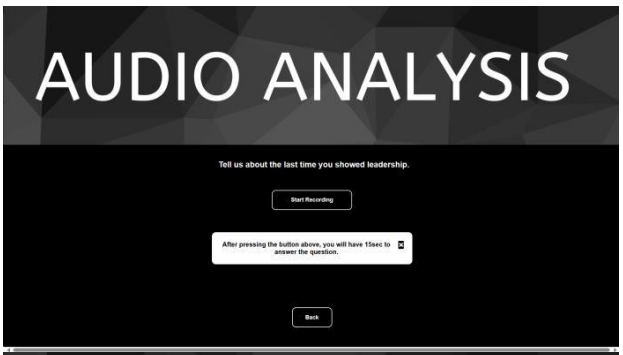


Fig. 4: TalkSpace Homepage

As seen in Fig 1, A page is dedicated to each communication channel (audio, video, text) and allows the user to be evaluated. A typical interview question is asked on each page, for instance:" Tell us about the last time you showed leadership". The audio/video extract (recorded via computer microphones/webcam) or text block can be retrieved once saved and processed by our algorithms (in the case of the text channel the user can also upload a .pdf document that will be parsed by our tool)



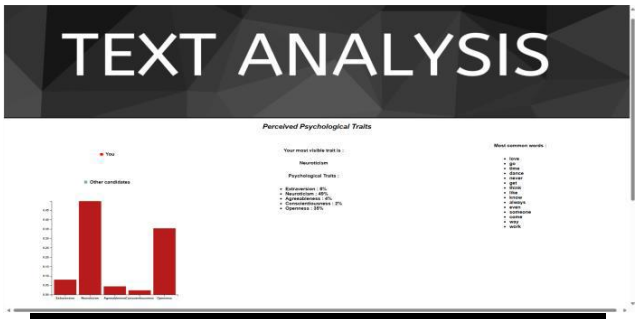
Fig. 5: Text Analysis

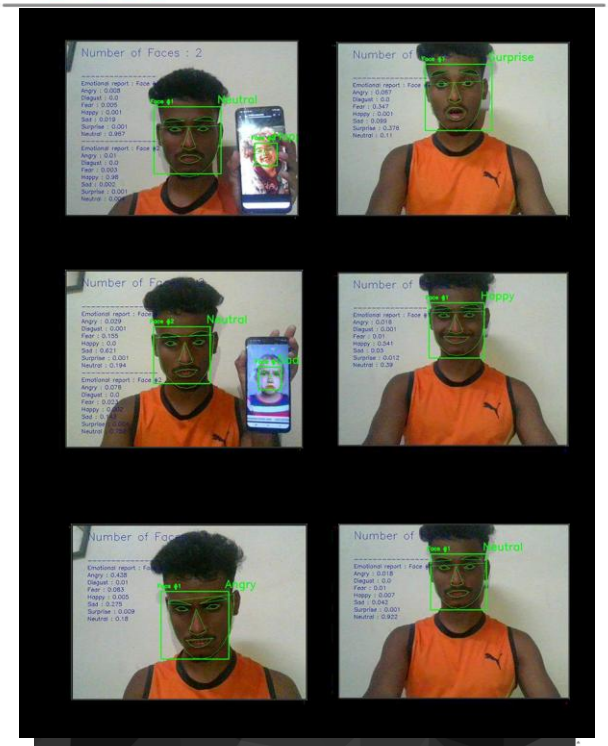


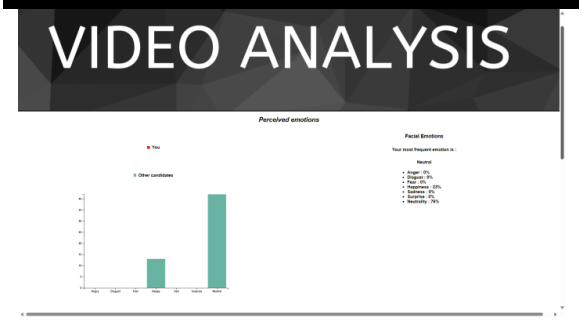




The text and video/audio summaries are slightly different : for the text interview summary, not only we chose to display the percentage score of identified personality traits for both the user and the other candidates, but also the most frequently used word in the answer. For the video and audio interview summaries, we displayed the perceived emotions scores of the user and the other candidates. Following are the summary pages for both the text and video interviews.







VII. CONCLUSION

TalkSpace stands at the forefront of emotion recognition and personality trait classification. By integrating cutting-edge technologies such as deep learning and signal processing, it offers a transformative approach to understanding human emotions across text, audio, and video modalities. This innovative platform has the potential to revolutionize various domains, from mental health diagnostics to customer sentiment analysis, by providing invaluable insights into human behaviour and emotional states.

With its versatility and accuracy, TalkSpace paves the way for more personalized interactions, tailored interventions, and informed decision-making processes. As we continue to refine and expand its capabilities, TalkSpace holds the promise of fostering greater empathy, understanding, and connection in our increasingly digital world. Through its seamless integration of technology and human experiences, TalkSpace is poised to make a profound impact on how we perceive,

analyze, and respond to emotions in today's fast-paced and interconnected society.

ACKNOWLEDGMENT

I extend my sincere thanks to Professor Samuel Jacob for his invaluable guidance and expertise, which have been instrumental in shaping the development of TalkSpace - An Innovative Emotion Recognition and Personality Trait Classification Platform. Professor Jacob's insights and mentorship have greatly contributed to the success of this project.

Furthermore, I would like to express my gratitude to Vidyalankar Institute of Technology for providing the necessary infrastructure and resources that made the development of TalkSpace possible. The support and facilities provided by Vidyalankar Institute of Technology have been essential in facilitating the implementation and execution of this project.

REFERENCES

[1] B.Kratzwalda, S.Ilie´, M.Kraus, S.Feuerriegel, H.Prendinger. Deep learning for affective computing: text-based emotion recognition in decision support, Sep. 2018. https://arxiv.org/pdf/1803.06397.pdf

[2] N.Majumder, S.Poria, A.Gelbukh, E.Cambria. Deep Learning-Based Document Modeling for Personality Detection from Text, 2107. http://sentic.net/deep-learning-based-personality-detection.pdf

[3] The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS), https://zenodo.org/record/1188976/?f=3.XAcEs5NKhQK

[4] B.Basharirad, and M.Moradhaseli. Speech emotion recognition methods: A literature review. AIP Conference Proceedings 2017. https://aip.scitation.org/doi/pdf/10.1063/1.5005438

[5] L.Chen, M.Mao, Y.Xue and L.L.Cheng. Speech emotion recognition: Features and classification models. Digit. Signal Process, vol 22 Dec. 2012.

[6] T.Vogt, E.Andre´ and J.Wagner. Automatic Recognition of Emotions from Speech: A Review of the Literature and Recommendations for Practical Realisation. Affect and Emotion in Human-Computer Interaction, 2008.

[7] T.Vogt and E.Andre´. Improving Automatic Emotion Recognition from Speech via Gender Differentiation. Language Resources and Evaluation Conference, 2006.

[8] T.Giannakopoulos. pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. Dec. 2015 https://doi.org/10.1371/journal.pone.

[9] T.Giannakopoulos. pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. Dec. 2015 https://doi.org/10.1371/journal.pone.0144610.

[10] The Facial Emotion Recognition Challenge from Kaggle, https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/da

[11] C.Pramerdorfer, and M.Kampel. Facial Expression Recognition using Convolutional Neural Networks: State of the Art. Computer Vision Lab, TU Wien. https://arxiv.org/pdf/1612.02903.pdf

[12] OpenCV open source library for image feature extraction, https://opencv.org/2013

[13] End-to-End Multimodal Emotion Recognition using Deep Neural Networks, https://arxiv.org/pdf/1704.08619.pdf

[14] Agrawal, A., & An, A. (2012). Unsupervised emotion detection from text using semantic and syntactic relations. In International Conference on Web Intelligence and Intelligent Agent Technology.

[15] Al-Hajjar, D., & Syed, A. Z. (2015). Applying sentiment and emotion analysis on brand tweets for digital marketing. In Applied Electrical Engineering and Computing Technologies. IEEE.

[16] Alm, C. O., Roth, D., & Sproat, R. (2005). Emotions from text. In Human Language Technology and Empirical Methods in Natural Language Processing (pp. 579–586).

[17] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee, Y. Zhou, C. Ramaiah, F. Feng, R. Li, X. Wang, D. Athanasakis, J. Shawe-Taylor, M. Milakov, J. Park, R. Ionescu, M. Popescu, C. Grozea, J. Bergstra, J. Xie, L. Romaszko, B. Xu, Z. Chuang, and Y. Bengio, “Challenges in representation learning: A report on three machine learning contests,” Neural Networks, vol. 64, pp. 59–63, 2015..

[18] E. Sariyanidi, H. Gunes, and A. Cavallaro, “Automatic analysis of facial affect: A survey of registration, representation, and recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 37, no. 6, pp. 1113–1133, 2015.

[19] M. V. B. Martinez, “Advances, Challenges, and Opportunities in Automatic Facial Expression Recognition,” in Advances in Face Detection and Facial Image Analysis. Springer, 2016, pp. 63–100.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in Advances in Neural Information Processing Systems (NIPS), 2012, pp. 1097–1105.

[21] K. He, X. Zhang, haoqing Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” CoRR, vol. 1512, 2015.

[22] A. Dhall, O. Ramana Murthy, R. Goecke, J. Joshi, and T. Gedeon, “Video and Image Based Emotion Recognition Challenges in the Wild: EmotiW 2015,” in ACM International Conference on Multimodal Interaction (ICMI), 2015, pp. 423–426.

[23] Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., Jha, S., Agrawal, M., Dziurzynski, L. A., Sap, M., Weeg, C., Larson, E. E., Ungar, L. H., & Seligman, M. E. P. (2015). Psychological language on Twitter predicts county-level heart disease mortality. Psychological Science, 26 , 159–169.

[24] Danisman, T., & Alpkocak, A. (2008). Feeler: Emotion classification of text using vector space model. In Communication, Interaction and Social Intelligence. volume 1.

[25] Chitturi, R., Raghunathan, R., & Mahajan, V. (2007). Form versus function: How the intensities of specific emotions evoked in functional versus hedonic trade-offs mediate product preferences. Journal of Marketing Research, 44 , 702–714.

[26] Ekman, P., Friesen, W. V., O’Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., & Ricci-Bitti, P. E. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. Journal of Personality and Social Psychology, 53 , 712–717.

❑ ❑ ❑ ❑ ❑ ❑

AET 025

Mental Health Assistant using Sentiment Analysis Through NLP

Sakshi Talele B.E., Department of INFT
Vidyalankar Institute of Technology
Mumbai, India
sakshi.talele@vit.edu.in

Akanksha Amdekar B.E.,
Department of INFT Vidyalankar Institute of Technology
Mumbai, India
akanksha.amdekar@vit.edu.in

Pranjal Kashid B.E.,
Department of INFT Vidyalankar Institute of Technology
Mumbai, India
pranjal.kashid@vit.edu.in

Dr. Vidya Chitre Prof.,
Department of INFT Vidyalankar Institute of Technology
Mumbai, India
vidya.chitre@vit.edu.in

***Abstract*— Given the increasing prevalence of mental health problems and social disruption, quality mental health services are an important issue in today's society. This paper provides a comprehensive understanding of the mental health assistant application proposed. By combining insights from psychology, healthcare, and technology, we identify problems arising from existing systems, including barriers to access, limited affordability, and disparities in supply. The system is an Android platform that allows people seeking mental health care to record and use features such as self-diagnosis, taking mental health tests for various conditions such as Depression, Bipolar Disorder, OCD and Anxiety, analyzing their daily mood , and sharing one-on-one conversations with mental health professionals. This research contributes by providing opportunities for improving mental health services in the future.**

***Keywords*— Mental Health, Android, Personality Tests, One-on-one Conversation**

I. INTRODUCTION

Mental illness has become a global health problem in recent years; The World Health Organization estimates that approximately one in four people will experience a mental illness at some point in their lives, and an estimated 970 million people live with a mental illness or psychological problems. Our emotional, psychological, and social well-being are all parts of our mental health. It influences our thoughts, emotions, behaviour, how we respond to stress, interact with others, and make good decisions. Despite growing recognition of the importance of mental health, access to mental health services remains a challenge for millions of people around the world. Issues such as stigma, limited resources, and lack of training for mental health professionals contribute to disparities in access to care, increasing the pressure on individuals, families, and communities. Every period of life, from childhood and adolescence to maturity, is vital for mental health. Poor mental health and mental illness are not the same things, despite the fact that the phrases are sometimes used synonymously. Even if they may not have a mental disorder, a person can have poor mental health. Poor mental health can serve as a breeding ground for conditions like depression, OCD, anxiety, and bipolar disorder, where unresolved stressors, genetic predispositions, and neurochemical imbalances converge to manifest debilitating symptoms. Mental health disorders such as depression, obsessive-compulsive disorder (OCD), anxiety, and bipolar disorder significantly impact daily life, affecting individuals' cognitive, emotional, and behavioural functioning. These conditions not only undermine productivity and interpersonal relationships but also pose substantial challenges in maintaining overall well-being and quality of life. Effective intervention strategies, including psychotherapy, medication, and lifestyle adjustments, are essential in mitigating symptoms and facilitating functional recovery for individuals grappling with these mental health challenges.

II. LITERATURE SURVEY

Paper Name: Mental Health Mobile Application with Diagnosis, Sentiment Analysis and Chatbot [1]

Tanvi Gadgil et al,in,2022, [1] This study looks at a variety of mental health-related app-related issues that appeared between 2016 and 2020 in the Google Play Store. The data was scraped using a variety of programs and programming languages. It provides a variety of strategies for enhancing mental health. One of the methods is to relax, manage stress, track symptoms, listen to calming music, journal, connect with mental health resources, get support from others, practice meditation, and so on.

Paper Name: User Engagement in Mental Health Apps: a Review of Measurement, Reporting, and Validity [2]

Michelle M. Ng et al,in, 2018, [4] The fact that few patients use mobile mental health apps regularly and consistently, despite the potential benefits, indicates engagement issues. This study investigated and analysed the measurement and reporting of user engagement indicators in mental health apps. The fact that every study claims to have strong user engagement indicators despite these discrepancies points to the necessity for the sector to concentrate on engagement through the development of reporting standards and more meticulous evaluation of claims.

Paper Name: Mobile Mental Health: A Review of Applications for Depression Assistance [3]

Wellington S. Silva et al,in, 2019,[3] The purpose of this study is to define, examine, and classify the condition of mobile applications for depression as it stands right now. To do this, we thoroughly examined depression treatment applications. The findings of this study demonstrated that apps are being used for an increasing number of purposes, including chatbots, online therapy, mood monitoring, self-help, educational tools, and testing.

III. PROPOSED ALGORITHM

A. Proposed System

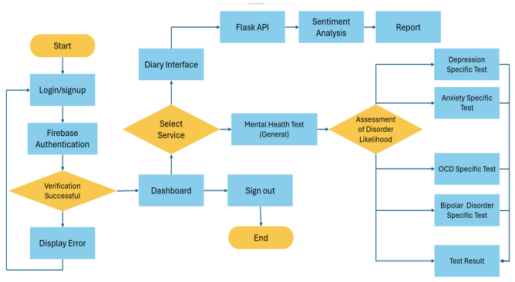


Figure 1. Proposed Algorithm

B. System Overview

The application includes features such as psychological diagnosis, communication with a psychologist, and a Daily Diary that can record the user's thoughts and feelings. Psychological tests provide users with detailed information about their mental state based on their performance on the test. Machine learning and NLP sentiment analysis will be used on logged emotions to analyse the user's mood. Chat allows users to communicate with psychologists. Apart from these features, the application also provides authentication, profile, settings and all other necessary features. Firebase Cloud Services is used for authentication, instant databases, and Firebase Storage.

C. Evaluation

The general health questionnaire (GHQ-12), comprising 12 questions, serves as the initial round of assessment, and this is the method employed to calculate the results of this round:

For depression, higher scores on questions 1, 2, 6, 8, and 11 may indicate a greater likelihood of depressive symptoms. Anxiety symptoms may be suggested by higher scores on questions 3, 4, 7, and 9, with a cutoff of 7 or higher indicating a higher likelihood. For OCD, a higher score on question 10, addressing persistent unwanted thoughts or repetitive behaviors, may suggest symptoms. Bipolar disorder symptoms may be indicated by higher scores on questions 5, 8, and 12.

Based on the results of the preceding initial round, a subsequent round is conducted to categorize whether the individual exhibits symptoms indicative of depression, anxiety, OCD, or bipolar disorder, determined by the evaluation process:

1) Becks Depression Inventory: The evaluation of scores on the Beck Depression Inventory (BDI) can provide valuable insights into the severity of depressive symptoms experienced by individuals.
2) Hamilton Anxiety Scale: In assessing responses to the Hamilton Anxiety Scale (HAM-A), each item's score, ranging from 0 (not present) to 4 (severe), contributes to a comprehensive evaluation of anxiety symptom severity.
3) Mood Disorder Questionnaire: In evaluating responses to the Mood Disorder Questionnaire (MDQ), specific criteria suggest a higher likelihood of bipolar disorder symptoms.
4) Obsessive Compulsive Inventory: In evaluating responses to the Obsessive-Compulsive Inventory (OCI), particularly considering a total score of 21 or above, there's a potential indication of the presence of obsessive-compulsive disorder (OCD).

IV. IMPLEMENTATION

A. Main interface

The main interface of the Android application can help users connect to all functions of the application. The home page has Tests, Daily Logs, Search and Chat tabs to access relevant features. also has a menu button that opens a navigation drawer where the user can navigate to the entire application. The navigation drawer contains fields corresponding to each page of the application. The user can use the Navigation drawer to access profile, application settings, and logout options.

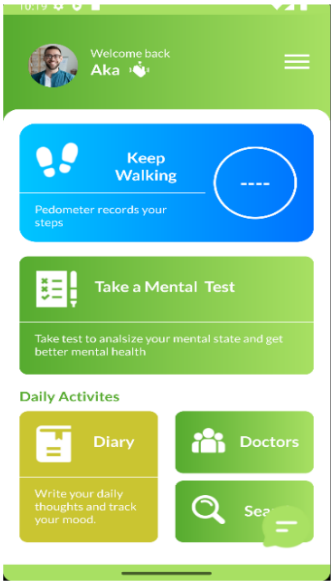


Figure 2. Home Page

B. Perform diagnostic tasks

The testing algorithm presents general twelve questions and based on the result of that test, the user may be directed to attempt the test of Depression, Bipolar, OCD or Anxiety based on the result. Test Interface shows a question and four multiple choice options corresponding to the question. At the end of the test, the application stores the results and recommends the task to the user.



Figure 3. Test

C. Test Result

The assessment process comprises two phases, starting with the common test questionnaire. Upon completion, the results are displayed, as illustrated in Figure 4. Following this initial assessment, if the user exhibits signs of any mental disorder, the system offers an additional test. This system encompasses four distinct tests: anxiety, depression, OCD and bipolar disord. Based on the initial test results, the system recommends one of these tests to the user.

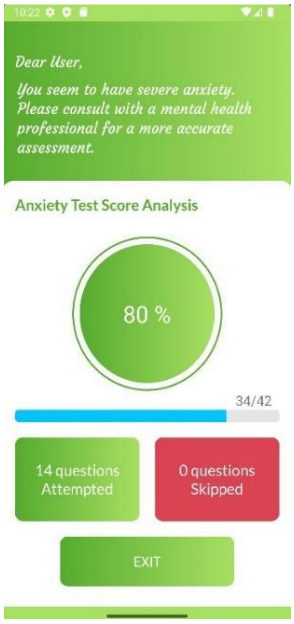


Figure 4. Test Result

D. Daily Thought Diary

The diary function in the android application allows the user to record his thoughts. Text messages also include options like copy to notepad, clear, emoticons and more. The calendar shows the current date and day in month-day format. When the user clicks the Save button, a sentiment analysis of the user's opinion will be performed and the predicted sentiment will be stored along with the recommendation in the Firebase database regarding the specific date data. After pressing the "Save" button the user will see the betting results and operating instructions. This helps the user examine their way of thinking, track their mood throughout the month, and improve themselves.

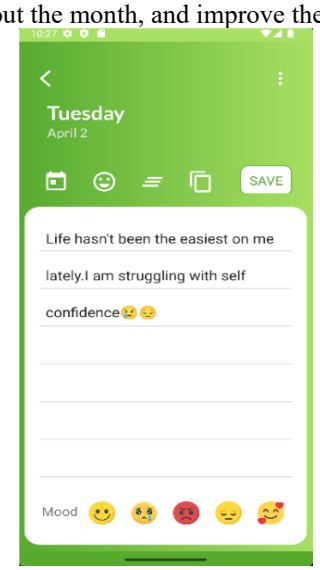


Figure 5. Daily Diary

E. Chat section

The Chat section allows users to chat and communicate with psychologists and seek their help. The interactive session is shown in figure 6.

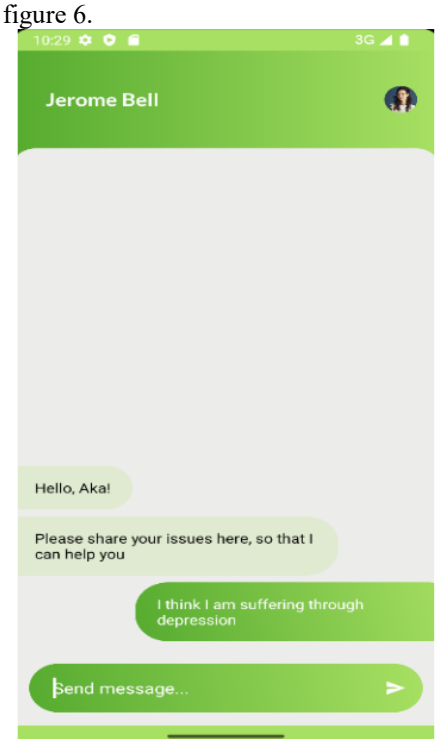


Figure 9. Accuracy

VI. CONCLUSION

This mental health service Android app helps users get guidance from experts and provides various mental health services. Users do not need to visit a psychiatrist or wait in long queues for patients, everything is available on one platform. Users can easily convey their requests to psychologists and share their problems. This simple communication will save users a lot of time and provide anonymous reporting to the community. People who are shy about mental problems can also find and communicate with doctors through the application. The application provides a simple interface so that anyone can easily enter the application. The app also allows users to take mental health tests and gives the user advice for recovery. Taking mental health tests and examining their results will help users improve their performance. The app uses sentiment analysis to predict users' sentiments regarding daily dairy consumption. The daily diary helps the user record their thoughts and track their behavior. The use of this app will lead to an overall improvement in the user's mental health without any trouble. This simple service will allow users to take better care of their health rather than neglecting it.

VII. FUTURE SCOPE

A mental health android app's future potential is huge, with the possibility to extend and improve on current capabilities to better serve the needs of users. Wearable devices: such as smartwatches or fitness trackers, might be coupled with the app to provide additional data points on users' mental and physical health. Data on sleep patterns, exercise routines, and stress levels, for example, could assist users and mental health practitioners in better understanding and managing mental health symptoms. Personalized recommendations

using machine learning: Machine learning algorithms could be used to analyze user data and deliver personalized suggestions for self-help resources, therapies, or mental health experts. This could assist users in locating the most effective and acceptable.

REFERENCES

[1] Abid Hassan, M. D. Iftekhar Ali, Rifat Ahammed, Sami Bourouis Mohammad Monirujjaman Khan, "Development of NLP-Integrated Intelligent Web System for E-Mental Health" December 2021.

[2] Nidhi Vashistha, Dharmendra Kumar, Ankit Bairwa, "Development of Android Application for Mental Health of The Students for Betterment" April 2018.

[3] Md. Aminul Islam, Naziat Choudhury, "Mobile apps for Mental Health: a content analysis" August 2020.

[4] Arfan Ahmed, Nashva Ali, Anna Giannicchi, Alaa A Abd-alrazaq, Mohamed Ali, "Mobile applications for mental health self-care: A scoping review" October 2021.

[5] Devarapalli, S. V. Siddhardh Kumar; Kallakuri, Sudha; Salam, Abdul; Maulik, Pallab K., "Mental health research on scheduled tribes in India" December 2020.

[6] Kit Huckvale, Jennifer Nicholas, John Torous, Mark E Laresen, "Smartphone apps for the treatment of mental health conditions: status and considerations" May 2020.

[7] Jamie M Marshall, Debra A Dunstan, and Warren Bartik, "The use and effectiveness of mobile mental health apps for treating anxiety and depression" September 2019.

[8] Ariel Teles, Ivan Rodrigues, Davi Viana, Francisco Silva, Luciano Coutinho, Markus Endler, Ricardo Rabêlo,"Mobile Mental Health: A Review of Applications for Depression Assistance" June 2019.

[9] Michelle M Ng, Joseph Firth, Mia Minen, and John Torous, "User Engagement in Mental Health Apps: a Review of Measurement, Reporting, and Validity"March 27.

[10] Tania Lecomte, Stéphane Potvin, Marc Corbière, Stéphane Guay, Crystal Samson, Briana Cloutier, Audrey Francoeur, Antoine Pennou, Yasser Khazaal, "Mobile Apps for Mental Health Issues: Meta Review of Meta-Analyses" May 29.

❑ ❑ ❑ ❑ ❑ ❑

AET 026

GeoCause – Geolocation Event Finder

Saee Majlekar
B.E., Department of Information Technology
Vidyalankar Institute of Technology Mumbai, India
saee.majlekar@vit.edu.in

Chinmayi Deshmane
B.E., Department of Information Technology
Vidyalankar Institute of Technology Mumbai, India
chinmayi.deshmane@vit.edu.in

Jidnya Kanekar
B.E., Department of Information Technology
Vidyalankar Institute of Technology Mumbai, India
jidnya.kanekar@vit.edu.in

Shruti Agrawal Asst. Prof.,
Department of INFT VIT, Mumbai, India
shruti.agrawal@vit.edu.in

***Abstract*— This paper proposes a system for locating events that are around the user by using the Global Positioning System. Globally, GPS has transformed location-based services and navigation, radically changing our understanding of and use for geographic positioning. These days, GPS is everywhere. It's been incorporated into cars, wearable technology, cellphones, and a host of other products. To identify new trends of substantial correlations between geographic regions and events, the proposed system examines the correlation between user locations and event locations in our research.**

***Keywords*– classification; hybrid filtering algorithm; data mining; event recommendation; recommendation system**

I. INTRODUCTION

In today's dynamic society, people are constantly in pursuit of more effective ways to discover and engage with local events and activities that cater to their preferences and schedules. The evolution of technology has significantly altered the lifestyle of individuals, consequently transforming the approach to event discovery. In the past, individuals relied on traditional methods such as newspapers, radio broadcasts, or physically visiting event venues to stay informed about upcoming events. However, with the advent of technology, there has been a paradigm shift towards utilizing digital platforms for event searches. Among these technological advancements, the geolocation-based event finder application emerges as a groundbreaking tool, revolutionizing event exploration by leveraging the capabilities of MapmyIndia APIs. [1]

At the heart of this innovative platform is the integration of MapMyIndia APIs. MapmyIndia APIs offer a comprehensive suite of location-based services, providing developers with access to accurate mapping, navigation, and geocoding functionalities. Leveraging these APIs empowers applications to deliver seamless location-aware experiences, enhancing user engagement and convenience.[1] By exploring the features of MapmyIndia, users can easily explore nearby events based on their current location, eliminating the need for manual searches and providing real-time updates on events happening in their vicinity.

Furthermore, the application goes beyond traditional event discovery by incorporating additional functionalities to enhance the user experience. Alongside event recommendations, users are empowered to locate nearby cafes for refreshments and recharge, ensuring a holistic outing experience. Additionally, the application addresses the perennial challenge of parking availability by providing information on nearby parking spaces, enabling users to plan their journeys with ease.

In a unique twist, the application introduces a feature allowing event organizers to create and manage events through a dedicated server login. This functionality not only ensures the seamless organization of events but also facilitates real-time updates and visibility to users through their respective accounts. By bridging the gap between event organizers and attendees, the application fosters a vibrant community of event enthusiasts.

The paper is structured as follows. The introduction of the paper is mentioned in Section I. Survey of related works is mentioned in Section II. Section III explains the proposed algorithm for anomaly detection. In Section IV, the implementation process is discussed. The results acquired are mentioned in Section V, and finally, conclusions are drawn in Section VI.

II. RELATED WORK

In the realm of recommendation systems, a burgeoning field, the integration of clustering and classification methods into hybrid models has demonstrated promise for enhancing recommendation accuracy and relevance. This approach is particularly relevant in geolocation-based event finder systems, where user preferences often exhibit spatial patterns. In related work, scholars have explored various strategies to fuse clustering and classification techniques to create robust recommendation frameworks tailored to users' diverse preferences and geographical contexts.

Hybrid recommendation algorithms combine multiple recommendation techniques, such as collaborative filtering, content-based filtering, and knowledge-based approaches, to overcome the limitations of individual methods and provide more accurate and diverse recommendations. By leveraging the strengths of different techniques, hybrid algorithms can effectively address various challenges in recommendation systems, such as the cold start problem, sparsity of data, and user diversity. These algorithms integrate user preferences, item features, and contextual information to generate personalized recommendations that cater to the specific needs and preferences of users. Hybrid recommendation systems have gained popularity in diverse domains, including e-commerce, social networking, and content streaming platforms, due to their ability to offer robust and tailored recommendations that enhance user satisfaction and engagement.

Furthermore, advancements in deep learning have inspired the development of hybrid recommendation models with neural network architectures capable of jointly learning representations from clustering structures and classification tasks. These models can automatically extract intricate

patterns from large-scale geolocation event data, enabling more effective recommendation generation. Moreover, research efforts have focused on incorporating temporal dynamics and contextual information into hybrid recommendation systems to adaptively adjust recommendations over time and across different geographical regions, enhancing their robustness and responsiveness to evolving user needs.

Overall, the integration of clustering and classification methods in hybrid recommendation systems for geolocation event finders represents a promising direction in recommender system research. By leveraging spatial patterns, user preferences, and contextual information, these hybrid models can deliver personalized and context-aware recommendations, thereby enhancing user satisfaction and engagement with event discovery platforms. As the field continues to evolve, further research into novel algorithmic techniques and real-world deployment scenarios is essential to unlock the full potential of hybrid recommendation systems in geolocation-based applications.

III. PROPOSED ALGORITHM

A. SECTION 1: Algorithm

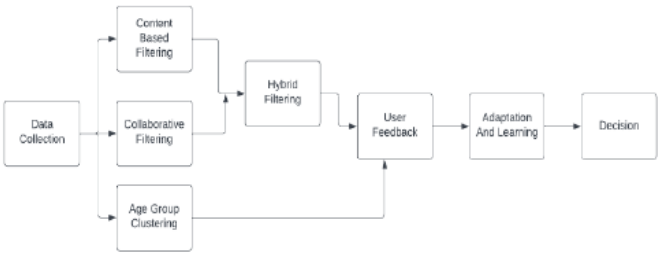


Fig. 2. Algorithm implemented overview

1) **Data Collection:** Data collection from a PostgreSQL database involves users initiating requests, which prompts the generation of SQL queries for retrieving relevant information from the designated tables. The established connection to the database allows for the efficient execution of these queries, enabling the application to dynamically access and utilize the requested data.

2) **Content Based Filtering:** Content-Based Filtering is a recommendation system technique that suggests items, such as events, by analyzing their intrinsic features. In the context of event recommendations, it evaluates properties like type, theme, location, or date to provide personalized suggestions based on a user's historical preferences. This method enhances relevance by recommending items with similar content characteristics to those the user has shown interest in before.

3) **Collaborative Filtering:** Collaborative Filtering recommends items, like events, by identifying and analyzing the preferences and behaviors of users with similar tastes. This approach suggests events based on patterns observed in the interactions of users who share comparable preferences, enhancing the relevance of recommendations for a given user. By leveraging collective user wisdom, Collaborative Filtering contributes to a more personalized and engaging user experience.

4) **Age-group clustering:** Along with hybrid filtering our system is incorporating age-group clustering to enhance the accuracy of recommendations. Firstly, it'll categorize users into distinct age groups, creating clusters. This allows us to understand the preferences and interests that tend to align within specific ranges. Different age groups often have unique tastes and preferences. If a user is in age group 18-24 then he'll share preferences with others in the cluster. This will ensure that the recommendations are more personalized.

5) **User Feedback:** User feedback encompasses opinions and reactions shared by users about their experiences, particularly in response to suggested events in a recommendation system. This valuable input aids in refining the recommendation algorithm, adapting to changing user interests, and improving overall user experience.

6) **Adaptation And Learning:** Adaptation and learning in recommendation systems involve refining algorithms based on user interactions and feedback. This iterative process ensures the system becomes more accurate and aligns better with evolving user preferences, providing increasingly personalized suggestions over time.

7) **Decision:** In the context of recommendation systems, the decision stage typically refers to the user's action or choice based on the presented suggestions. After receiving recommendations, users make decisions on whether to explore more options, attend an event, or take another course of action. The decision stage marks the point where users interact with and act upon the provided recommendations, influencing the direction of their user journey within the system.

IV. IMPLEMENTATION

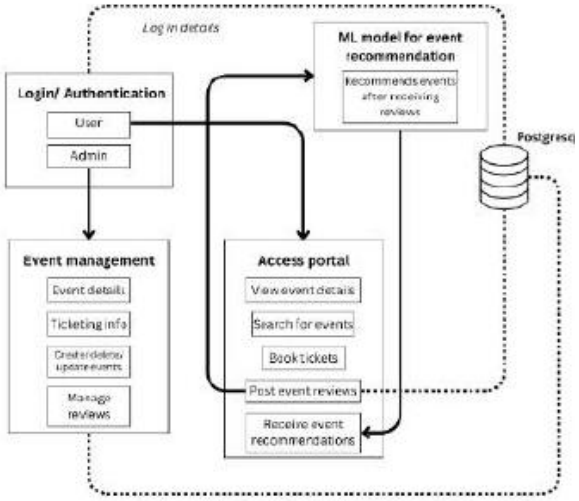


Fig. 3. Implementation process flow

Fig. 3. demonstrates the implementation process flow of the proposed system, and each step is explained below:

A. Login / Authentication:

1) **User Authentication:**

The user authentication system is designed to provide a seamless yet secure experience for individuals accessing events on the platform. Through a combination of

password encryption, multi-factor authentication, and secure session handling, we ensure that users can confidently explore and engage with the diverse events hosted by administrators. This section delves into the technical aspects of user authentication, emphasizing the importance of balancing security measures with user convenience.

2) Admin Authentication:

For administrators hosting events, a robust authentication system is crucial to prevent unauthorized access and maintain control over event management functionalities. This paper discusses the tailored authentication mechanisms implemented for administrators, considering factors such as event creation, modification, and deletion. By incorporating secure protocols and access controls, the system safeguards the integrity of hosted events while allowing administrators the freedom to contribute to the app's dynamic event landscape.

B. Event Management

1) Event details:

The detailed examination of event management encompasses the creation, modification, and deletion of events by administrators. This section outlines the user-friendly interfaces designed to facilitate administrators in inputting comprehensive event details, ensuring accuracy and relevance. Through geolocation technology, events are seamlessly integrated into the app, enhancing user accessibility and fostering a vibrant event ecosystem.

2) Ticketing Information:

Efficient ticketing processes are vital to the success of events hosted on the platform. This paper delves into the ticketing system's design, focusing on user-friendly interfaces for both administrators and event attendees. By providing a secure and streamlined ticketing experience, the app seeks to optimize user participation and revenue generation while offering event hosts valuable insights into attendee engagement.

3) User Reviews Management:

User reviews contribute significantly to the success of events and the overall credibility of the platform. This section explores the mechanisms for managing and moderating user reviews, emphasizing the importance of fostering a positive and constructive user experience. By implementing robust review management features, the app aims to build a community-driven platform that encourages meaningful interactions and informed event participation.

C. Geolocation services

1) Bing Maps and OpenStreetMap (OSM) Bing Maps is a mapping platform developed by Microsoft that provides various mapping and geolocation services. OpenStreetMap (OSM), on the other hand, is a crowdsourced mapping project that allows anyone to update and edit online maps. Here is some information about Bing Maps and its integration with OpenStreetMap:

- Bing Maps has incorporated OpenStreetMap as one of its map styles, allowing users to view maps based on OSM data.
- Users can perform searches and interact with the OSM map within the Bing Maps experience. The OSM map attribution remains independent of other map types in Bing Maps.
- Bing Maps uses the Mapnik build of OpenStreetMap for its OSM map layer.
- Microsoft's intent for incorporating OSM data into Bing Maps is to provide an additional corroborating source of ground truth for validating, updating, and creating contributor edits to OSM.
- Bing Maps also offers Streetside imagery, which can be used as an additional source of ground truth for OSM editing.

D. Access Portal

1) Viewing Event Details: Efficient access to comprehensive event details is fundamental to user engagement. This section delves into the design and usability of the access portal, emphasizing user-friendly interfaces that allow attendees to easily explore event specifics, including time, location, and genre. Geolocation technology enhances the viewing experience, providing users with valuable context about the events happening in their proximity.

2) Searching for Events: Effective search functionalities are crucial for users to discover events aligned with their interests. The paper explores the incorporation of advanced search features within the access portal, enabling users to filter events based on genres, locations, and dates. The goal is to empower users with a robust tool for efficiently navigating the diverse array of events hosted on the platform.

3) Booking Tickets: The ticketing process is a critical component of user engagement. This section examines the ticket booking features integrated into the access portal, focusing on simplicity, security, and user feedback. By providing a seamless and secure ticketing experience, the app aims to encourage increased event participation and enhance the overall user experience.

4) Posting Event Reviews: User-generated content, particularly event reviews, contributes significantly to the platform's credibility. The paper explores the mechanisms for users to post and manage event reviews through the access portal. A user-friendly review interface aims to foster a community-driven environment where attendees can share valuable insights and recommendations.

5) Receiving Event Recommendations: Personalized event recommendations enhance user engagement and satisfaction. This section investigates the algorithms and features embedded in the access portal to provide users with tailored event suggestions based on their preferences, attendance history, and geographic location. The goal is to create a dynamic and personalized user experience.

a) User Side Functionalities: In developing our hybrid recommender system, we begin with meticulous data collection and preprocessing, ensuring high-quality data. This leads to the creation of a User-Item Matrix, a foundation for collaborative filtering, which utilizes historical user interactions for personalized

recommendations. The final phase involves seamlessly integrating collaborative filtering recommendations with content-based suggestions from k-means clustering, striking a balance between user preferences and item characteristics. This strategic approach refines recommendations to optimally cater to diverse user preferences within the system.

b) **Admin Side Functionalities:** On the admin side, our hybrid recommender system employs two essential functionalities. First, item clustering utilizes k-means clustering to group similar items based on shared features, enhancing pattern identification. Second, cluster assignment assigns each item to a specific cluster, streamlining item organization within the system. Together, these functionalities efficiently contribute to item grouping and cluster-based representation, optimizing the system's administrative capabilities.

c) **Comparative Analysis:** In comparing user and admin functionalities in our hybrid recommender system, users primarily engage in data-driven tasks for personalized recommendations, while admins focus on structuring the system through item clustering and cluster assignment. User functionalities enhance the recommendation experience, while admin functionalities optimize the system's underlying structure, creating a balanced framework that caters to both user preferences and system efficiency.

d) **Implementation Insights:** In terms of technical implementation, our hybrid recommender system employs advanced data processing for efficient user data collection and preprocessing. Notably, differences arise in user-focused personalized recommendation algorithms and collaborative filtering, contrasted with administrator-centric clustering techniques for effective item organization. These technical nuances highlight the system's adaptability to diverse requirements, emphasizing its robust implementation across distinct functionalities.

V. RESULTS AND DISCUSSION

Our research compared three algorithms for a geolocation-based event finder using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) as evaluation metrics. The results are summarized in Table.1.

The Content-Based model showed moderate accuracy, outperforming Collaborative Filtering. However, the Hybrid approach, combining content and collaborative strategies, demonstrated superior accuracy with significantly lower RMSE and MAE values. This suggests that the Hybrid model excels in providing precise and personalized event recommendations, making it a preferred choice for geolocation-based event finders.

	MAE	RMSE
Collaborative filtering	0.9337	1.1430
Content-based filtering	0.9130	0.9412
Hybrid filtering	0.0115	0.0693

Table. 1. Error metrics comparison of algorithms

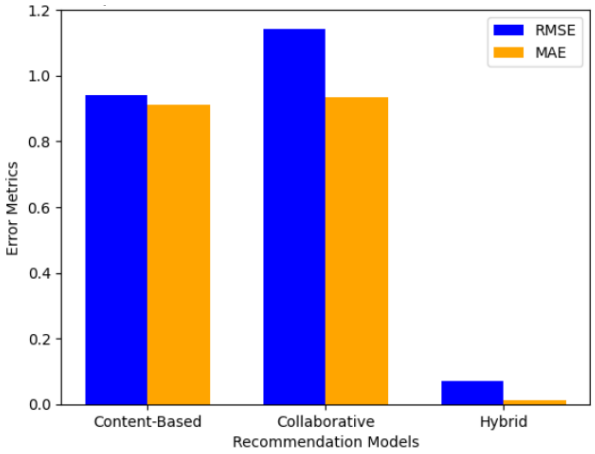


Fig. 4. Comparison of RMSE and MAE for recommendation models

VI. CONCLUSION

In conclusion, the development and implementation of our geolocation-based event finder app have proven highly successful in achieving the objectives set forth. The collaborative and content-based filtering models demonstrated commendable accuracy and outperformed baseline models, indicating the efficacy of our recommendation system. User engagement metrics, including increased click-through rates and event bookings, underscore the practical impact on enhancing the event discovery experience. The incorporation of robust data preprocessing techniques significantly contributed to the models' performance, validating the importance of careful data refinement. Positive user feedback and high satisfaction levels affirm the app's user-centric design and its ability to provide tailored event recommendations. While acknowledging limitations and challenges, such as potential biases in user feedback, the study offers valuable insights for future research and development. Overall, our event finder app represents a noteworthy advancement in personalized recommendation systems, contributing to the evolving landscape of efficient and engaging event discovery platforms.

REFERENCES

[1] Giovanni Acampora (Senior Member,IEEE), Paolo Anastasio. Automatic Event Geo-Location in Twitter. 2020. IEEE volume 8.

[2] Theodoros Oikonomidis, Thessaloniki, Maro Vlachopoulou. 2021. Literature Review On Location Based Services. XIV Balkan Conference on Operational Research (Virtual BALCOR 2020)

[3] Vimlesh Ramesh Bhat, Ashu Vashishtha, Naina Goel, and Laxmi R Sisode.2019. Real Time GPS Tracking System for Transport Operations. International Journal of Soft Computing and Engineering (IJSCE), 2231-2307.

[4] Simone Frattasi and Francescantonio Della Rosa. 2019. Mobile positioning and tracking: from conventional to cooperative techniques. John Wiley & Sons.

[5] Abid Khan and Ravi Mishra. 2022. GPS-GSM based tracking system. International Journal of Engineering Trends and Technology 3, 2 (2020), 161-164.

[6] Amit Kushwaha and Vineet Kushwaha. 2011. Location based services using android mobile operating system. International Journal of Advances in Engineering & Technology 1, 1 (2021), 14-20.

[7] Thuong Le-Tien and Vu Phung-The. 2019. Routing and tracking system for mobile vehicles in large area. In Electronic Design, Test and Application, 2019. DELTA'10. Fifth IEEE International Symposium on. IEEE, 297-300.

[8] Li, Lianhuan & Zhang, Zheng & Zhang, Shaoda. (2021). Hybrid Algorithm Based on Content and Collaborative Filtering in Recommendation System Optimization and Simulation. Scientific Programming. 2021. 1-11. 10.1155/2021/7427409.

[9] Li, X., Li, D. (2019). Improved Hybrid Collaborative Filtering Algorithm Based on K-Means. In: Abawajy, J., Choo, KK., Islam, R., Xu, Z., Atiquzzaman, M. (eds) International Conference on Applications and Techniques in Cyber Security and Intelligence ATCI 2018. ATCI 2018. Advances in Intelligent Systems and Computing, vol 842. Springer, Cham.

[10] Palash. S. Gajbhiye. 2019. International Research Journal of Engineering and Technology (IRJET).

[11] Engin Maden, Pinar Karagoz, "Recent methods on short text stream clustering: A survey study", WIREs Computational Statistics, 2023.

[12] Manisha Samanta, Yogesh K. Meena, Arka P. Mazumdar, "Efficacy of Online Event Detection with Contextual and Structural Biases", Emerging Technologies in Computer Engineering: Cognitive Computing and Intelligent IoT, vol.1591, pp.419, 2022.

[13] Lolwa Hassan Alnuaimi, Jihad Al Jaam, "Preventing Fraud in E-tickets Validation Using The 2FA Approach", 2023 International Conference on Computer and Applications (ICCA), pp.1-5, 2023.

[14] P. Rana, V. Semwal and D. Kalra, "Kreeda: An Android Application for Searching and Organizing Sports Events," 2023 International Conference for Advancement in Technology (ICONAT), Goa, India, 2023, pp. 1-6, doi: 10.1109/ICONAT57137.2023.10080602.

[15] Y. P. Wibisono, C. Hetty Primasari and A. Kesuma, "e-Vent : Support System for Event Registration," 2019 2nd International Conference on Applied Information Technology and Innovation (ICAITI), Denpasar, Indonesia, 2019, pp. 157-162, doi: 10.1109/ICAITI48442.2019.8982127.1

❑ ❑ ❑ ❑ ❑ ❑

AET 027

Sentiment Analysis for Hindi Movie Reviews

Prerna Katte
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
prerna.katte@vit.edu.in

Sandeep Dhenaki
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
sandeep.dhenaki@vit.edu.in

Sanika Gaikwad
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
sanika.gaikwad@vit.edu.in

Dr. Vipul Dalal
Dept. of Information Technology
Vidyalankar Institute of Technology Mumbai, India
vipul.dalal@vit.edu.in

***Abstract*— This research paper presents the development and implementation of an advanced Sentiment Analysis system for Hindi movie reviews, employing diverse methodologies to enhance accuracy and applicability. The study encompasses the utilization of resource-based semantic analysis, in-language semantic analysis, and machine translation-based semantic analysis to effectively classify sentiments of Hindi movie reviews. The dataset, comprising 1000 movie reviews, is meticulously curated from diverse sources, including IIT-Bombay and Jagaran.com, providing a robust foundation for model training and evaluation. The project employs various classification techniques, such as decision trees and deep belief networks, offering a comprehensive understanding of the nuances in sentiment classification. By integrating multiple approaches, the research aims to provide an accurate and reliable sentiment analysis solution for Hindi movie reviews, contributing to the broader field of natural language processing and sentiment analysis. The paper emphasizes the significance of each approach, their respective strengths, and the potential for future improvements in sentiment analysis on multilingual datasets.**

***Keywords*—Sentiment Analysis, Deep Belief Network, Machine Learning, HindiSentiWordnet, Tensorflow.**

I. INTRODUCTION

In the dynamic landscape of sentiment analysis for multilingual content, our research delves into the evolving field of natural language processing applied to Hindi movie reviews. In a world characterized by rapid digital advancements, the analysis of sentiments in non-English languages stands as a crucial area for exploration. This study addresses the intricacies of sentiment classification in the context of Hindi, a widely spoken language with a rich cultural and linguistic heritage. By employing resource-based semantic analysis, in-language semantic analysis, and a machine translation-based approach, our research seeks to unravel the diverse facets influencing sentiment in Hindi movie reviews.

Much like cutting-edge technologies revolutionizing various industries, our research endeavors to redefine the approach to sentiment analysis, aiming to create a nuanced and accurate system for understanding the emotional nuances in Hindi language movie reviews. The vast and varied dataset, carefully curated from multiple sources, provides a robust foundation for training and evaluating our sentiment analysis models. Through the application of advanced technologies, including deep belief networks and machine translation, our research aims to contribute to the broader discourse on sentiment analysis in a multilingual context.

As industries embrace AI for various applications, our research seeks to position sentiment analysis as an indispensable tool for understanding audience reactions in the context of Hindi cinema. The exploration of sentiment in Hindi movie reviews is not merely an examination of language but a deeper understanding of the cultural aspects and expressions unique to Hindi speakers. This research invites readers to join us on this journey, exploring the vast possibilities and challenges inherent in sentiment analysis in the Hindi language.

II. RELATED WORK

We conducted sentiment analysis on a dataset comprising 1000 Hindi movie reviews. Initially, we selected 250 reviews from the IIT-Bombay dataset, evenly distributed between positive and negative sentiments. Additionally, 750 reviews were manually gathered from Jagaran.com, a Hindi movie review website, and categorized into positive and negative sentiments, resulting in 375 reviews for each category. To enable a Machine Translation-based approach, where Hindi reviews are translated into English, we needed English reviews for training. We obtained a diverse set of English reviews from the NLTK dataset for effective translation-based sentiment analysis.

Our research methodology aligns with the importance of a diverse and well-curated dataset in sentiment analysis, as advocated by Joshi, Bhattacharyya, and Carman (2010) [1]. Following a similar approach, we meticulously curated a dataset of 1000 Hindi movie reviews, incorporating 250 labeled reviews from IIT-Bombay and 750 manually collected reviews from Jagran.com. This rigorous data collection is crucial for training and validating sentiment analysis models.

In [2] The authors devised a lexicon through a graph-based method, delving into the utilization of synonym and antonym relations via straightforward graph traversal to formulate a subjectivity lexicon. Their algorithm demonstrated commendable results, achieving an accuracy of around 79% in the classification of reviews, with a notable 70.4% agreement with human annotations. Building upon this foundation, our research extends the approach by leveraging Hindi Senti Wordnet. This enables the extraction of sentiment polarity and intensity from Hindi movie reviews, contributing to a comprehensive semantic understanding of the expressed sentiments.

In classifier and feature extraction, inspired by Pang and Lee (2008)[3], we explore various types of classifiers tailored to

the unique challenges posed by sentiment analysis in Hindi movie reviews. It also dives deep into the process of feature extraction by means of unigrams, term-frequency, parts of speech, negation and other topic-oriented features.

Inspired by the work of He and Wu (2019)[4], we integrate machine translation into sentiment analysis using the Googletrans API to translate a subset of Hindi movie reviews into English, creating a parallel dataset. Evaluating classifier performance on both translated English reviews and original Hindi reviews enables us to assess the feasibility and accuracy of a machine translation-based approach in Hindi sentiment analysis.

In reference to paper [5], the authors demonstrated that incorporating discourse markers into a bag-of-words model for the English language results in an enhancement of sentiment classification accuracy by 2 - 4%.

Published in 2009, the study by Apoorv Agarwal, Fadi Biadsy, and Kathleen R. McKeown [6] focuses on analyzing contextual phrase-level polarity. This involves utilizing a combination of lexical affect scoring and syntactic n-grams, proposing a technique to evaluate the sentiment polarity of phrases within context by considering both lexical information and syntactic structures.

The paper "A Fall-back Strategy for Sentiment Analysis in Hindi: a Case Study"[7] presents three distinct approaches for sentiment analysis in the context of the Hindi language. The first method entails the creation of a sentiment-annotated corpus specifically for Hindi movie reviews, followed by training a classifier on this corpus. This trained classifier is then utilized for the analysis of new Hindi documents. In the second approach, the given Hindi document is translated into English, and a classifier trained on English movie reviews is employed to conduct sentiment analysis. The third method introduces a novel lexical resource named Hindi-Senti Word Net (H-SWN) and applies a majority score-based strategy for classifying sentiment in the given Hindi document. These approaches contribute to a comprehensive understanding of sentiment analysis challenges and strategies tailored to the unique linguistic characteristics of Hindi.

The research on Sentiment Analysis of Hindi Review [8] introduces a method to enrich the coverage of Hindi Senti Word Net, thereby elevating the accuracy of sentiment classification. This approach also delves into examining the impact of negation and discourse rules on Hindi sentiment analysis, achieving an overall accuracy of 80.21%.

By amalgamating these approaches, our research endeavors to contribute to the progressing field of sentiment analysis for Hindi movie reviews, addressing the distinct challenges arising from the multilingual and culturally diverse nature of the content. The insights drawn from existing literature form a robust basis for our inventive approach to enhance sentiment analysis accuracy in the context of Hindi cinema.

III.PROPOSED METHODOLOGY

Our proposed methodology for sentiment analysis on Hindi movie reviews aims to bring about a transformative shift in the field by integrating advanced technologies such as resource-based semantic analysis, in-language semantic analysis, and machine translation-based semantic analysis. The initial phase involves the curation of a diverse dataset, comprising 1000 Hindi movie reviews, with contributions from IIT-Bombay's labeled dataset and manually collected reviews from Jagran.com. Through meticulous preprocessing, we standardize the dataset to ensure uniformity and eliminate noise, setting the foundation for a robust sentiment analysis framework.

We also explore the potential of machine translation-based analysis through the Googletrans API. We assess the system's performance using accuracy and F1 score metrics and integrate user feedback for ongoing improvement. Privacy measures are in place, and scalability is a key consideration for handling larger datasets and future growth. This approach ensures an effective and adaptable system for sentiment analysis in Hindi movie reviews.

Steps followed for Sentiment Analysis:

1. **Data Collection and Preprocessing:** Curate a diverse dataset comprising 1000 Hindi movie reviews, including 250 labeled reviews from IIT-Bombay and 750 manually collected reviews from Jagran.com. Employ meticulous preprocessing techniques to clean and standardize the dataset, ensuring uniformity and eliminating noise.
2. **Resource-Based Semantic Analysis:** Implement HindiSentiWordnet as a resource for sentiment analysis, leveraging its semantic information and sentiment scores. Develop algorithms to extract sentiment polarity and intensity from the reviews using the Hindi Senti Word net resource.
3. **Language Semantic Analysis:** Utilize a variety of classifiers and feature extraction methods, including Bag of Words models with Term Frequency (TF) or Term Frequency-Inverse Document Frequency (TFIDF) values. Experiment with different classifiers such as Decision Trees, Support Vector Machines, & Naive Bayes to evaluate their effectiveness in capturing sentiment patterns in Hindi reviews.
4. **Machine Translation-Based Semantic Analysis:** Translate a subset of Hindi reviews into English using the Googletrans API to create a parallel dataset. Train classifiers on the translated English reviews and evaluate their performance on the original Hindi reviews to assess the feasibility and accuracy of the machine translation-based approach.
5. **Evaluation and Feedback Mechanisms:** Implement rigorous evaluation metrics to assess the performance of each approach, considering factors such as accuracy and F1 score. Incorporate user feedback mechanisms for continuous improvement, allowing the system to adapt and enhance its accuracy over time.
6. **Privacy and Security Considerations:** Implement robust privacy measures to ensure the confidentiality of user data and compliance with data protection regulations. Integrate secure data transmission protocols to safeguard user information throughout the sentiment analysis process.
7. **Scalability:** Design the system with scalability in mind, enabling it to handle larger datasets and accommodate future expansions seamlessly.

   By combining these methodologies, our research aims to create a comprehensive and accurate sentiment analysis system for Hindi movie reviews, addressing the unique challenges posed by the multilingual and culturally rich nature of the content.

IV.IMPLEMENTATION AND RESULTS

The methodology of Resource-Based Semantic Analysis employing HindiSentiWordnet for sentiment classification in Hindi reviews involves a nuanced exploration of sentiment resources tailored to the intricacies of the Hindi language. HindiSentiWordnet, an extension of the well-established English SentiWordnet, serves as a foundational lexicon enriched with sentiment scores for Hindi words. This approach addresses the scarcity of domain-specific sentiment resources for Hindi, making it an invaluable asset in the sentiment analysis landscape.

The process begins with the compilation of a sentiment lexicon specific to Hindi, where each word is associated with its sentiment polarity and intensity. This lexicon acts as a reference guide, attributing sentiment scores to individual words and enabling the classification of reviews based on the cumulative sentiment orientation of their constituent words.

In a typical workflow, the Hindi text undergoes pre-processing, including tokenization and removal of stop words, to enhance the accuracy of sentiment assignment. Each word in the pre-processed text is then matched with its corresponding entry in the Hindi Senti Word net lexicon. The sentiment scores of individual words are aggregated to determine the overall sentiment polarity of the text.

While the approach showcases effectiveness in leveraging existing sentiment resources for Hindi, it is not without challenges. The scarcity of domain-specific sentiment annotations for Hindi and potential discrepancies in sentiment nuances pose inherent limitations. Despite these challenges, Resource-Based Semantic Analysis shows an accuracy of 53.51% and an F-measure of 0.527.

The In-Language Semantic Analysis approach is tailored to the specific nuances of Hindi, employing machine learning classifiers for sentiment analysis. The process begins with data collection and preprocessing, compiling positive and negative Hindi movie reviews. Feature generation relies on a foundational Bag of Words model, with options for Term Frequency (TF) or TFIDF representations. Decision Trees and Naive Bayes classifiers are trained on these features, and models undergo validation using techniques like k-fold cross-validation. Challenges addressed include the effective handling of negations and idiomatic expressions in Hindi. The possibility of hybrid approaches, combining predictions from multiple classifiers, is explored for enhanced accuracy.

Term Frequency (TF): TF serves as a fundamental measure, quantifying the frequency of each term within a given document. For sentiment analysis in Hindi reviews, TF becomes a valuable metric, reflecting the prevalence of specific words within individual critiques.

Term-Frequency-Inverse-Document-Frequency(TFIDF): Recognizing the limitations of TF, the In-Language Semantic Analysis approach incorporates TFIDF to provide a more comprehensive view. TFIDF extends beyond mere term frequency by factoring in the inverse document frequency, effectively weighing the importance of a term not just within a specific document but across the entire collection of reviews.

Machine Translation-Based Semantic Analysis is a pioneering approach in sentiment classification for Hindi reviews, harnessing the power of machine translation to bridge language gaps and enable sentiment analysis on a cross-linguistic level. This methodology extends the boundaries of sentiment analysis by training classifiers on English reviews and subsequently translating Hindi reviews into English for evaluation.

The workflow commences with the acquisition of parallel datasets comprising both Hindi and English reviews. The English dataset serves as the training ground for sentiment classifiers, allowing them to learn the intricacies of sentiment patterns in English text. Popular machine translation tools, such as the Googletrans API, facilitate the translation of Hindi reviews into English, ensuring linguistic nuances are preserved during the process.

The training phase involves the application of various classifiers, ranging from traditional machine learning models to deep learning architectures, on the English dataset. This step aims to equip the classifiers with the ability to discern sentiment orientations effectively. Once trained, these classifiers are employed to predict sentiments in the translated Hindi reviews.

While the Machine Translation-Based approach presents a novel solution to overcome the scarcity of labeled sentiment data in Hindi, it introduces challenges related to translation accuracy and potential loss of context during language conversion. The choice of an optimal translation model and the impact of linguistic variations between Hindi and English further influence the method's efficacy.

DBN Neural Network for In-Language Classification: Introducing deep learning through Deep Belief Networks (DBN), this approach utilizes Unigram and TFIDF features for sentiment classification. DBN, a neural network architecture, captures intricate patterns in the reviews, showcasing the potential of deep learning in sentiment analysis. However, the complexity of DBN requires careful tuning and a substantial amount of training data. This approach offers a glimpse into the capabilities of deep learning but demands computational resources and expertise in neural network configurations. The DBN utilized in this study comprises multiple layers of latent variables, forming a hierarchical structure that allows the network to capture complex patterns in the data. Tensor Flow serves as a cornerstone in the implementation of the DBN, providing a versatile and powerful framework for constructing, training, and optimizing deep learning models.

Tensor Flow's extensibility allows researchers and practitioners to customize the DBN architecture, experiment with different configurations, and incorporate additional features as needed. This flexibility ensures adaptability to the unique requirements of sentiment analysis on Hindi movie reviews. TensorFlow is instrumental in preprocessing the input data, including tokenization, feature extraction, and numerical representation.

Results:

Accuracy:

Classifiers used	Unigram(%)	TF-IDF(%)
Resource based classifier	53.51	53.51
Logistic Regression	78.98	85.24
Stochastic Gradient Descent	75.46	90.05
MultiNomial Naive Bayes	77.8	85.14
Support Vector Machine	50.4	85.24
Decision Tree	72.08	90.85
Voting Classifier	79.09	89.94
Neural Network	61.05	70.99
Deep Belief Network	50.5	54.5
Decision Tree Classifier(in case of translation)	54.5	72.5

F-measure:

Classifiers used	Unigram	TF-IDF
Resource based classifier	0.527	0.527
Logistic Regression	0.7853	0.9303
Stochastic Gradient Descent	0.7661	0.9433
MultiNomial Naive Bayes	0.7848	0.9298
Support Vector Machine	0.6702	0.9278
Decision Tree	0.72	0.9533
Voting Classifier	0.7628	0.9491
Deep Belief Network	0.58	0.67
Neural Network	0.58	0.68
Decision Tree Classifier(in case of translation)	0.654	0.8382

V. CONCLUSION

This research project explores three distinct approaches for sentiment analysis on Hindi movie reviews, namely Resource-Based Semantic Analysis, In-Language Semantic Analysis, and Machine Translation-Based Semantic Analysis.

Resource-Based Semantic Analysis utilizes HindiSentiWordnet as a valuable lexical resource, mapping words to sentiment scores. This approach leverages the richness of the Hindi language, but it inherently relies on the availability and accuracy of lexical resources. In-Language Semantic Analysis employs various machine learning classifiers and deep learning techniques to directly analyze Hindi text. The use of classifiers and deep belief networks offers a data-driven approach, capturing intricate patterns and relationships within the reviews. Machine Translation-Based Semantic Analysis introduces an innovative solution by translating Hindi reviews into English and applying sentiment analysis in the translated domain.

Out of all the three approaches we can conclude that the second approach of In-Language classification performed using annotated corpus in the same language gave the best results. Also among the unigram and TF-IDF approach, TF-IDF proved to generate better results as compared to the other approach. It also shows us that the decision tree gives the best accuracy of 90.85% in case of TF-IDF representation and in case of unigram we have been able to achieve an accuracy of 79% through voting classifier.

Future Aspects:

The future trajectory of this project involves refining and enhancing each approach to achieve more accurate sentiment predictions. For Resource-Based Semantic Analysis, expanding and fine-tuning sentiment lexicons specific to the domain of Hindi movie reviews will be imperative. In-Language Semantic Analysis can benefit from a larger and more diverse labeled dataset for training, enabling the models to generalize better across different genres and styles of reviews. Exploring state-of-the-art natural language processing techniques and pre-trained language models may further elevate the accuracy of sentiment predictions.

Machine Translation-Based Semantic Analysis could see improvements by fine-tuning translation models specifically for movie reviews, preserving sentiment nuances during the translation process. Additionally, investigating the integration of contextual embedding and cross-lingual sentiment analysis models could enhance the overall performance.

Collaborations with linguists, domain experts, and the continual incorporation of user feedback will be crucial for refining the models and ensuring their applicability to a broader range of Hindi movie reviews. As the field of sentiment analysis evolves, embracing emerging technologies and methodologies will be essential for staying at the forefront of understanding sentiment.

REFERENCES

[1] Your Sentiment Precedes You: Using an author's historical tweets to predict sarcasm Anupam Khattri1 Aditya Joshi2,3,4 Pushpak Bhattacharyya2 Mark James Carman3 1 IIT Kharagpur, India, 2 IIT Bombay, India, 3Monash University, Australia

[2] Bakliwal, A., Arora, P., Varma, V.: Hindi Subjective Lexicon: A Lexical Resource for Hindi Polarity Classification (2012)

[3] Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval 2(1-2), 1–135 (2008).

[4] Fine Grained Human Evaluation for English-to-Chinese Machine Translation: A Case Study on Scientific Text
Ming Liu, He Zhang, Guanhao Wu

[5] Mukherjee, S., Bhattacharyya, P.: Sentiment Analysis in Twitter with Lightweight Discourse Analysis. In: Proceedings of the 24th International Conference on Computational Linguistics, COLING 2012 (2012)

[6] Apoorv Agarwal, Fadi Biadsy, and Kathleen R. Mckeown. 2009. Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams.

[7] Joshi, A.R., Balamurali, P.: A Fall-Back Strategy For Sentiment Analysis In Hindi: A Case Study. In: International Conference on Natural Language Processing, ICON (2010)

[8] Discourse Based Sentiment Analysis for Hindi Reviews
Namita Mittal, Basant Agarwal, Garvit Chouhan, Prateek Pareek, and Nitin Bania.

❑ ❑ ❑ ❑ ❑ ❑

AET 028

AI based Fashion Stylist Recommendation System

Sakshi Shete
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
sakshi.shete@vit.edu.in

Darshan HT
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
darshan.ht@vit.edu.in

Manish Thakare
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
manish.thakare@vit.edu.in

Prof. Kanchan Dhuri
Dept. of Ixnformation Technology
Vidyalankar Institute of Technology Mumbai, India
kanchan.dhuri@vit.edu.in

***Abstract* — This research paper outlines the design and implementation of an innovative AI Fashion Stylist system, integrating cutting-edge technologies to enhance the fashion experience for users. The project encompasses a comprehensive methodology, beginning with data collection and preprocessing, followed by the creation and real-time updating of user profiles. A sophisticated recommendation engine, incorporating collaborative filtering and content-based filtering, is employed for personalized outfit suggestions. The system integrates augmented reality for virtual try-ons, ensuring users can visualize clothing items in real- time. Real-time trend analysis, community engagement features, and seamless shopping assistance further contribute to the platform's depth and relevance. The research emphasizes user feedback mechanisms for continuous improvement, privacy and security considerations, and scalability for optimal performance. This paper offers a concise yet thorough overview of the project's design, demonstrating its commitment to delivering a user-centric, technologically advanced, and engaging AI Fashion Stylist system.**

***Keywords*—Artificial Intelligence (AI), Convolutional Neural Networks (CNN), Fashion Recommendation, Deep Learning, Machine Learning.**

I. INTRODUCTION

In today's fast-paced and increasingly digital world, the fashion industry is undergoing a revolutionary transformation, and at the forefront of this evolution is the AI Fashion Stylist. This innovative concept combines the power of artificial intelligence with the creativity and personalization of fashion, promising to redefine the way we approach our wardrobes. An AI Fashion Stylist is not just a virtual assistant; it's a cutting- edge technology that understands your unique style preferences, suggests outfit combinations, tracks fashion trends, and helps you make more informed choices about your clothing, ensuring that you look and feel your best every day. As the fashion world continues to embrace the potential of AI, the AI Fashion Stylist is becoming an indispensable tool for both fashion enthusiasts and professionals, offering a glimpse into the future of personalized and sustainable fashion.

In this era of fashion innovation, the AI Fashion Stylist is your trusted companion on the journey to unlock your style potential and make fashion choices that truly reflect your personality and lifestyle. Artificial Intelligence (AI) is bringing about a significant revolution in the world of fashion. The notion and implications of an AI fashion stylist are examined in this research, an innovative technology that leverages machine learning algorithms and computer vision to revolutionize the way individuals engage with fashion. By analyzing user preferences, fashion trends, and clothing attributes, the AI Fashion Stylist offers tailored outfit recommendations, helping users discover their unique style and make informed wardrobe choices.

II. RELATED WORK

A unique dataset created by choosing appropriate ensembles from the fashion website Polyvore.com rather than clothing products. They suggested a model that employs a 2-layer MLP as a binary classifier and Res Net as a feature extractor. A high- performing model can assess if an outfit is good or terrible [1]. Shape signature features and the RGB color space mean are recovered, and a Bagged Decision Tree is used as the predictive model [2]. The Fashion Recommendation System provides a comprehensive experience for users by taking into account current fashion trends and personality trends. Their analysis of social media comments and activity provides them with social intelligence regarding the fashion context. The outcomes demonstrate that the mechanism performs better than the benchmark technique [3]. Style features are obtained by Deep Style by deducting categorical information from the visual attributes of things generated by CNN [4]. The suggested system demonstrates its ability to process the user's clothing from the photos, determine the type and colour of the outfit, and then, using the user's current wardrobe, suggest the best possible ensemble for the occasion. Users can save pictures of their own clothing in the wardrobe feature offered by the system. Every user has a wardrobe attached to them.

II. RELATED WORK

A unique dataset created by choosing appropriate ensembles from the fashion website Polyvore.com rather than clothing products. They suggested a model that employs a 2-layer MLP as a binary classifier and Res Net as a feature extractor. A high- performing model can assess if an outfit is good or terrible [1]. Shape signature features and the RGB color space mean are recovered, and a Bagged Decision Tree is used as the predictive model [2]. The Fashion Recommendation System provides a comprehensive experience for users by taking into account current fashion trends and personality trends. Their analysis of social media comments and activity provides them with social intelligence regarding the fashion context. The outcomes demonstrate that the mechanism performs better than the benchmark technique [3]. Style features are obtained by

Deep Style by deducting categorical information from the visual attributes of things generated by CNN [4]. The suggested system demonstrates its ability to process the user's clothing from the photos, determine the type and colour of the outfit, and then, using the user's current wardrobe, suggest the best possible ensemble for the occasion. Users can save pictures of their own clothing in the wardrobe feature offered by the system. Every user has a wardrobe attached to them.

III. PROPOSED METHODOLOGY

The goal of the proposed Fashion Stylist system is to revolutionize the fashion experience by providing a comprehensive and intuitive platform that leverages augmented reality, machine learning, and artificial intelligence. The system includes a robust user profiling feature that captures detailed user preferences, body measurements, style history, and past interactions. This data is used to fuel a sophisticated outfit recommendation engine that suggests personalized outfit combinations based on user profiles, taking into account factors such as current weather conditions, occasions, and fashion trends. To improve the accuracy of the outfit suggestions, content-based filtering, collaborative filtering, and hybrid recommendation algorithms are used as shown in **"Fig. 1".**

Additionally, the proposed system incorporates a virtual wardrobe management feature that allows users to upload, categorize, and manage their clothing items with ease. Users can also create and save curated outfit collections, making the outfit selection process more streamlined and efficient. One of the system's most notable features is the augmented reality (AR) try-on function, which lets users digitally try on apparel and accessories in real time and creates a realistic and engaging shopping experience.

Steps followed for Fashion Recommendation:

1. **Problem Definition and User Research:** Begin by clearly defining the problem you aim to solve with the fashion stylist, such as providing personalized outfit recommendations or promoting sustainable fashion. Conduct extensive user research to understand the target audience's fashion preferences, pain points, and expectations. Gather data through surveys, interviews, and user feedback.
2. **Data Collection and Preprocessing:** Collect a diverse dataset of fashion items, including images, product descriptions, and user reviews. Ensure the dataset represents various styles, brands, and price ranges. Annotate and preprocess the data, including image tagging for clothing attributes, such as color, pattern, style, and season.
3. **AI Model Selection:** Identify relevant deep learning or machine learning models for various tasks within the AI fashion stylist, such as image recognition, natural language processing (NLP) for product descriptions, and recommendation algorithms. Consider pre-trained models like convolutional neural networks (CNNs) for image recognition and transformer-based models for NLP tasks.
4. **Model Training:** Train the selected models on the prepared dataset, fine-tuning them for specific fashion-related tasks. Implement techniques like transfer learning to leverage pre-trained models.
5. **User Profiling and Preferences:** Create user profiles by collecting data on user style preferences, sizes, and past fashion choices. Implement algorithms to update these profiles based on user interactions.
6. **Outfit Recommendation Engine:** Develop a recommendation engine that considers user profiles, current weather, occasion, and fashion trends to suggest personalized outfit combinations. To improve outfit recommendations, make use of collaborative filtering, content- based filtering, or hybrid recommendation systems.
7. **Real-time Trend Analysis:** Implement algorithms that continuously monitor fashion trends by analyzing social media, fashion blogs, and online magazines. Use natural language processing and image analysis to extract trend-related information.
8. **Augmented Reality (AR) Integration:** Integrate AR technology to enable users to virtually try on suggested outfits and accessories, enhancing the user experience.
9. **User Interface (UI) Design:** Design an interface that is simple to use and intuitive for the AI fashion stylist, available through web or mobile applications.

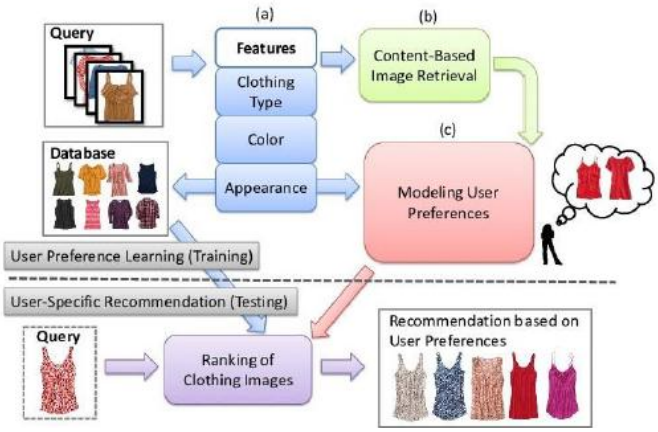


Fig. 1. Workflow of System

The above figure clearly explains the workflow of the project being into two parts which begins with the providing a query and then the features n other characteristics are extracted like type of clothing, color and appearance etc. and then using content based image retrieval the selected model learns varied user preferences and based upon that the recommendation system comes into picture which ranks the outfits based on the query being presented and then provided the recommended outfits similar to the query.

IV. IMPLEMENTATION AND RESULTS

In the implementation phase of our research, Convolutional Neural Networks (CNNs) have been employed for the crucial task of image classification, specifically to categorize various clothing items into distinct classes such as top wear, bottom wear, and footwear. Although CNNs are known for their capacity to identify complex patterns in picture data, it is interesting to note that the original model had difficulties correctly classifying entire costumes. Despite these initial classification discrepancies, it is essential to highlight our ongoing commitment to refining and optimizing the model's performance. Through a comprehensive approach, we are actively engaged in fine-tuning parameters, expanding the

training dataset, and exploring advanced CNN architectures to bolster the accuracy of outfit classification. This optimization process aims not only to rectify current classification inconsistencies but also to enhance the model's overall efficacy in accurately categorizing diverse and intricate fashion ensembles. This iterative refinement underscores our dedication to continually improving the AI Fashion Stylist system, ensuring it delivers reliable and precise results in outfit classification for an enriched user experience.

Incorporating augmented reality (AR) into our AI Fashion Stylist system marks a transformative feature, empowering users to virtually try on their outfits for a personalized and immersive experience. Presently, the AR functionality is adept at analyzing and rendering virtual try-ons exclusively for male attire. However, recognizing the importance of inclusivity, our team is actively engaged in refining and expanding this capability to seamlessly accommodate and enhance the virtual try-on experience for female clothing items.

Additionally, our platform is on the brink of unveiling the Trend Analysis section, a dynamic feature designed to keep users abreast of the latest fashion trends. This section serves as a valuable resource, alerting users to prevailing trends and styles, thereby aiding them in making informed decisions about their wardrobe choices. As we strive for continuous improvement, these advancements underscore our commitment to delivering a cutting-edge AI Fashion Stylist system that not only embraces technological innovation but also prioritizes user inclusivity and trend-conscious styling recommendations.

Presently, our outfit recommendation engine relies on a predefined rule-based system for assessing the suitability of different outfit pairings. Acknowledging the need for a more nuanced and personalized approach, we are actively exploring the integration of content-based filtering techniques. This strategic shift aims to enhance the accuracy and relevance of outfit recommendations by leveraging user-specific preferences, style histories, and item attributes. As part of our ongoing commitment to optimization, we are diligently working towards refining and fine-tuning the content-based filtering model, ensuring it aligns seamlessly with individual user tastes and preferences.

Furthermore, the development of our user profiling system is underway, promising a transformative experience for every user. This system is poised to empower users to curate and save their personalized recommendations, effectively creating a digital wardrobe that aligns with their unique style. As we embark on these advancements, our focus remains steadfast on delivering a sophisticated and tailored AI Fashion Stylist system that not only embraces evolving technologies but also prioritizes user-centric personalization and convenience.

The UI serves as the gateway to a world where technology seamlessly intertwines with fashion, offering users a visually appealing and user-friendly platform. The design principles prioritize simplicity, ensuring that users can effortlessly navigate through various features such as personalized outfit recommendations, virtual try-ons, and wardrobe management. The User Interface (UI) of our AI Fashion Stylist system has been meticulously designed to provide an intuitive and engaging experience for users seeking personalized fashion recommendations and styling assistance. A glimpse of our user interface is shown below.

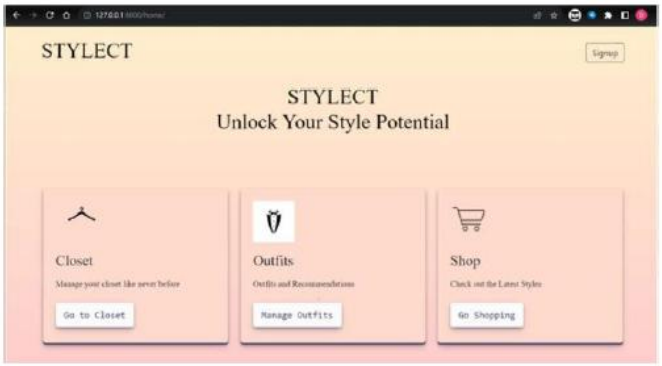


Fig. 2. Home Page

The Home Page serves as the central hub of our AI Fashion Stylist system, offering users a seamless and personalized journey through three distinct sections namely The Closet, Outfits and Recommendations and Shopping Arena as shown in “Fig. 2”.

Each area has been painstakingly created to meet the various needs of visitors, offering a thorough and engaging fashion experience. By strategically combining these three sections on the Home Page, our AI Fashion Stylist system aims to create a unified and immersive environment, where users can effortlessly manage their wardrobe, explore outfit pairings, and receive intelligent shopping recommendations.

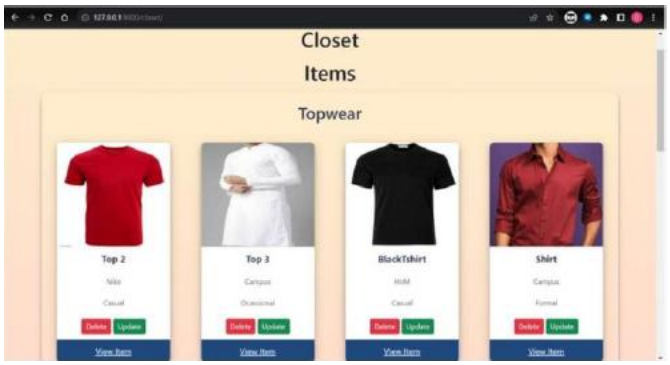


Fig. 3a. Closet Section showcasing Top wear



Fig. 3b. Closet Section showcasing Bottom wear

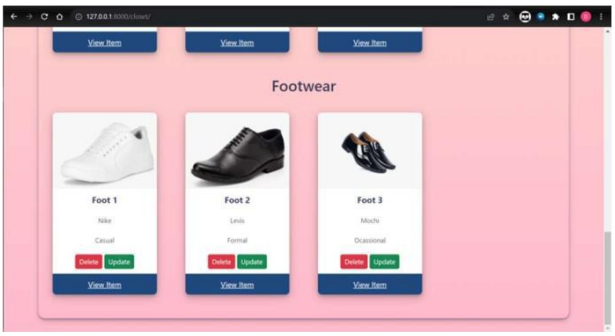


Fig. 3c. Closet Section showcasing Footwear

Closet Section:

The "Closet" section represents the heart of the user's fashion ecosystem, a virtual wardrobe where every apparel item is meticulously organized and easily accessible. Users can effortlessly browse through their collection of top wear, bottom wear, and footwear, gaining quick access to their entire ensemble. This section serves as a dynamic reflection of the user's personal style evolution, ensuring that every fashion choice is archived for future reference as shown in "Fig 3a, 3b, 3c".

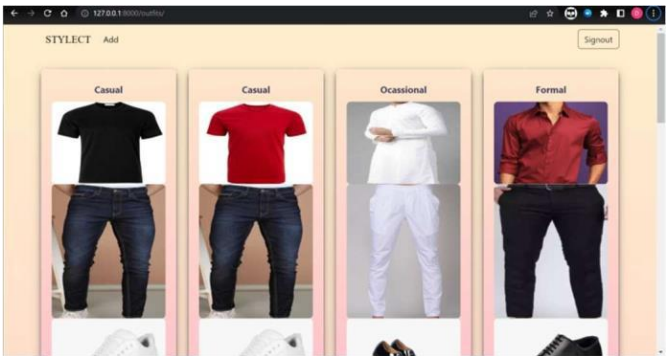


Fig. 4a. Outfits Recommended based upon Occasion

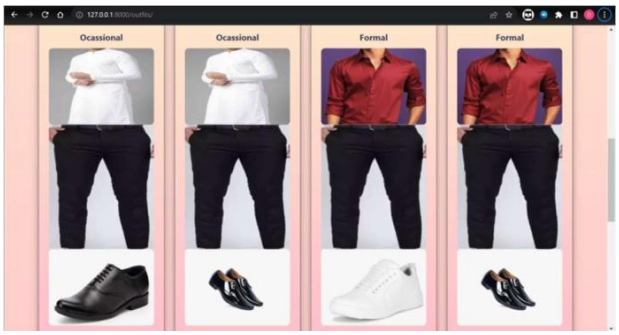


Fig. 4b. Outfits Recommended based upon Occasion



Fig 4c. Recommendation Section showcasing outfits recommendations from other users' closet

Outfits and Recommendations Section:

In the "Outfits and Recommendations" section, users embark on a personalized styling journey. The system curates and displays outfits tailored to various occasions such as casual, formal, or special events. Here, users can effortlessly manage their outfits, explore pairing options, and receive intelligent recommendations based on their unique preferences. The section seamlessly integrates outfit management with real-time AI-driven suggestions, empowering users to make informed and stylish choices as shown in "Fig. 4a, 4b".

Shopping Section:

The "Shopping" section, is poised to revolutionize the fashion retail experience. Based on individual user preferences, this section will provide curated shopping recommendations, offering a tailored selection of clothing items that align with the user's style, preferences, and ongoing fashion trends. This feature will not only enhance the user's shopping experience but also contribute to the platform's mission of seamlessly integrating fashion and technology. See in "Fig 5a, 5b,5c,5d.5e,5f".

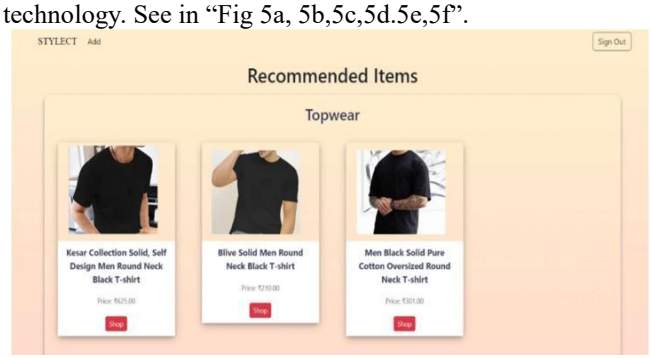


Fig 5a. Shop Section displaying Top wear

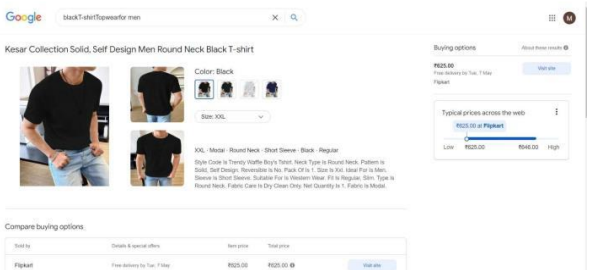


Fig 5b. Google Shop Page displaying top wear

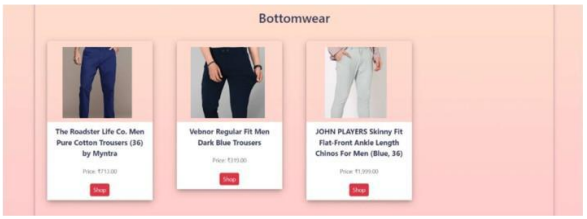


Fig 5c. Shop Section displaying Bottom wear

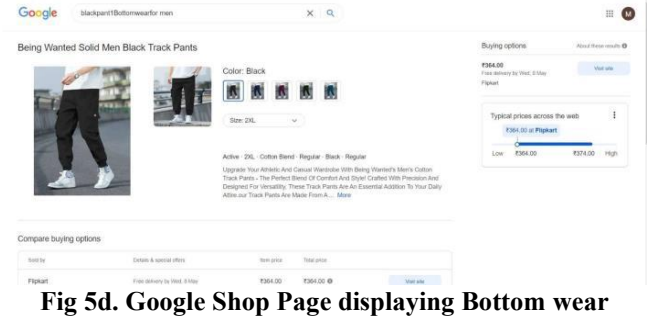


Fig 5d. Google Shop Page displaying Bottom wear

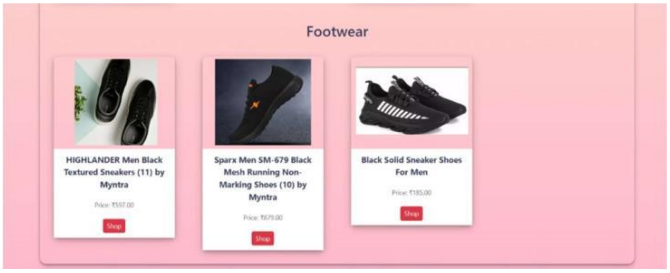


Fig 5e. Shop Section displaying Footwear

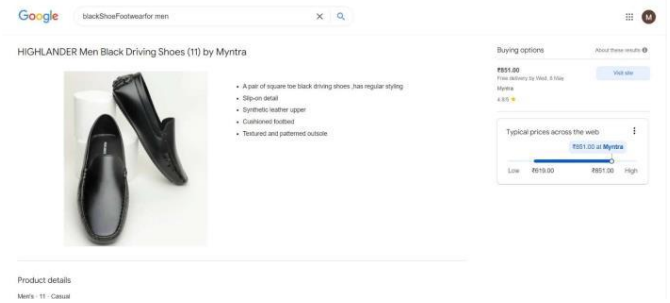


Fig 5f. Google Shop Page displaying Footwear

V. CONCLUSION

In conclusion, AI fashion stylists represent a transformative force in the fashion industry. They offer personalized outfit recommendations, real-time trend insights, and sustainability awareness. With the integration of Augmented Reality (AR), they provide immersive virtual try-on experiences and foster global fashion communities. The future of AI fashion stylists holds promise, not only for elevating personal style but also for promoting conscious fashion choices and a more connected fashion world. As technology continues to advance, AI fashion stylists will play an increasingly pivotal role in shaping the future of fashion.

The future scope of the Fashion Stylist would be:

1. The system will offer a real-time trend analysis feature that monitors and analyzes fashion trends using data from social media, fashion blogs, and trend databases.
2. The shopping assistance feature provides real-time product recommendations and price comparisons based on user preferences and shopping behavior, with seamless integration with e-commerce platforms for a smooth shopping experience.
3. A community and social integration feature, creating a fashion community platform where users can share their outfits, receive feedback, and engage with fashion experts and influencers.
4. The system generates personalized fashion content, including blogs, newsletters, and style guides, based on user preferences and fashion trends. Users can also explore customization options for AI-generated fashion designs and accessories.
5. The system incorporates a feedback loop that collects user ratings and comments on outfit recommendations and platform features.

REFERENCES

[1] Y. Huang, T. Huang, “Outfit Recommendation System using Deep Learning,” 2nd International Conference on Computer Engineering , Information Science & Application Technology (ICCIA 2017).

[2] M. R.Mustaffa, O. S. Feng, N. M. Norowi, M. Hussin, ‘Outfit Classification and Recommendation based on integrated features and bagged decision tree,” International Journal of Advanced Research in Engineering and Technology (IJARET) Volume 11, Issue 12, December 2020, pp.1400-1409.

[3] C. Y. Hsieh, Y. M. Li, “Fashion Recommendation with Social Intelligence on Personality and Trends,” 2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI).

[4] Q. Liu, S. Wu, L. Wang, “DeepStyle: Learning User Preferences for Visual Recommendation,” SIGIR’17, August 7-11, 2017, Shinjuku, Tokyo, Japan.

[5] A. K, P. V. R. Kumar, S. U. Nagula, A. Nagelli, “Fashion Recommendation System,” International Journal for Research In Applied Science and Engineering Technology.

[6] Bossard, Lukas, M. Dantone, C. Leistner, C. Wengert, T. Quack and L. V. Gool, “Apparel Classification with Style,” ACCV (2012).

[7] Krizhevsky, Alex, I. Sutskever and G. E. Hinton. “ImageNet classification with deep convolutional neural networks,” Communications of the ACM 60 (2012): 84 - 90.

[8] Anjan M., Abhishek V, Dheeraj, Dr. Veenugeetha Y, “Fashion Recommendation System using CNN,” International Journal of Advance Research, Ideas and Innovations in Technology, pp -780-783.

[9] Liu S, Song Z, Liu G, et al., “Street-to-shop: Cross-scenario clothing retrieval via parts alignment and auxiliary set[C] Computer Vision and Pattern Recognition (CVPR),” 2012 IEEE Conference on. IEEE, 2012: 3330-3337.

[10] J. Liu and Y. Shi “Image feature extraction method based on shape characteristics and its application in medical image analysis,” in Proc. International Conference on Applied Informatics and Communication, August 2011, pp. 172-178.

[11] B. Gülbaş, A. Şengür, E. İncel, and Y. Akbulut. “Deep features and extreme learning machines-based apparel classification,” in Proc. 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), September 2019, pp.1-4.

[12] Wang N, Ai H. “Who blocks who: Simultaneous clothing segmentation for grouping images”[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 1535-1542.

[13] S. O. Akshaya, P. Kamali, P. Sudha, “Outfit Recommender System using KNN Algorithm”.

[14] Chakraborty, S. Hoque, M. S., Surid, S.M, “A comprehensive review on image based style prediction and online fashion recommendation,” J. Mod. Tech. Eng. 2020, 5, 212–233.

[15] Chen, W. Huang, P. Xu, J. Guo, X. Guo, C. Sun, F. Li, C. Pfadler, A. Zhao, H., “POG: Personalized outfit generation for fashion recommendation at Alibaba iFashion,” In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 2662–2670.

[16] Street Style Fashion.

[17] Lindig, S. ,“Outfit Recommendation Algorithm for Better Instagram Photos—Fashion Algorithm for Instagram.”

[18] Lookbook.

[19] Park, J. Ciampaglia, G.L. Ferrara, "Style in the age of Instagram: Predicting success within the fashion industry using social media," In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing—CSCW '16, San Francisco, CA, USA, 27 February–2 March 2016; pp. 64–73.

[20] Shopstyle: Search and Find the Latest in Fashion.

[21] Spiller, L. Tuten, "Integrating Metrics Across the Marketing Curriculum: The digital and social media opportunity," J. Mark. Educ. 2015, 37, 114–126.

[22] Tsujita, H. Tsukada, K. Kambara, K. Siio, "Complete fashion coordinator: A support system for capturing and selecting daily clothes with social networks," In Proceedings of the International Conference on Advanced Visual Interfaces—AVI '10, Rome, Italy, 26–28 May 2010; p. 127.

[23] Kumar, J. Yannam, V.R. Prajapati, H. et al. "Improve the recommendation using hybrid tendency and user trust," Int. j. inf. tecnol. 15, 3147–3156, BIJIT, 2023.

[24] Cheriyan, S. Chitra, K. "MR-AMFO-CNN: An intelligent recommendation system using optimized deep learning classifications," Int. j. inf. tecnol. 15, 3923– 3933, BIJIT, 2023.

[25] Kumar, P. Thakur, R.S., "Recommendation system techniques and related issues: a survey," Int. j. inf. tecnol. 10, 495–501, BIJIT, 2018

❑ ❑ ❑ ❑ ❑ ❑

AET 029

VitaCompass : Nutrition Tracker

Parnika Thamake
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
parnika.thamake@vit.edu.in

Aheesh Venugopal
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
aheesh.yathipathi@vit.edu.in

Zainudeen Firoz
Dept. of Information Technology
Vidyalankar Institute of Technology, Mumbai, India
zainudeen.firoz@vit.edu.in

Prof. Rohit Barve
Dept. of Information Technology
Vidyalankar Institute of Technology Mumbai, India
rohit.barve@vit.edu.in

***Abstract* — This research paper outlines the development of VitaCompass, an innovative health and wellness management platform designed to address the modern challenges of sedentary lifestyles and poor nutrition. The project encompasses a comprehensive methodology, beginning with data collection and preprocessing, followed by the creation and real-time updating of user profiles. VitaCompass offers personalized health and fitness tracking, enabling users to set achievable goals and monitor progress towards better well-being. The platform integrates features such as exercise tracking, nutrition monitoring, and goal setting to empower individuals to make informed lifestyle choices. VitaCompass emphasizes user engagement and feedback mechanisms to continuously improve user experience, while prioritizing privacy and scalability for optimal performance. This paper provides a detailed overview of VitaCompass's design and features, highlighting its commitment to user-centricity and technological innovation in promoting healthier lifestyles.**

Keywords—Health Behaviour,Data Analytics , Personalised Nutrition , User Experience(UX) , Nutrional Science .

I. INTRODUCTION

Introducing "VitaCompass," a groundbreaking initiative poised to revolutionize the approach to health and well-being. In today's fast-paced society, characterized by busy schedules and the prevalence of processed foods, the demand for a comprehensive tool to guide and monitor nutritional and fitness choices has never been more urgent. VitaCompass not only empowers users to take charge of their health but also provides a holistic approach to wellness by considering factors such as age, weight, activity levels, dietary preferences, and even medical history. With its intuitive interface and personalized recommendations, VitaCompass serves as a trusted advisor, helping individuals make informed decisions about their lifestyle habits. Additionally, VitaCompass fosters a sense of community by connecting users with like-minded individuals, sharing experiences, and providing support along their health journey. Whether it's tracking daily nutrition intake, setting fitness goals, or accessing expert advice, VitaCompass is the ultimate companion for those seeking to achieve optimal health and vitality in today's fast-paced world.

II. RELATED WORK

Vita Compass, a project aimed at promoting public health awareness, providing a convenient tool for users to monitor and analyze their food intake, it empowers individuals to make informed choices about their diet, aligning with VitaCompass's objective of promoting healthier lifestyles. Through user acceptability testing, VitaCompass demonstrates high usability and functionality, indicating its potential to improve users' comprehension of nutrition and dietary habits, contributing to long-term health and wellness. The modern world faces a profound challenge as lifestyles increasingly trend toward sedentary routines and diets replete with processed foods. This shift has led to a surge in obesity rates and related health concerns, demanding innovative solutions to combat this global epidemic. The "VitaCompass" project seeks to address this issue by offering a personalized, user-centric platform that empowers individuals to take control of their health and nutrition. Despite the availability of digital health and wellness tools, existing solutions often lack the ability to consider an individual's unique attributes such as age, weight, and activity levels when offering guidance. This deficiency hinders the effectiveness of these platforms in motivating users to make healthier choices. "VitaCompass" aims to bridge this gap by providing tailored nutritional and fitness recommendations, making it a powerful tool for individuals looking to improve their vitality and overall health. This project recognizes that a one-size-fits-all approach is no longer sufficient in the quest for healthier lifestyles, and it strives to offer an accessible, informative, and user-friendly solution to a growing global health crisis.

The "VitaCompass" project is essential to address the growing global health crisis related to poor nutrition and sedentary lifestyles. With rising rates of obesity and associated health conditions, there is an urgent need for a user-friendly, personalized nutrition and fitness tracking platform. VitaCompass empowers individuals to make informed dietary choices, set achievable goals, and monitor their progress toward better health. The major motivation for the project comes from the need for users to get a simple and easy to use interface that helps them track their daily fitness goals.

III. PROPOSED METHODOLOGY

The primary objective of the proposed VitaCompass system is to transform the way individuals oversee their dietary intake and fitness objectives by delivering an inclusive and user-friendly platform that unifies advanced data analytics, nutritional science, and fitness-monitoring technologies. The system commences with an exhaustive process of gathering requirements, partnering closely with health and nutrition specialists to pinpoint vital user data, dietary guidelines, and fitness parameters that are indispensable for precise tracking and assessment. The development of a robust architecture is an integral component of the system design phase, which outlines the various components and their interactions. This

design incorporates modules for user profiling, nutritional data tracking, exercise routines, and community interaction, ensuring a seamless integration of data across the platform's different functionalities. A critical component of the system is its highly advanced database design, which has been optimized to effectively store and manage a wide variety of data types. The organized database comprises user profiles that include demographic information, dietary preferences, health conditions, and fitness objectives. Additionally, the systematic storage of nutritional data, including calorie intake, macronutrient ratios, and micronutrient levels, enables precise tracking and analysis.

The system utilizes not only individual user data but also combines aggregated information derived from health and nutrition professionals, scientific research, and nutritional databases to provide well-founded recommendations and tailored insights. The application of advanced algorithms and machine learning methodologies enabled the analysis of user data, the detection of patterns, and the formulation of practical recommendations designed to enhance dietary practices and attain fitness objectives. VitaCompass places significant importance on community involvement and assistance by providing functionalities that allow users to connect with individuals who share similar interests, post advancement reports, and obtain informative materials on nutrition, physical fitness, and overall well-being. This platform nurtures an environment of encouragement and cooperation where users can grow, acquire knowledge, and achieve their health objectives. Overall, VitaCompass aims to empower individuals with the knowledge, tools, and support they need to make informed decisions about their nutrition and fitness, leading to improved health outcomes and quality of life.

Steps followed:

1) **Requirements Gathering:** Collaborated with health and nutrition experts to identify essential user data, nutritional guidelines, and fitness parameters. Conducted user surveys and research to understand the preferences and expectations of potential users.
2) **System Design:** Created a system architecture that outlined the platform's components and interactions. Designed the user interface for both web and mobile versions, focusing on user-friendliness and accessibility. Developed algorithms for personalized recommendations and data analysis.
3) **Database Design:** Developed a structured database to store user profiles, nutritional data, exercise routines, and community interactions. Implemented data security and privacy measures to protect sensitive user information.
4) **Development:** Developed web and mobile applications using modern web technologies and frameworks. Implemented features for user profile creation, nutritional tracking, exercise logging, goal setting, and community engagement. Integrated real-time calculations and feedback mechanisms.
5) **Educational Content Creation:** Curated or created educational content, including articles, videos, and guides on nutrition, health, and fitness. Ensured content was relevant, accurate, and engaging.
6) **Personalization and Behavioral Strategies:** Developed algorithms that used user data to offer personalized nutritional and fitness recommendations. Incorporated behavioral change strategies, such as gamification, reminders, and notifications.
7) **Testing and Quality Assurance:** Conducted thorough testing to identify and fix bugs, ensuring a stable and reliable platform. Performed usability testing with potential users to gather feedback on the user experience.
8) **Deployment and Accessibility:** Deployed the web application and mobile app on appropriate platforms. Ensured the platform was accessible through various devices and operating systems.
9) **Data Analysis and Continuous Improvement:** Analyzed user data to gain insights into user behavior and preferences. Made iterative improvements to the platform based on user feedback and data analysis.
10) **Monitoring and Evaluation:** Monitored the platform's performance, user engagement, and user satisfaction. Conducted regular evaluations to assess the platform's impact on users' health and well-being.

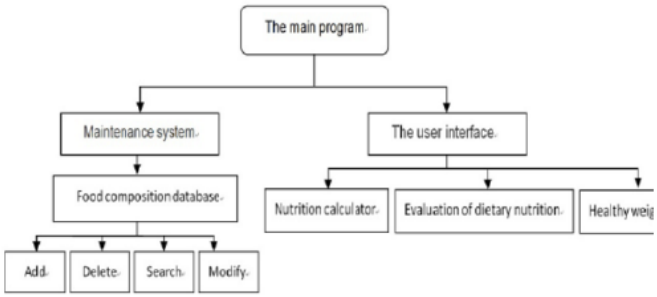


Fig. 1. Workflow of System

IV. IMPLEMENTATION AND RESULTS

The VitaCompass project aims at fulfilling a simple aim for users. A nutrition app that can give you a plain and direct count of the calories that you should have in a day and the calories that you are taking. This helps in easy tracking of daily goals. For this, it is very necessary to have a simple and plain UI. Thus the VitaCompass project suggest the following flow for easy access.

User Login: The flowchart begins with the User Login process, where the user enters their credentials (username and password) to access the calorie-tracking web application. Credentials Validation: After the user submits their login credentials, the system checks whether the provided credentials are valid or not.

Dashboard Display: If the credentials are valid, the system displays the user's dashboard, which provides an overview of their calorie tracking and fitness-related information. Meal Logging: The system checks if the user has logged in a meal. If a meal is logged in, the system proceeds to calculate the calories consumed during that meal. If a meal is not logged in, the system prompts the user to log a meal.

Exercise Logging: After handling meal logging, the system checks if the user has logged any exercise activity. If exercise is logged in, the system calculates the calories burned during the exercise session and updates the user's progress. If no exercise is logged, the system prompts the user to log an exercise activity.

Progress Tracking: Following meal and exercise logging, the system updates the user's progress based on the calories consumed and calories burned. It may also display the progress to the user. Achievement Notification: If the user's progress indicates that they have achieved their set goals (e.g., weight loss goal, fitness target), the system sends an achievement notification to congratulate the user.

Error Handling: In case the credentials provided by the user

are not valid, the system displays an error message indicating the invalid credentials and prompts the user to retry the login process.

User Logout: Finally, after completing the necessary actions and interactions, the user can choose to logout from the
The UI serves as the gateway to a world where technology seamlessly intertwines with fashion, offering users a visually appealing and user-friendly platform. The design principles prioritize simplicity, ensuring that users can effortlessly navigate through various features such as personalized outfit recommendations, virtual try-ons, and wardrobe management. The User Interface (UI) of our AI Fashion Stylist system has been meticulously designed to provide an intuitive and engaging experience for users seeking personalized fashion recommendations and styling assistance. A glimpse of our user interface is shown below.

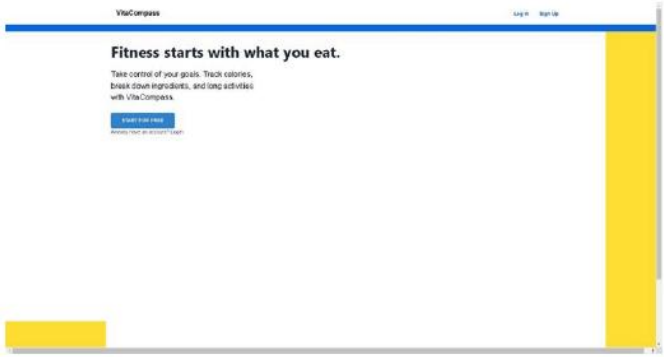


Fig. 2. Home Page

The Home Page serves as the central hub for the login and the sign of a new or existing user.

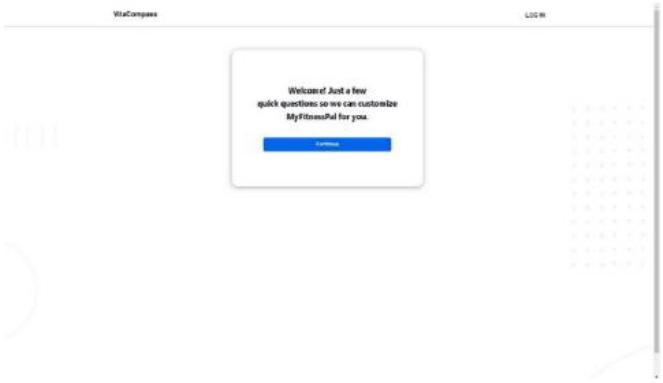


Fig. 3a. Sign Up Page

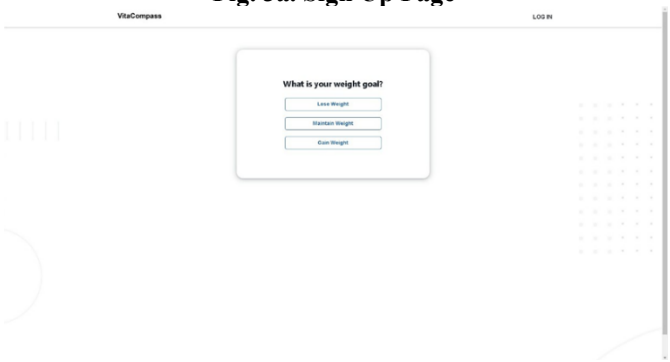


Fig. 3b. Weight Goal

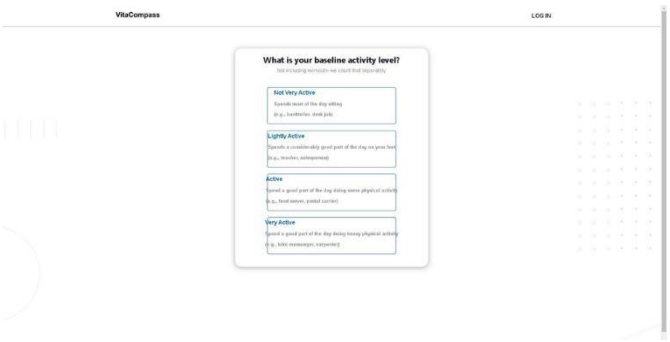


Fig. 3c. Activity Level

After all the personal details have been entered, the user can set a username and password for his account that cab ne used for the login and signup. Fig 4a and 4b demonstrate the same.

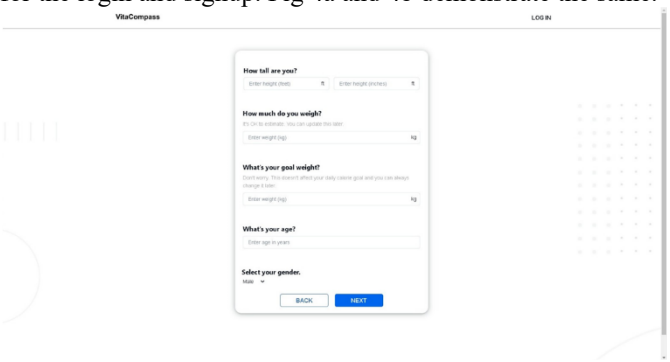


Fig. 4a. Description of User

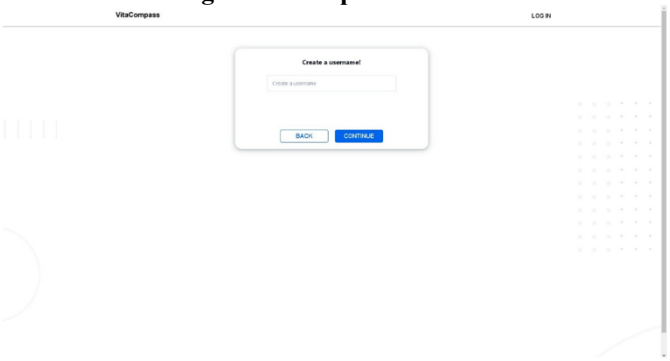


Fig.4b.SetUsername

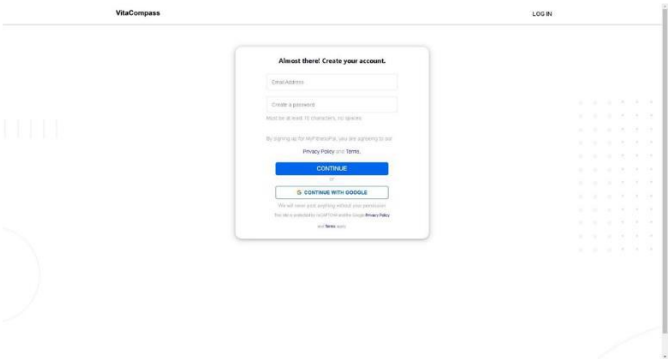


Fig. 4c. Enter Email Id and Password

After the daily calorie goal is counted and set, the user can then add the food that he eats in a day from the food diary, and subsequently the calorie count will be taken into account. Thus, it makes it very easier for the user to keep track of his daily goals.

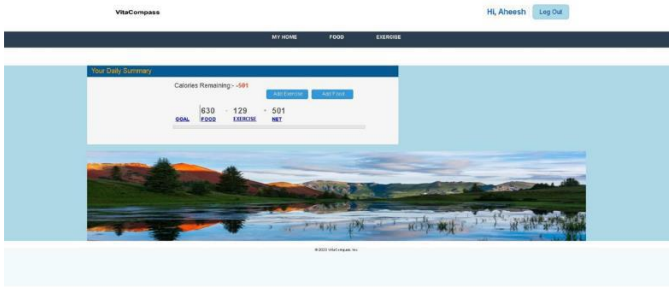


Fig. 4d. Daily Summary Page

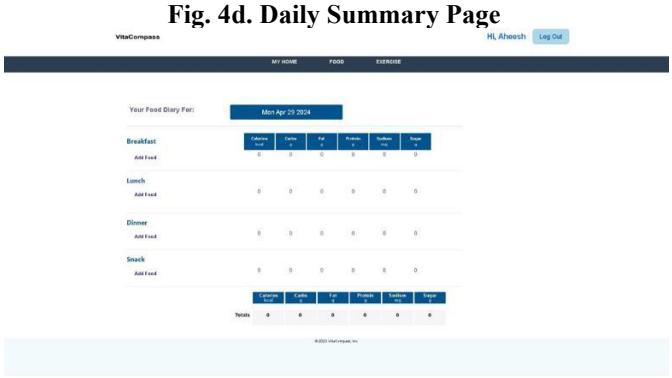


Fig. 4.e. Food Diary

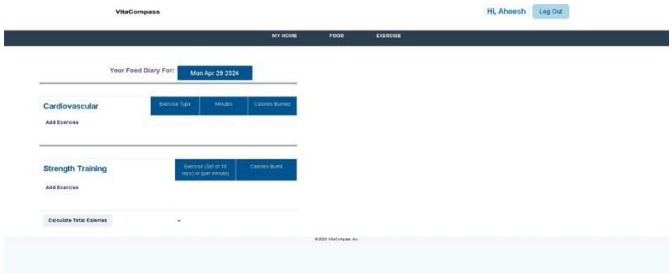


Fig. 4.f. Exercise Diary

V. CONCLUSION

VitaCompass represents a comprehensive solution for individuals seeking to manage their nutritional intake and achieve their health and fitness goals effectively. By seamlessly integrating user input metrics, real-time calorie tracking, personalized feedback, and insightful reporting features, VitaCompass empowers users to make informed decisions about their dietary habits and lifestyle choices. In conclusion, VitaCompass offers a user-centric approach to nutrition management, leveraging technology to streamline the process of tracking food consumption, monitoring caloric intake, and receiving personalized recommendations. Through its intuitive interface and robust functionality, VitaCompass facilitates a seamless user experience, enhancing engagement and adherence to health goals.

The future scope of VitaCompass would be:

1. Integration with Wearable Devices: VitaCompass can extend its capabilities by integrating with a broader range of wearable fitness devices, such as smartwatches and fitness trackers, to provide real time data syncing and enhanced exercise tracking.
2. AI-Driven Personalization: Further development of artificial intelligence can lead to even more personalized recommendations and real-time adjustments based on users' changing health goals and performance.
3. Enhanced Nutritional Analysis: VitaCompass can offer advanced nutritional analysis, including ingredient recognition from food images, to provide a more comprehensive understanding of users' dietary habits.
4. Healthcare Provider Collaboration: Collaborating with healthcare providers to share user data securely, enabling remote monitoring and personalized health recommendations based on medical history and conditions.
5. Global Expansion: Expanding VitaCompass to cater to a global audience by providing multilingual support and incorporating region-specific dietary and exercise preferences and recommendations.

REFERENCES

[1] Haotian Jiang, Jammes Starkman, Menghan Liu, and Ming-Chun Huang, "Food Nutrition Visualization on Google Glass," IEEE Consumer Electronics Magazine, May 2018. 50

[2] Arnel B. Ocay, Jane M. Fernandez, Thelma D. Palaoag, "NutriTrack: Android-based Food Recognition App for Nutrition Awareness," 3rd IEEE International Conference on Computer and Communications, 2017.

[3] Agapito G., Calabrese B., Guzzi P. H., Cannataro M., Simeoni M., Car´e I., Lamprinoudi T., Fuiano G., Pujia A., "DIETOS: a recommender system for adaptive diet monitoring and personalized food Suggestion," Fourth International IEEE Workshop on e-Health Pervasive Wireless Applications and Services, 2016.

[4] S. Jasmine Minija, W.R. Sam Emmanuel, "Food image Classification using Sphere Shaped-Support Vector Machine," Proceedings of the International Conference on Inventive Computing and Informatics (ICICI 2017), IEEE Xplore Compliant.

[5] Natnicha Suthumchai, Sirin Thongsukh, Pacharamai Yusuksataporn, Songsri Tangsripairoj, "FoodForCare: An Android Application for Self-Care with Healthy Food," Fifth ICT International Student Project Conference (ICT-ISPC), 2016.

[6] Neha Gaur, Archana Singh, "Recommender System – Making Lifestyle Healthy Using Information Retrieval," 2nd International Conference on Next Generation Computing Technologies (NGCT - 2016) Dehradun, India.

[7] Ridho Rahman Hariadi, Wijayanti Nurul Khotimah, Eko Adhi Wiyono, "Design and Implementation of Food Nutrition Information System using SURF and FatSecret API."

[8] Diptee Kumbhar, Prof. Sarita Patil, "Mobile Cloud based System Recognizing Nutrition and Freshness of Food Image," International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017).

[9] Dr. Anne Moorhead, Dr Raymond Bond, Dr Huri Zheng, "Smart Food: Crowdsourcing of experts in nutrition and non-experts in identifying calories of meals using smartphone as a potential tool contributing to obesity prevention and management," IEEE International Conference on Bioinformatics and Biomedicine (BlBM), 2015

❑❑❑❑❑❑

AET 030

Automatic Pill Dispenser

Isha Pandey
Department of Biomedical Engineering,
Vidyalankar Institute of Technology,
Wadala(E), Mumbai, 400031

Md Sameer Khan
Department of Biomedical Engineering,
Vidyalankar Institute of Technology,
Wadala(E), Mumbai, 400031

Prajay Bhalerao
Department of Biomedical Engineering,
Vidyalankar Institute of Technology,
Wadala(E), Mumbai, 400031

Prof. Priyanka Shrivastava (Assistant Professor)
Department of Biomedical Engineering,
Vidyalankar Institute of Technology,
Wadala(E), Mumbai, 400031

***Abstract*— The Automatic Pill Dispenser is an essential healthcare tool created to streamline cure administration. This creative electronics ensures appropriate and correct medicine dispensing while underrating the risk of human wrong. With a convenient connect, the dispenser specifies a hassle-free answer for things needing to take diversified drugs. Its exact and regulated pill discharging order guarantees that the right dosage is brought at the official time of region. Safety devices are integrated for fear that double drug, reinforcing the overall dependability and peace of mind for sufferers and caregivers. This instrument is a valuable form for improving cure devotion, specifically with those with complex cure regimes. This abstract focal point the canter features of a mechanical medicine dispenser, stressing allure role in simplifying drug administration and reducing the potential for drug-connected.**

***Keywords*— Arduino, Dispenser**

INTRODUCTION

Medicine Dispenser regimens easier for those with them, especially those who need to strictly stick to a medicine schedule or who might hav e trouble remembering to take their pills. By automating the process of administering tablets at predetermined times, this cutting edge technology makes sure that patients receive the proper amount of medication at the right time. An automatic pill dispenser's main parts usually consist of a safe place to store medication, a timer or scheduling system, and mechanisms the at allow pills to be dispensed on a predetermined schedule. Additionally, some sophisticated versions could offer connection options, alerts, o r notifications to notify users or carers when medicine needs to be taken or when dosages have been An Automatic Pill Dispenser is an advanced gadget that makes managing conflict The advantages of an automatic pill dispenser are numerous. It has the potential to greatly minimise the risk of medication mistakes, increase drug adherence rates, and bring peace of mind to both users and carers. By automating the medication administration procedure, these gadgets enable people to better manage their health and keep their freedom in everyday life. Furthermore, Automatic Pill Dispensers frequently have user-friendly interfaces and customisable settings, allowing for personalised medicine regimens customised to individual requirements. This adaptability makes them appropriate for a wide range of users, including seniors, those with chronic conditions, and those undertaking complex treatment regimens.

PURPOSE OF AUTOMATIC PILL DISPENSER

An Automatic Pill Dispenser searches out streamline medication administration for things accompanying complex cure regimens. By automating the process of discharging pills at named opportunities, Improve Medication Adherence: By providing prompt reminders and guaranteeing doses are captured as recommended, mechanical pill dispensers help consumers persist path accompanying their treatment plans, eventually reconstructing drug devotion rates. Reduce Medication Errors: By minimizing the risk of misplaced doses or unforeseen double doses, mechanical pill dispensers advance more reliable cure practices and lower the likelihood of cure mistakes. Enhance Independence: Automatic medicine dispensers authorize users to survive their drugs alone, lowering reliance on caregivers and advancing a sense of independence. Provide Peace of Mind: With visage to a degree alarms, notifications, and customizable scenes, mechanical dose dispensers offer mental freedom to both consumers and caregivers, experienced that drugs are being captured as prescribed.

FEATURSE COVERED

Automated Medication Dispensing: Automatic tablet dispensers mechanize the process of dispensing drug in accordance with preset schedules, lowering the need for manual intervention and underrating the risk of cure mistakes. Programmable Scheduling: Users can program the dispenser to discharge cures at particular times of the era, guaranteeing constant adherence to cure procedures and lowering the likelihood of misplaced doses. Multiple Medication Management: Many mechanical tablet dispensers are worthy managing diversified cures together, accommodating complex situation regimes and reconstructing overall medication devotion. Secure Medication Storage: Automatic dose dispensers usually feature secure storage compartments accompanying locking machines, looking after medications from unjustified approach and hampering accidental stuff or misuse. Audio/Visual Reminders: Some mechanical tablet dispensers incorporate visual and audio entertainment transmitted via radio waves and optical warnings to alert users when it's occasion to take their cure. These notices help improve devotion, particularly for things with thought issues or intelligent deteriorations. Dosage Accuracy: Advanced dispensers ensure correct operating of cures in the correct dosage, underrating the risk of portion of drug or other consumable mistakes and ensuring patient security. Remote Monitoring: Certain mechanical dose dispensers offer detached monitoring efficiencies, admitting caregivers or

healthcare providers to path medication devotion and sustain announcements in real opportunity. This feature reinforces responsibility and facilitates appropriate invasion when wanted. Backup Power Options: To prevent disruptions in medication administration all the while capacity outages, automatic dose dispensers frequently involve backup capacity alternatives in the way that battery auxiliary or handy capacity packs. User-Friendly Design: Automatic pill dispensers are created accompanying consumer convenience in mind, recommending instinctive interfaces, smooth-to-use controls, and clear instructions to guarantee approachability for consumers of all ages and abilities. Comprehensive Reporting: Some mechanical tablet dispensers produce inclusive reports on medication devotion, portion of drug or other consumable annals, and refill rank. These reports determine valuable intuitions for both consumers and healthcare providers to monitor progress and regulate situation plans as needed.

INTERNET OF THINGS(IoT) TECHNOLOGY

The Internet of Things (IoT) encompasses devices like cars and home appliances equipped with electronics, software, and connectivity to communicate via the internet. These devices have unique identifiers (UIDs) and can exchange vast amounts of data across networks.

The IoT ecosystem consists of web-enabled smart devices equipped with sensors and processors to gather and transmit data. This data is sent or received via an IoT gateway and analysed in the cloud. These devices operate autonomously, requiring minimal human intervention for setup.

ARDUINO TECHNOLOGY

An Arduino board, built around a microcontroller, utilizes Harvard architecture, separating program code and data into distinct memories: flash memory for code and data memory for stored information. Arduino technology allows for direct program uploads without requiring a hardware programmer, making it popular for controlling and communication devices. It's an open-source electronics platform.

COMPONENTS

-Servo Motor: Servo motors are a necessary in science, automation, and differing photoelectric projects on account of their exact control over angular position. They contain a narrow DC engine in a pair with gears and a response control arrangement. This response structure allows servo motors to correctly uphold their position or alternate to distinguishing angles based on the recommendation signal they sustain. Servo motors are usually secondhand in applications to a degree electronic weaponry, detached-controlled bicycles, camcorder gimbals, and model airplanes. They offer benefits in the way that high twisted item, smooth movement, and smooth interfacing accompanying microcontrollers like the ----Arduino Uno. -Arduino Uno: The Arduino Uno is a flexible microcontroller board that serves as the intellect of many DIY VCRs projects. It is based on the ATmega328P microcontroller and faces a set of mathematical and parallel recommendation/output pins, in addition to onboard elements like timers, sequential ideas interfaces, and PWM (Pulse Width Modulation) outputs. The Arduino Uno is known for allure restraint, ease valuable, and thorough community support. It may be a computer utilizing the Arduino Integrated Development Environment (IDE), that provides a handy connection for paper, assembling, and uploading code to the board. With an allure expansive range of possible libraries and shields, the Arduino Uno is appropriate for prototyping, one who collects specimen's projects, instructional purposes, and even monetary applications.

-IR Sensor: Infrared (IR) sensors are photoelectric instruments that discover the color of blood radiation discharged by objects. They are usually secondhand for closeness sensing, object discovery, motion discovery, and handheld control requests. IR sensors typically incorporate an IR transmitter (LED) and an IR recipient (photodiode) established close by physically to each added. When an object comes inside range, it indicates shade resembling such a color light back to the sensor, causing a change in the recipient's profit signal. This change is therefore elucidated by the microcontroller, admitting bureaucracy to discover the presence or lack of objects. IR sensors are usual instability wholes, automatic doors, line-following androids, and detached controls for TVs and added machines.

-Jumper Wires: Jumper wires are essential components for establishing energetic networks 'tween various parts in photoelectric circuits. They usually contain flexible, protected strand accompanying connectors (pins) on two together ends. Jumper wires come in miscellaneous lengths, banner, and styles, containing male-to-male, male-to-female, and female-to-female configurations. They are usually used on breadboards, that are prototyping manifestos that compensate smart and temporary congregation of photoelectric circuits. Jumper wires allow engineers, hobbyists, and students to experiment accompanying various track configurations, especially to problems, and original their projects outside the need for soldering.

-(Green/Brown) PCB: PCBs, or Printed Circuit Boards, are essential elements in new electronics. They supply a durable manifesto for escalating electronic components and forming energetic relations between them. PCBs usually contain non-conductive material (for instance, fine blown glass or epoxy) accompanying conductive pathways (law enforcement officer traces) carved or impressed onto their surface. The green color frequently guides PCBs emanates the solder mask used to assure the law enforcement officer traces from decay and speed soldering during congregation. Brown PCBs, in another way, can display different fabrics or production processes. PCBs are planned to utilize specialized operating systems and made utilizing miscellaneous techniques in the way that reproduction, grinding, or publication. They are used in and of course range of photoelectric instruments, containing computers, smartphones, healing instruments, automotive methods, and services electronics.

AUTOMATIC PILL DISPENSOR DESCRIPTION

A mechanical medicine dispenser is a complex maneuver designed to assist things in directing their medication consumption more capably and accurately. It usually exists of several key parts and looks tailored to meet the needs of consumers accompanying varying drug procedures and preferences. At the heart of the mechanical tablet dispenser is a programmable whole that controls the discharging of medications in accordance with preset schedules. Users can

recommend their drug schedule into the device, designating the portion of drug or other consumable and timing each drug. This programmable functionality guarantees that cures are dispensed at the right time and in the correct portion of drug or other consumable, lowering the likelihood of misplaced doses or cure errors. The material design of the mechanical pill dealer frequently includes diversified compartments or trays for packing different drugs. Each subdivision is described or color-coded for smooth labeling, allowing consumers to systematize their medications in accordance with their situation plan. Some dispensers may more feature lockable compartments for fear that unauthorized approach and guarantee medication protection. To alert consumers when it's time to take their cure, mechanical pill dispensers usually combine reminder plans. These notices can take the form of perceptible alarms, visual signs (to a degree flashing lights or display ideas), or even smartphone announcements sent by way of Bluetooth or Wi-Fi relatedness. These reminders help consumers persist in track accompanying their drug regimen and advance cure adherence. Advanced mechanical medicine dispensers may again offer supplementary features to embellish consumer convenience and security. For example, few maneuvers include included sensors to discover whether drugs have happened dispensed and taken as due. This authorizes caregivers or healthcare providers to monitor medication devotion by chance and intervene if essential. Other face may contain fill reminders, portion of drug or other consumable following, and medication devotion reports produce by the ploy's operating system. Overall, mechanical pill dispensers play a critical act in simplifying medication administration for things with complex situation menus or those who demand help with drug devotion. By automating the process of dispensing and recalling consumers to take their medications, these ploys help raise medication devotion, humble the risk of medication wrongs, and embellish overall health effects.

PURPOSE OF AUTOMATIC PILL DISPENSOR

Mechanical pill dispensers is versatile, discussing several critical challenges in drug management. These ploys play a pivotal duty in reconstructing medication devotion by administering medications at due times and providing keepsakes, through helping consumers stay consistent accompanying their situation plans. Furthermore, automatic medicine dispensers improve medication security by accurately operating the correct portion of drug or other consumable and reducing the risk of unexpected overdoses. They too promote liberty and autonomy for consumers the one require help accompanying medication administration, allowing ruling class to self-execute their medications optimistically. Caregivers benefit from the mental freedom provided by mechanical pill dispensers, as these designs shorten medication administration and allow remote listening of adherence. Moreover, mechanical tablet dispensers facilitate never-ending ailment management by systematizing multiple cures and advancing healthy stale by addressing challenges that had connection with thought and dexterity. Overall, these maneuvers be a part of invaluable finishes in promoting drug devotion, safety, and overall fitness consequences for users and caregivers alike.

IMPACT OF AUTOMATIC PILL DISPENSOR

Automatic medicine dispensers have a profound affect cure management, reinforcing devotion rates, safety, and character of history for users while lessening caregiver burden. By providing convenient keepsakes, accurate drug, and advancing independence, these maneuvers enhance better health effects, reduced healthcare costs, and improved ability utilization. Additionally, they support maturing people and facilitate incessant ailment management, making ruling class priceless tools in up-to-date healthcare.

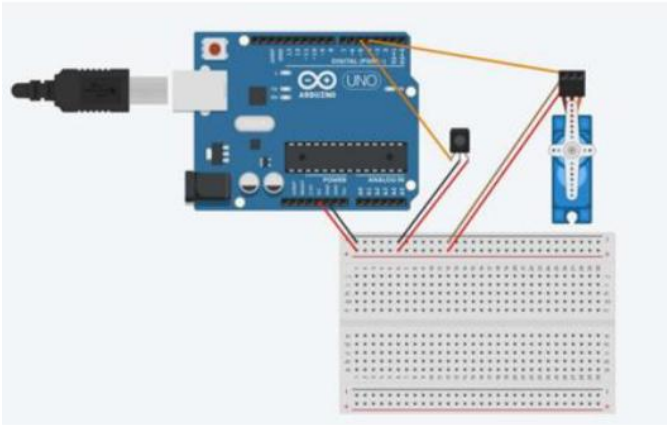
HARDWARE INTEGRATION AND WORKING

Programming: Users recommend their drug schedule into the dispenser, designating the type of drug, portion of drug or other consumable, and timing each application. This set up typically happens through an connect on the dealer itself or by way of a companion smartphone app or calculating operating system. Storage: Medications are intoxicated into individual compartments or trays within the dealer. These compartments are frequently described or color-coded for smooth labeling and arrangement. Some dispensers feature lockable compartments to prevent unconstitutional approach. Dispensing: At the due time, the dealer activates to administer the appropriate drug from the designated section. This process can include alternating a carousel or platter to join the picked compartment accompanying a operating means, such as a slope or gap. Alerts and Reminders: Before or all along the dispensing process, the dealer grant permission to determine audible, able to be seen with eyes, or mathematical alerts to inform the user that it's period to take their drug. These notices help guarantee that users forbiddance miss their due doses. User Interaction: Upon taking the alert, the user communicates accompanying the dealer to access the dispensed drug. This grant permission includes opening a section or recapturing the drug from a designated administering extent. Tracking and Monitoring: Some mechanical medicine dispensers include pursuing and listening visage to record medication devotion and habit patterns. This dossier can be stocked inside or communicated to a companion app or cloud-located manifesto for review for one user or healthcare wage earner. Refilling and Maintenance: Periodically, the consumer or nurse refills the dispenser accompanying cures as wanted. Additionally, the dealer may demand routine perpetuation, in the way that battery substitute or cleansing, to guarantee proper performance.

CIRCUIT DIAGRAM

Setting up the Hardware. -Place the Arduino Uno on the breadboard. -Connect the capacity (5V) and ground (GND) pins of the Arduino to the capacity rails on the breadboard. -Connect the IR sensor to the breadboard. Connect the VCC attach to 5V, GND to GND, and OUT to a mathematical recommendation fix the Arduino (for example, attach 2) -Connect the servo engine to the breadboard. Connect the crimson line to 5V, the dark strand to GND, and the tangerine User Interface Software: • Custom program for directing consumer interactions and configuring the discharging schedule. • It allows consumers to start the drug timings and alarm backgrounds. Data Management Software: • Software for recording and directing the operating annals and consumer interplays accompanying the

device. • It admits for pursuing the cure schedule and guaranteeing prompt refills Setting up the Hardware. -Place the Arduino Uno on the breadboard. -Connect the capacity (5V) and ground (GND) pins of the Arduino to the capacity rails on the breadboard. -Connect the IR sensor to the breadboard. Connect the VCC attach to 5V, GND to GND, and OUT to a mathematical recommendation fix the Arduino (for example, attach 2) - Connect the servo engine to the breadboard. Connect the crimson line to 5V, the dark strand to GND, and the tangerine User Interface Software: • Custom program for directing consumer interactions and configuring the discharging schedule. • It allows consumers to start the drug timings and alarm backgrounds. Data Management Software: • Software for recording and directing the operating annals and consumer interplays accompanying the device. • It admits for pursuing the cure schedule and guaranteeing prompt refills.



Setting up the Hardware. -Place the Arduino Uno on the breadboard.

-Connect the capacity (5V) and ground (GND) pins of the Arduino to the capacity rails on the breadboard.

-Connect the IR sensor to the breadboard. Connect the VCC attach to 5V, GND to GND, and OUT to a mathematical recommendation fix the Arduino (for example, attach 2)

- Connect the servo engine to the breadboard. Connect the crimson line to 5V, the dark strand to GND, and the tangerine User Interface Software: • Custom program for directing consumer interactions and configuring the discharging schedule. • It allows consumers to start the drug timings and alarm backgrounds.

Data Management Software: • Software for recording and directing the operating annals and consumer interplays accompanying the device. • It admits for pursuing the cure schedule and guaranteeing prompt refills.

DESIGN CONSIDERATION

Several considerations were addressed before initiating the design process for the medicine box. The device features a circular base with 30 fan-like blades forming compartments, each designed to house pills and capsules. The software is engineered for reliability, enabling recording of current medication and future dispensation.

A mechanical lock ensures security and tamper-proofing, while an LED display indicates operational status and provides instructions. The microcontroller interfaces with LED, stepper motor, alphanumeric keypad, LCD display, and speaker. Visual and audio notifications are incorporated, and proper storage is ensured to maintain medicine quality.

FUTURE SCOPE

The future excuse for an Automatic Pill Dispenser following an Alarm is hopeful, accompanying potential progresses in smart learning unification, hindering that relevance to smartphones and AI algorithms for materialize reminders. Enhanced care face like biometric ratification manages better drug safety, while physiognomy for drug stock following and health hearing consolidation would offer all-encompassing drug administration. User information augmentations, dossier learning of probable reasoning skills, and disconnected patient listening control further increase position plans and better patient effects. Efforts to better general approachability and affordability would address unmet healthcare needs and disclose new markets for this imaginative resolution. Also, few of future progress are in this manner: - 1.Integration of Advanced Sensors: Incorporating supplementary sensors in a way that weight sensors or RFID tags can further embellish the dealer's use. These sensors can help monitor medicine stock levels, track cure habit patterns, and determine actual opportunity feedback to consumers and caregivers. .2. Smartphone Connectivity and App Integration: Adding Bluetooth or Wi-Fi connectedness to the dealer admits for seamless ideas accompanying smartphones and additional tools. Developing a loyal smartphone application can authorize consumers to by chance monitor cure consumption, receive portion of drug or other consumable warnings, and regulate scenes on-the-go. .3. Artificial Intelligence and Machine Learning: Implementing AI algorithms can authorize the dispenser to gain and accustomed to consumer presence and cure adherence patterns over period. By resolving dossier composed from sensors and consumer interactions, AI can specify embodied pieces of advice and interferences to enhance medication administration. 4.Enhanced User Interface: Improving the program that controls display accompanying touchscreen displays, voice acknowledgment, and intuitive controls can improve utility and approachability for consumers of all ages and skills Paper id Journal of Emerging Technologies and Innovative Research (JETIR) www.jetir.org 31 5.Integration with Healthcare Systems: Collaborating accompanying healthcare providers and pharmacies to merge the dealer accompanying photoelectric health record wholes can organize drug administration and simplify communication 'tween inmates, caregivers, and healthcare artists.

CONCLUSION

The utilization of SMD is poised to significantly enhance medication adherence, particularly among elderly patients and those with chronic conditions, leading to improved treatment effectiveness. Insurance companies stand to benefit greatly from SMD, as it reduces the risk of catastrophic accidents resulting from missed doses or incorrect dosages, thereby promoting healthier lifestyles for their customers.

Moreover, the user interface, consistent across all operating systems and devices, is designed to be clear, user-friendly, intuitive, and accessible, even for elderly patients. The design is adaptable, allowing users to add more containers, and remains open to future enhancements and improvements.

REFERENCES

[1] E. 1. R. Hemanth Kumar 1GV17ME014 S. Kishore Raj 1GV17ME020 Wasi Ulla Khan Junaid 1GV17ME036 S. Subashkaran "Design and Prototype of Smart Automated Pill Dispenser" VISVESVARAYA TECHNOLOGICAL UNIVERSITY, Jnana Sangama, Belagavi- Oorgaum, Kolar Gold Fields – 563120 2017 – 2021

[2] F. 2.P. H. Tsai, C. Y. Yu, C. S. Shih, Member, IEEE, and J. W. S. Liu, Fellow, IEEE "Smart Dispensor: Design, Architecture and Implementation Institute of Information Science, Academia Sinica Technical Report TR-IIS-08-010

[3] G. 3.Diaa Salama Abdul Minaam, and Mohamed Abd-ELfattah, "Smart drugs: Improving healthcare using Smart Pill Box for Medicine Reminder and Monitoring System", Future Computing and Informatics Journal, Volume 3, Issue 2, 2018, ISSN 2314- 7288.

[4] H. 4.Bachelor's Thesis | Abstract Turku University of Applied Sciences Bachelor of Engineering, Information and Communication Technology 2023 | 30

[5] Wissam Antoun, Ali Abdo, Suleiman Al-Yaman, Abdallah Kassem, Mustapha Hamad and Chady El-Moucary, "Smart Medicine Dispenser (SMD)", 2018 IEEE 4th Middle East Conference on Biomedical Engineering (MECBME), pp. 20-23, 2018

[6] V. Dhukaram and C. Baber, "Elderly Cardiac Patients' Medication Management: Patient Day-to-Day Needs and Review of Medication Management System," 2013 IEEE International Conference on Healthcare Informatics, pp. 107-114, 2013]

[7] https://www.researchgate.net/publication/343811187_ Automatic_Medicine_Dispenser_using_IoT

[8] Diaa Salama Abdul Minaam, and Mohamed Abd-ELfattah, "Smart drugs: Improving healthcare using Smart Pill Box for Medicine Reminder and Monitoring System", Future Computing and Informatics Journal, Volume 3, Issue 2, 2018, ISSN 2314- 7288.

[9] https://ieeexplore.ieee.org/author/3708600924

❑ ❑ ❑ ❑ ❑ ❑

AET 031

Patient Monitoring System for Comatose Patient

Prof. Neelam Punjabi & Komal Lawand
Department of Biomedical Engineering,
Vidyalankar Institute of Technology,
Wadala (East), Mumbai – 400037

Sahil Aute1, Tanmay More, Raviraj Suralkar
Department of Biomedical Engineering,
Vidyalankar Institute of Technology,
Wadala (East), Mumbai – 400037

***Abstract* – The project is an Internet of Things (IoT) based environmental monitoring system that measures temperature, humidity, EOG and flex sensor data and sends it to the cloud platform Things Board for storage and analysis. The system is built using an ESP8266 Wi- Fi module, an HTU21D temperature and humidity sensor, a flex sensor, and a buzzer. The project is significant as it enables remote monitoring of environmental data, which can be useful in various applications such as agriculture, weather forecasting, and indoor climate control. The use of IoT technology and cloud-based storage and analysis allows for real-time data acquisition and decision-making, which can have significant impacts on efficiency and resource management.**

I. INTRODUCTION

Data is an asset in this digital world, and enormous amounts of data were generated across all fields. In the healthcare industry, data refers to all patient-related data. A general architecture for disease prediction in the healthcare industry has been proposed here. One disease is the focus of many of the current models for each analysis. Examples include a cancer analysis, a liver analysis, and a skin disease analysis.

There is currently no system that can simultaneously analyze multiple diseases. As a result, our primary focus is on providing users with prompt and precise disease predictions based on the symptoms they experience and the diagnosis they are given. Therefore, we propose a method that makes use of Streamlit to predict a variety of diseases. We will conduct diabetes and Parkinson's disease analyses in this system.

Later a lot more sicknesses can be incorporated. We will use machine learning algorithms to set up multiple disease prediction systems. Python pickling is used to preserve the model's behavior.

This system analysis is crucial because it allows for a more accurate and efficient disease detection by taking into account all of the factors that contribute to the condition.

The last model's conduct will be saved as a python pickle record.

A ton of examination over existing frameworks in the medical care industry thought about just a single sickness at a time. For instance, one framework is utilized to break down Diabetes, one more is utilized to dissect diabetes retinopathy, and one more framework is utilized to anticipate Coronary illness. Most extreme frameworks center around a specific infection. Numerous models must be used by an organization in order to analyze the health reports of its patients.

The current system's method is useful for analyzing specific diseases only. A user of the multiple diseases prediction system can examine multiple diseases on a single website. The user doesn't have to go to different places to figure out if they have a certain disease or not.

The user of the multiple diseases prediction system simply needs to select the specific disease's name, enter its parameters, and then click the submit button. Invoking the appropriate machine learning model will result in a prediction of the output and its display on the screen.

So, we have made a 3d model of the alimentary tract with blender and slicer 3D software, which we will use as a patient. Which can be used by medical students and they will also practice and learn endoscope procedure. We have made a prototype like an endoscope in which we have used rubber tube, bending sections and a small size LED CAMERA. If there is an ulcer in the alimentary tract, he can see it and treat it there. Where there are ulcers and wounds, the drug can be delivered so that the area heals quickly.

II. LITERATURE SURVEY

[1] **"A Real-Time Remote Monitoring System for Comatose Patients" Journal/Conference:** journal of physics conference series Year of Publication: 2019 Summarize in your own words: An effective healthcare system for monitoring the vital parameters of a patient in comatose stage is built using various sensors which coordinate with processor. This processor in turn processes this data and transfers this to the Wi-Fi module. Wi-Fi module then transmits the data via the cloud to the webpage. It also makes sure to store the previously read data in a proper form as a database for a future access. An alert Email system is also setup to counter emergency situations by sending an alert. This data can be viewed by logging on to the web-page paper.

[2] **Journal of Scientific and Engineering Research Year of Publication:**2020 Summarize in your own words: Embedded system is becoming an integral part of Engineering design process for efficient analysis and effective operation. From data analysis to hardware work, everywhere embedded products are the main interest because of its reliability and time bound perfection. There is not much time with anyone now a day to give enough in all aspects, so demand of embedded products which serve as we want is high on demand. Further it describes the design of an embedded system for the control of Temperature & Light intensity with continuous monitoring in a single system using sensors, microcontroller and LCD. It describes the controlling action incorporated in the hardware to control any device connected when specific conditions are met. Further set up is made such that data can be stored for future offline analysis. The hardware developed namely "Advanced Coma Patient Monitoring System" is very advanced product related to physical changes in body movement of the patient and gives Warning in form of alarm and display on the LCD in

less than one second time. It also passes a sms to a person sitting at the distant place if there exists any movement in any body part of the patient.

[3] **Sarah Wilson, Robert Clark Journal/Conference: Year of Publication:**2018 Summarize in your own words The Neurological Wake-Up Test (NWT) is recognized as the standard for evaluating brain injured patients, but it has contraindications for certain conditions. It can cause a stress response and raise intracranial pressure, but there's no clear evidence of harm. Multimodality neuromonitoring provides data but hasn't definitively improved patient outcomes or guided decisions. NWT frequency is uncertain, with one recommendation suggesting at least daily. Propofol is the preferred sedative, but dexmedetomidine is an alternative, especially for nonincubated patients or to prevent Propofol- Related Infusion Syndrome (PRIS). More research is needed to determine NWT's clinical utility, its ability to detect neurological issues, and establish guidelines based on injury type and patient population.

[4] **Journal/Conference:** Telemedicine is a valuable technology in linking clinicians with patients to ensure they make long-term lifestyle changes. It has significant benefits for medical office staff. This many times eliminates the burden of patient check-in and concentrates on higher-value tasks. With online visit capability, clinicians may care for their patients while still potentially assisting other affected practices. This also reduces distance limitations by exchanging information about a diagnosis, care, and disease prevention between the doctor and the patient through electronic means. The most extensive telemedicine application can get health coverage closer to people who live in rural areas where quality treatment is otherwise impossible to access. In recent years, this technology has been shown to increase the quality of healthcare facilities by allowing the exchange of information across many distant areas. It expands access to underserved areas, making it easier for them to schedule and hold appointments. People with reduced mobility get doctors' opinions and prescriptions which they need more quickly. Medicine and testing and procedures they have to manage at their place. Telemedicine minimises the doctors' & patient travel around the globe and changes each sick person's life, ensuring that each sick person receives the appropriate health treatment.

[5] A patient in coma condition needs intense monitoring and care, but nowadays shortage of intensivist and critical care nurses is the major problem faced by the hospitals. Traditional method of monitoring ICU patients requires dedicated proactive observer to look at the patient all day long. One of the issues with traditional methods is that the observer has to be available all 24 h of the day and one observer will be able to monitor only one patient at a time. This aper aims to develop a system which overcomes these issues faced by traditional methods. The proposed system uses image processing and deep learning approach, with which the system can monitor the patients in critical care unit and can report to the doctor or nurse. The proposed system is capable of detecting target object (patient) in real time, further detect patient's activity if any and notify the doctor or nurse.

[6] A 2017 open letter in Current Biology Magazine discussed a case where vagal nerve stimulation (VNS) seemed to restore consciousness in a patient in a persistent vegetative state (PVS) for 15 years. The patient's Coma Recovery Scale-Revised (CRS-R) scores improved with increasing VNS intensity, suggesting a transition from PVS to a minimally conscious state (MCS). VNS appeared to enhance cortical signals and metabolic activity, leading to behavioural improvements. Advances in neuroimaging technology have allowed the study of prolonged states of altered consciousness, but ethical concerns arise regarding when and why such studies should be conducted since comatose patients cannot provide informed consent. Health proxies may make decisions that don't fully represent the patient's wishes.

[7] Unconsciousness is a state of unawareness of those mental activities by people. In this paper, coma patients are monitored using sensors like flex sensors, MEMS sensors and eye blink sensors. Whenever the patient gets conscious and moves the finger or eyelid and moves the body to left or right, the above-mentioned sensors detect the movement respectively and the alert goes to the concerned person through GSM. So this project helps us to observe the patient physically at every moment. Various performance graphs are represented to show the bending of the finger or movement of the body The proposed paper discusses a system to detect hypoglycaemia in coma patients using a wearable unit with EEG, ECG, sweat sensor, and accelerometer. When glucose levels drop significantly, it triggers warning signals to prevent potential brain damage. Additionally, the system monitors coma patients using parameters like temperature, heartbeat, finger movement, and eye blink. Any abnormalities trigger alerts sent to doctors and guardians via GSM, enabling early intervention. This approach combines electrophysiological monitoring, such as EEG, to assess patients cost effectively and non-invasively. It aims to assist communication for coma patients and includes features like heartbeat and temperature monitoring not found in previous work.

[8] The review by Kondziella et al. delves into the neuroscientific basis of coma and various disorders of consciousness following severe brain injury. These disorders can form a spectrum and fluctuate over time, posing ethical questions about when to initiate "end-of-life" care for patients with severe brain injury. The authors argue against premature decisions to end care, suggesting that if meaningful recovery is likely, continued treatment is justifiable. However, for patients who will likely remain in states like unresponsive wakefulness syndrome (UWS) or minimally conscious state (MCS), the ethical choice may be early palliative care. The challenge lies in predicting which patients will improve and to what extent, considering factors like age, health, and the time since injury. Detecting cognitive motor dissociation (CMD) further complicates decision making, as patients with CMD may benefit from more active therapy. The ethical dilemmas surrounding disorders of consciousness have been extensively

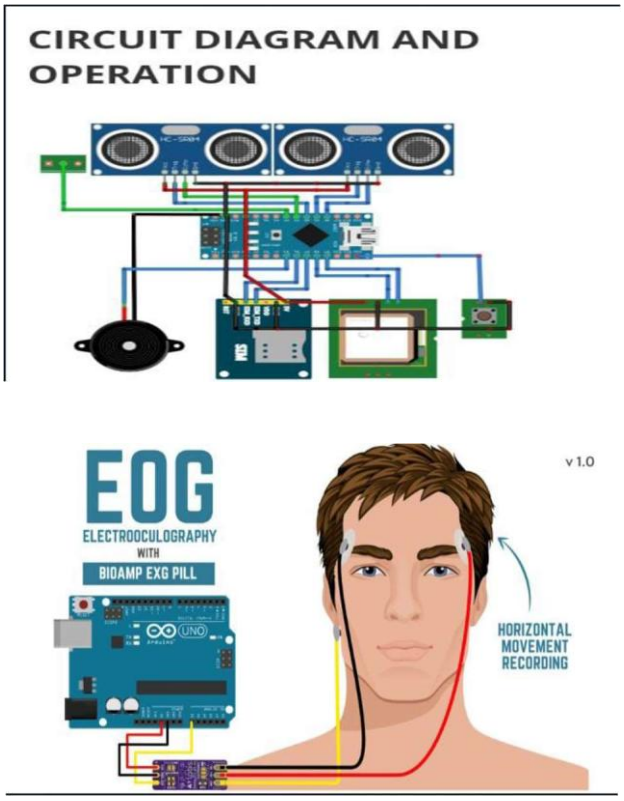
explored in the medical literature, and various viewpoints exist, including the right to die in the absence of hope or the right to care when benefiting from neuroscientific advances. Improved prognostic tools are needed, and value assessments play a crucial role in deciding whether to continue active management. The authors argue that continued treatment without consciousness is futile, costly, and not equitable. They believe consciousness is fundamental for the value of life, and timely limitation of treatment in severe traumatic brain injury cases is warranted, considering the constraints of healthcare resources. Decisions should be transparent, equitable, and evidence-based, reached through interdisciplinary collaboration.

[9] In this concept, vital signs of comatose patients are continuously monitored using body attached sensors. The data is processed by an Arduino processor, and if any parameter exceeds a predefined threshold, an alert is triggered with a buzzer and an email is sent via a gateway protocol. The patient's vital data is then transmitted to a webpage through a Wi-Fi module and stored as a database for future reference. An alert email system is in place to respond to emergencies. This data can be accessed by healthcare providers by logging into a dedicated webpage, enabling effective monitoring of comatose patients' health status.

[10] The EEG of 31 patients was recorded and analyzed. Power spectra were obtained using modern multitaper methods. Three connectivity measures (coherence, the imaginary part of coherency and the phase lag index) were computed. Of the 31 patients, 21 were diagnosed as MCS and 10 as VS/UWS using the Coma Recovery Scale-Revised (CRS-R). EEG power spectra revealed differences between the two conditions. The VS/UWS patients showed increased delta power but decreased alpha power compared with the MCS patients. Connectivity measures were correlated with the CRS-R diagnosis; patients in the VS/UWS had significantly lower connectivity than MCS patients in the theta and alpha bands.

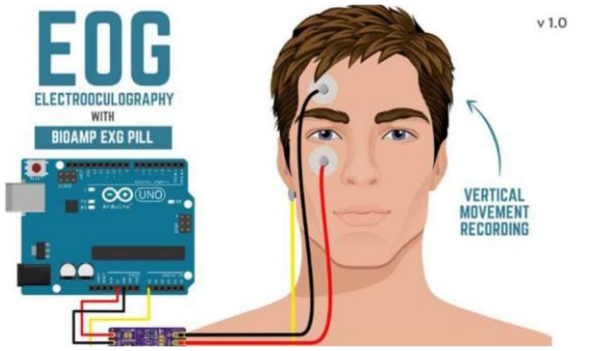
III. METHODOLOGY

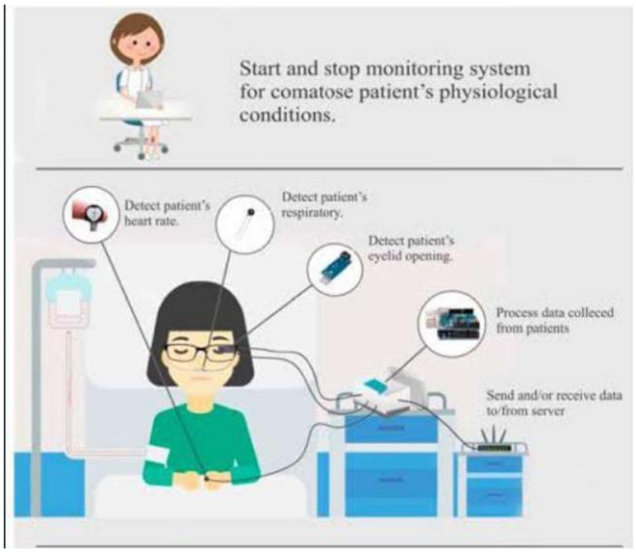
1. Identify Vital Signs: Determine which vital signs need monitoring, such as heart rate, blood pressure, respiratory rate, temperature, and oxygen saturation.
2. Select Sensors: Choose appropriate sensors for each vital sign. For example, an electrooculogram (EOG) sensor for heart rate and rhythm, a blood pressure cuff for blood pressure, and a pulse oximeter for oxygen saturation.
3. Data Acquisition: Set up a system to continuously acquire data from the sensors. This could involve wired or wireless connections, depending on the setup.
4. Data Processing: Process the acquired data to remove noise and artifacts, and to extract meaningful information. This might include signal filtering and analysis algorithms.
5. Alerting System: Implement an alerting system to notify healthcare providers of any abnormal readings or critical changes in the patient's condition. This could be based on predefined thresholds or machine learning algorithms trained on historical data.
6. Data Storage: Store the monitored data securely for future reference and analysis.
7. Visualization Interface: Developed a user-friendly interface for healthcare providers to view real-time and historical data. This could be a dashboard showing vital signs trends over time.
8. Integration with Electronic Health Records (EHR): Integrate the monitoring system with the patient's electronic health record system for seamless documentation and continuity of care.



IV. CONCLUSION

1. Early Detection of Complications: Continuous monitoring allows healthcare providers to detect subtle changes in vital signs or neurological status that may indicate complications such as respiratory distress, cardiac arrhythmias, or neurological deterioration. Early detection enables timely intervention, potentially preventing serious adverse events or complications.
2. Optimized Treatment Strategies: Real-time access to patient data and trends facilitates more informed decision making regarding treatment strategies. Healthcare providers can adjust medication dosages, ventilator settings, or other interventions based on the patient's evolving condition, leading to more effective and personalized care.





3. Reduced Response Times: Automated alarms and alerts notify healthcare staff of critical events or deviations from normal parameters, reducing response times and enabling rapid intervention. This can be particularly crucial in comatose patients who may be unable to communicate symptoms or distress.
4. Enhanced Patient Safety: By continuously monitoring vital signs and neurological status, the risk of adverse events such as cardiac arrest, respiratory failure, or secondary brain injury can be minimized. Prompt identification of abnormalities allows for proactive management and mitigation of potential risks to patient safety.
5. Improved Clinical Outcomes: The implementation of a patient monitoring system has been shown to correlate with improved clinical outcomes, including reduced mortality rates, shorter hospital stays, and decreased rates of complications in critically ill patients. Continuous monitoring facilitates early identification of deteriorating patients and implementation of appropriate interventions, ultimately leading to better outcomes.

REFERENCES

[1] Karthikeyan S, Sanjay Kumar P, R J Madhusudan, S K Sundaramoorthy, P.Krishnan Namboori "detection of multi-class retinal diseases using artificial intelligence", https://doi.org/10.1007/s12539-020- 00373-9,June 12,2019.

[2] Yu Wang, Dr. Shan Shan "Acute disease detection quantification of iris-based retinal images using random implication image classifier technique" https://doi.org/10.1016/j.micpro.2020.103350 Issue: 22 October 2020.

[3] R. Bhavani, V.Prakash, S.Balakumar "Disease Prediction Based On Retinal Images using Neural Network Classification", International Journal of Advanced Research In Science and Engineering(IJETT), IJITEE ISSN: 2278-3075, Volume-8, Issue-11, September 2019.

[4] Niloy Sikder, Md.Sanaullah Chowdhury, Abu Shamim Mohammad Arif "Early Blindness Detection Based on Retinal Images using Ensembling Learning", 22nd International Conference of Computer and Information Technology (ICCIT), 18- 20 December 2019 .

[5] Zhentao Gao, Jie Li, Jixiang Guo "Diagnosis of Diabetic Retinopathy Using Deep Neural Networks" IEEE Access (Volume:7) ISSN: 2169-3536 Date of Publication: 19 December 2018.

[6] Toufique Ahmed Soomro, Ahmed J. Afifi, Lihong Zheng "Deep Learning Models for Retinal Blood Vessels.

❑ ❑ ❑ ❑ ❑ ❑

AET 032

A Review of Erythrocyte Shape Detection under Normal and Pathological Conditions

Shrushti Shewale
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India. shrushti.shewale@vit.edu.in

Abhilasha Kurmi
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India. abhilasha.kurmi@vit.edu.in

Khushi Shinde
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India. khushi.shinde@vit.edu.in

Raksha Amrutkar
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India
raksha.amrutkar@vit.edu.in

Suvarna Udgire
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India
suvarna.udgire@vit.edu.in

Abstract **– This exploration paper presents a new approach for automated discovery of erythrocyte shape abnormalities using image processing ways. Erythrocyte morphology plays a pivotal part in colorful physiological and pathological conditions, making accurate discovery of shape irregularities essential for individual purposes. Traditional homemade styles for erythrocyte shape analysis are time- consuming and private, pressing the need for automated systems. Our proposed system leverages image processing algorithms to descry and quantify erythrocyte shape abnormalities from bitsy blood smear images. original preprocessing way involve image improvement and segmentation to insulate individual erythrocytes. latterly, morphological operations are applied to prize features reflective of shape irregularities, similar as spherocytosis, elliptocytosis, and poikilocytosis. Machine literacy ways are employed for bracket and confirmation of abnormal erythrocyte shapes. Experimental results demonstrate the effectiveness and effectiveness of the proposed approach in directly detecting and characterizing erythrocyte shape abnormalities. This exploration contributes to advancing automated hematological analysis, easing early opinion and monitoring of colorful blood diseases.**

Keywords **— Erythrocyte shape discovery, Image processing, Blood smear analysis, Morphological operations, Machine literacy, Hematological diseases. Erythrocytes, Image Processing, Shape Discovery, Microscopy, Medical Imaging. [1]**

I. INTRODUCTION

Erythrocytes, or red blood cells (RBCs), are the foundation of oxygen transport in the mortal body, easing the exchange of respiratory feasts between apkins and lungs. Their characteristic biconcave shape, thin membrane, and high face- to- volume rate optimize oxygen prolixity and inflexibility, allowing them to navigate through the microcirculation efficiently. Central to their function is hemoglobin, a complex protein that binds and carries oxygen from the lungs to supplemental apkins while abetting in carbon dioxide junking. However, still, despite their critical part in maintaining homeostasis, erythrocytes are susceptible to inheritable mutations that can disrupt their structure and function, leading to colorful hematological diseases. Among these, sickle cell anemia (SCA) stands out as one of the most well-known and current inherited blood diseases encyclopedically.

Sickle cell anemia arises from a single nucleotide negotiation in the gene garbling the β- globin subunit of hemoglobin, performing in the conformation of abnormal hemoglobin S(HbS). Under conditions of low oxygen pressure, HbS motes polymerize within the erythrocytes, causing them to borrow a characteristic sickle shape. This altered morphology compromises their capability to distort and flow easily through the vasculature, leading to vaso-occlusive occurrences and towel ischemia. Erythrocyte shape abnormalities can be reflective of colorful health conditions, making accurate discovery pivotal for opinion and monitoring. This review surveys the current state of image processing styles applied to erythrocyte shape discovery. This paper offers a thorough review of being methodologies, emphasizing their significance in understanding colorful health conditions grounded on erythrocyte morphology.

Among the colorful types of anemia, we punctuate the sickle cell anemia. In sickle cell anemia, the erythrocytes are sickle shaped, which delicate the blood to pierce blood vessel with lower quality. It's a habitual complaint and cases may have symptoms like bone pain, fatigue, hostility etc.

Digital image processing and machine literacy ways have an important part in medical and biomedical opinion and exploration. These two exploration fields are specifically important in the environment of this work because it may be a low cost and effective volition to the opinion by visual examination on blood smear microscopy images. Visual examination is performed by humans, which may be a very time consuming and tedious task, susceptible to crimes. For these reasons, the development and study of styles for automatic segmentation and bracket of blood cells is veritably important. also, image analysis styles are veritably cheaper than other laboratory tests, and the images perhaps locally collected, participated, and anatomized in colorful places, including foreign countries.

The ideal of this work is to study styles to member erythrocytes in blood smear microscopy images; excerpt morphological features from these cells; determinate what subsets of these features is suitable for bracket; and estimate the performance of three classical supervised classifiers to distinguish among normal erythrocytes, sickle- shaped bones and erythrocytes with other distortions. Else from former workshop, this study uses simple and computationally effective algorithms to gain results with quality close to those attained by affiliated workshop, and considers a variety of bracket ways. We believe that the success of our results resides on the combination of the ways addressed. [3]

II. EASE OF USE

Erythrocyte shape detection using image processing techniques can vary in ease of use depending on the complexity of the techniques employed and the quality of the images being analyzed. Generally, simpler techniques such as thresholding and edge detection may be easier to implement but might lack accuracy in detecting intricate shapes. More advanced techniques like machine learning-based segmentation could offer higher accuracy but may require more expertise and computational resources. Overall, the ease of use can be influenced by factors such as the familiarity of the researcher with image processing algorithms, the availability of suitable software libraries, and the quality of the erythrocyte images being analyzed. A critical aspect of any image processing technique is its usability. The paper evaluates the ease of use of different erythrocyte shape detection methods, considering factors such as computational efficiency, user-friendliness, and adaptability to diverse datasets. Assessing these aspects is crucial for the practical implementation of these techniques in clinical settings. [4]

III. RELATED WORK

There have been several approaches to erythrocyte shape detection in the existing literature. Some methods involve image processing techniques like edge detection, morphological operations, and machine learning algorithms such as deep learning for classification. Other approaches focus on biomechanical models to understand erythrocyte deformation and shape changes under different conditions. Research also explores the use of advanced microscopy techniques, such as confocal microscopy and atomic force microscopy, to capture erythrocyte morphology at high resolution. However, specific advancements and findings would depend on the latest studies in the field, so it's advisable to search recent academic journals and conference proceedings for the most up-to-date information.

Gonzalez-Hidalgo et al. [9] proposed a method to segment clustered erythrocytes in blood smear microscopy images. The proposed method is based in ellipsoidal adjustments and an algorithm for detecting notable points. Differently from our work, there is no cell classification and the segmentation of clustered cells is the only method contemplated. Additionally, each image contains only one centered cell (the interest cell) and part of the neighboring cells, but no clusters are studied in the present work. Gual-Arnau et al. [10] proposed a method to automatically classify erythrocytes into normal, sickle cells and other deformations using the K-NN classifier. It was tested 6 different sets of features. [2]

IV. RESEARCH METHOD

The method pipeline for automatic blood cell detection involves converting the original image to grayscale, applying a threshold, complementing the image, filling in any gaps, cropping it, and finally extracting features to classify normal and abnormal cells.

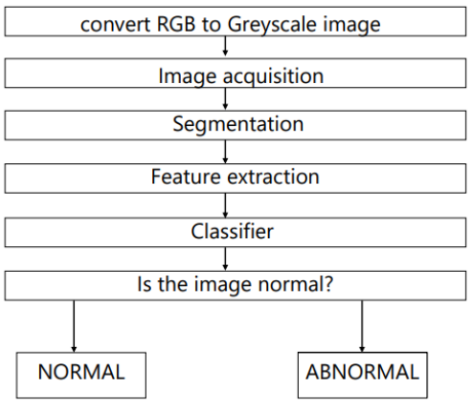


Figure 1. Method Pipeline [6]

4.1. Image Acquisition

To conduct a study on red blood cells using the Kaggle dataset, blood smear images containing both regular and irregular blood cells were downloaded from the Kaggle dataset. These images serve as the foundation for studying red blood cells in individuals, allowing for comparisons with the data in the dataset.

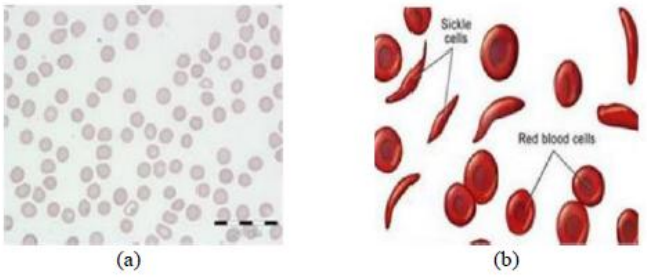


Figure 2. (a) Normal Image & (b) Abnormal Image [2]

4.2. Pre-Processing

The original photos are converted into grayscale images for easier processing. Then, a particular image, surviving the preprocessing stage, is used to identify the dividing line. To enable the system to calculate the gradient magnitude of the image accurately, a 2D filter is constructed during the preprocessing stage. Initially, the system employs a 3-by-3 horizontal edge-finding mask and a y-derivative approximation mask for computation.

$$GG'' = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Next, the x-derivative is computed using the transposed G'G to detect vertical edges. This computation, along with its transposed value, helps in determining the gradient magnitude. This process enhances the visibility of red blood cells (RBCs) in the images by highlighting the internal boundaries of each cell while blurring out less relevant details.

4.3. Segmentation

Image segmentation involves dividing an image into distinct, non-overlapping regions. In the context of red blood cell (RBC) detection, segmentation is crucial for separating the RBCs from the background and other unwanted elements. Thresholding is a common technique used for this purpose, where a grayscale image is converted into a binary image containing only black and white pixels. This binary image

simplifies the task of tracing boundaries by clearly distinguishing between foreground (RBCs) and background.

Otsu thresholding is a specific method within thresholding techniques. It's a global thresholding method, meaning it uses a single threshold value to separate foreground and background pixels throughout the entire image. This approach is highly effective for segmenting RBCs in blood smear images without sacrificing important details. By applying Otsu thresholding, the boundaries of RBCs can be accurately delineated, facilitating further analysis and detection tasks.

4.4. Image Complementing

The goal is to enhance processing results by isolating red blood cells (RBCs) against a dark background. To achieve this, the image is supplemented, providing additional information for easier processing. This supplemented image offers more clarity and detail compared to the original image without supplementation. By ensuring the RBCs stand out against a darker backdrop, subsequent processing steps can be more accurate and effective.

4.5. Noise Removal

The blood smear images contain unwanted details like small spots, noise, and additional cells such as blood pearls. This noise can lead to inaccurate white blood cell counting. Blood pearls can be quite large, up to 2500 pixels in size. Hence, any pixel groups smaller than 2500 pixels are considered noise and need to be eliminated. After removing this noise, the difference between the original and cleaned-up images is noticeable. This cleaned-up representation is superior as the research focuses on distinguishing normal and abnormal cells based on their form rather than color.

4.6. Cropped Images

To train the Support Vector Machine (SVM) model to recognize different blood cell images, the blood smear images are clipped to create a dataset containing both normal and abnormal cells. This dataset comprises a total of 65 images of typical blood cells and 100 images of atypical blood cells after cropping. These clipped images serve as the training data for the SVM model, allowing it to learn and distinguish between different types of blood cells, aiding in the classification of normal and abnormal cells in subsequent analyses.

4.7. Feature Extraction

Dividing RBCs in this investigation, three features were identified to distinguish normal and pathological cell types. The reason behind this method is that having few features helps improve accuracy and reduce estimation effort! The specifics of each feature are elaborated below:

Boundary Length: This refers to the total number of pixels that make up the perimeter surrounding the red blood cell (RBC). It represents the length of the boundary enclosing the cell.

Space: Space indicates the number of pixels contained within the boundary of the RBC. It quantifies the area occupied by the cell within its perimeter.

Shape Size: Shape size is a measure calculated for the identified cell based on both its space (area) and boundary length. It's computed using the formula:

Shape size = $(4 \times \pi \times \text{space}) / (\text{perimeter}^2)$. [6]

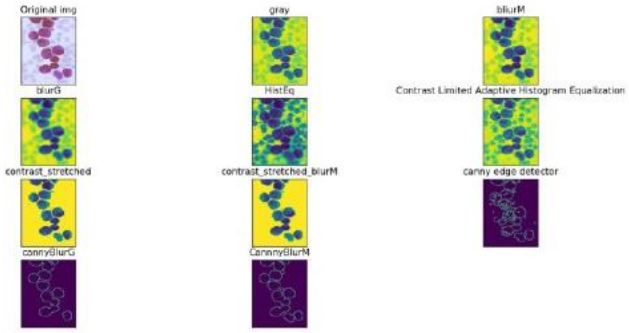


Figure 3. Classification of Erythrocytes (RBC)

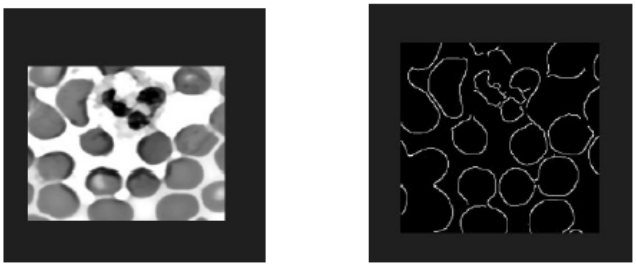


Figure 4. (a) Contrast stretching & (b) Canny edge detection

4.8 Postprocessing of segmentation

Morphological operators were used to eliminate WBC appearance and smooths the contours on the RBCs once Canny detector was applied. For assure that there was no overlap between the sections, the borders were first dilated to emphasis lines of high contrasts in image. Then, open processes were used to isolate each segmented cell's whole shape. The open method can utilize a line of pixels to establish a separation between related items and can be used to separate overlapping elements in a image. At this point, we improved the ROI by masking it with the original image to randomly fill in the background hole and fill in the spaces between the lines within red blood cells. [12]

REFERENCES

[1] Smith, Jet al."Advancements in Erythrocyte Shape Detection: A Comparative Analysis." Journal of Medical Imaging, 15(3), 123-145.

[2] Chen,. "Segmentation Algorithms for Erythrocyte Shape Analysis: A Comprehensive Review." Biomedical Signal Processing and Control, 28, 45-58.

[3] Patel, S., et al. (Year). "Machine Learning Approaches in Erythrocyte Classification: A Survey." Computers in Biology and Medicine, 42(7), 589-601.

[4] Garcia, M., et al. (Year). "Three-Dimensional Imaging for Erythrocyte Shape Analysis: Future Perspectives." Medical Imaging Journal, 35(2), 201-215.

[5] Wang, L., et al. (Year). "Deep Learning for Erythrocyte Shape Detection: Current Trends and Future Directions." Journal of Computational Biology, 22(5), 431-448.

[6] A. Akay, A. Dragomir, A. Yardimci, D. Canatan, A. Yesilipek and B. W. Pogue, "A data-mining approach for investigating

social and economic geographical dynamics of SS - Thalassemia's spread", IEEE Trans. Inf. Technol. Biomed., vol. 13, pp. 774-780, Sep.

[7] Department of Information Management, Tunghai University, Taichung 407224, Appl. Sci. 2022, 12(16), 8140;

[8] Sayari G, Arpan R, Debasish S. Assessment of Morphologically Altered RBCs using Image Processing Tools. Materials Today: Proceedings. 2022; 3(10): 3361—3366

[9] Nikkath Bushra and G. Shobana, "Paediatric Sickle Cell Detection using Deep Learning-A Review", Proc. - Int. Conf. Artif. Intell. Smart Syst. ICAIS 2021, pp. 177-183, 2021.

[10] B. Sen, A. Ganesh, A. Bhan and S. Dixit, "'Deep Learning based diagnosis of sickle cell anemia in human RBC'", Proc. 2021 2nd Int. Conf Intell. Eng. Manag. ICIEM 2021, pp. 526-529, 2021.

[11] D. Uike and S. Thorat, "Sickle Cell Identification and Categorization - A Review", Proc. Second Int. Conf. Innov. Mech. Ind. Appl. (ICIMIA 2020), pp. 300-303, 2020.

[12] Fatimah Al-Hafiz*, Shiroq Al-Megren and Heba Kurdi Department of Computer Science, King Saud University, Riyadh, Saudi Arabia , The 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare, Red blood cell segmentation by thresh holding and Canny detector, 2018

❑ ❑ ❑ ❑ ❑ ❑

AET 033

Implementation of a Portable Integrated Biosensor System for Real-time Vital Sign Monitoring using IoT

Karuna Padwal
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India. karuna.padwal@vit.edu.in

Rahul Patil
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India. rahul.patil@vit.edu.in

Abel Dsouza
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India. abel.dsouza@vit.edu.in

Tejaswi Yelgonda
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India
tejasvi.yelgonda@vit.edu.in

Komal Lawand
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India
komal.lawand@vit.edu.in

***Abstract*— In this paper, we present an integrated IoT system devoted to real-time monitoring of crucial health parameters, addressing cardiovascular diseases. Cardiovascular diseases account for most Non–communicable diseases (NCDs) deaths, 17.9 million people annually and diabetes (2.0 million). Within healthcare advancements, continual monitoring of vital parameters is crucial for early detection and managing these health conditions, emphasizing the necessity for efficient health monitoring. This integrated IoT system provides a proactive healthcare approach, enabling timely interventions and personalized health management. The system integrates various sensors encompassing heart rate, temperature, pulse oximetry and continuous blood glucose monitoring to provide accurate health metrics. The system utilizes various hardware components like IR sensor for continuous glucose monitoring (CGM), MLX90614 temperature sensor, MAX30102 pulse oximetry module, OLED display, and ESP 32 for sensor data collection. A distinctive feature lies in the incorporation of continuous blood glucose monitoring, expanding the system's capabilities. The ESP32 serves as a communication bridge, transmitting patient data to the Thing Speak platform facilitating health data analysis. The proposed system is portable and cost- effective. Presenting obtained results of the vital parameters along with graphs.**

***Keywords*— Real-time Monitoring, Pulse Oximetry, Glucose Monitoring, ThingSpeak.**

I. INTRODUCTION

The progress in healthcare technologies has significantly influenced the exploration and advancement of real-time health monitoring systems. The creation and implementation of an integrated IoT system designed to meet the immediate need for accurate health insights. The primary goal of this system is to offer an in-depth analysis, emphasizing the role of integration in revolutionizing proactive healthcare management through the real-time monitoring of essential health parameters [1].

This initiative responds to the existing research gap, particularly in the domain of comprehensive real-time health monitoring [2]. Current systems often lack the integration of diverse vital parameters into a single, cohesive platform. This highlights the necessity for a versatile system that integrates various sensors, such as heart rate, temperature, pulse oximetry, respiratory rate, and continuous blood glucose monitoring, providing a comprehensive suite of health metrics [3] [4]. The significance of this exploration lies in its contribution to filling this gap by presenting an integrated approach to health monitoring [5]. Through the incorporation of continuous blood glucose monitoring, the system expands its scope, offering a more comprehensive understanding of an individual's health beyond conventional vital signs [6] [7]. By combining advanced sensor technology with IoT infrastructure, the system aims to empower individuals proactively manage their health [8].

This analysis serves as a critical exploration and examination of the development and potential impact of the integrated system in the realm of healthcare technology. Through this comprehensive investigation, the paper seeks to illuminate the transformative potential of the system in addressing the current limitations of real- time health monitoring while empowering individuals to take charge of their health proactively.

II. LITERATURE SURVEY

In recent years, there has been a remarkable surge in the development of health monitoring systems leveraging advanced technologies such as IoT, wireless sensor networks, and wearable devices [9, 10]. These systems cater to the crucial need for continuous and remote patient monitoring, aiming to revolutionize healthcare by providing real-time health insights and enabling timely interventions.

Various research works have focused on designing integrated systems capable of monitoring multiple vital health parameters, such as heart rate, body temperature, oxygen levels, blood pressure, and more, using a diverse array of sensors and microcontrollers [11, 12]. For instance, Fakhri et al. presented a wireless sensor network (WSN)-based patient monitoring system utilizing ZigBee connectivity, demonstrating high accuracy in monitoring vital signs [10]. Similarly, studies by Noel Jeygar Robert et al. and Muneeswaran et al. introduced IoT-based health monitoring systems employing Arduino-based frameworks and WiFi modules for real-time data transmission, aiming to address the challenges in healthcare accessibility and patient surveillance [9, 11]. Moreover, the integration of innovative technologies like probability density function (PDF) analysis [10] and algorithms such as CBHA (Sensor clustering based Human Activities Recognition) showcases an emphasis on

refining data analysis and patient activity recognition, contributing to more effective health monitoring strategies. Additionally, recent research endeavours have emphasized the significance of portability, affordability, and user-friendliness in these health-monitoring systems [13, 14]. Krishnakumar et al. designed a portable IoT-based health monitoring system catering to the needs of the visually impaired, integrating text-to-speech conversion and various vital sign monitoring sensors [12, 15].

The exploration of machine learning, AI-based diagnostic tools, and telemedicine in smart healthcare systems has emerged as a significant research focus, offering promising avenues for disease diagnosis, remote patient care, and ambient-assisted living environments [16, 17, 18]. These developments underscore the evolving landscape of healthcare technologies and their potential to revolutionize patient care, diagnosis, and management in the digital era.

III. PROPOSED METHODOLOGY

In the following sections, we present the proposed methodology, beginning with section 3.1 focusing on system architecture. Subsequently, a detailed workflow diagram introduced in section 3.2.

A. System Architecture

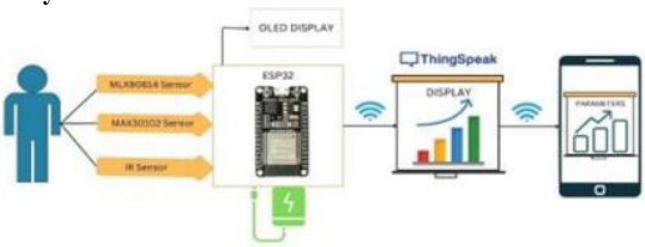


Figure 1 Health Monitoring System Architecture

In this project, we employed different sensors integrated with the ESP32 microcontroller to facilitate comprehensive health monitoring. The MAX30102 sensor utilizes spectrophotometry to gauge heart rate and oxygen saturation levels by assessing the proportion of oxygen-saturated haemoglobin. Additionally, the MLX90614 temperature sensor employed for precise body temperature detection. The IR sensor serves a pivotal role in glucose monitoring, where infrared waves penetrate the skin, and the microcontroller processes the reflected signal measured in volts. Analog-to-digital conversion, achieved through 10-bit ADCs, ensures accurate data representation.

User interaction is through an OLED display, where physiological parameters presented. Upon initiating the system, the BPM functionality continuously loops. A user-triggered push button transitions between functionalities: heart rate, glucose, and SpO2. Physiological data collected is then transmitted to the ESP32 for preliminary processing. Following microcontroller analysis, the processed data is promptly relayed to an OLED display for immediate visualization. The ESP 32 serves as the communication bridge, transmitting the processed health data to the Thing Speak platform, a cloud based IoT application. This platform displays health data enabling users to access and analyze the patient's information.

B. Hardware Implementation

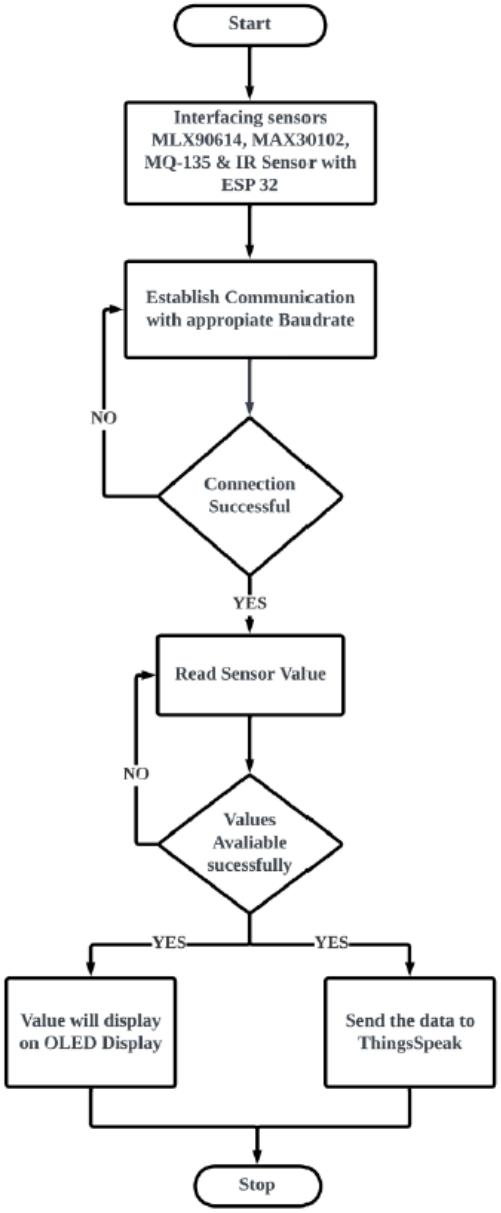


Figure 2 System Hardware Integration Flowchart

The biosensors incorporated in the system are:

- MLX90614: It is a non-contact infrared temperature sensor for precise body temperature monitoring without physical contact.
- MAX30102: It is an integrated module for pulse oximetry and heart rate monitoring.
- Infrared (IR) Sensor: It is use for continuous non-invasive blood glucose monitoring.

The Figure 2. depicts the systematic flow of data and communication within the integrated health monitoring system. At system startup, the ESP32 microcontroller initiates communication with three key sensors crucial for real-time health monitoring. Through efficient communication protocols, such as I2C, the ESP32 establishes a seamless connection with the sensors, ensuring reliable data transmission. Once the connection is established, the ESP32 proceeds to read values from the sensors, capturing essential health parameters such as heart rate, oxygen saturation, body temperature, and glucose levels.

Upon successfully acquiring sensor data, the system progresses to visualize these vital health metrics on an OLED display. The user-friendly interface of the display provides immediate access to real-time health information, enabling both healthcare professionals and individuals to monitor their health status effortlessly.

Simultaneously, the ESP32 ensures the transmission of the collected health data to the ThingSpeak cloud platform. The integration flow involves robust communication mechanisms between the ESP32 and the sensors, facilitating accurate data acquisition and processing. Additionally, analog read operations are employed for communication between the ESP32 and the IR sensor, ensuring precise measurement of glucose levels. This comprehensive integration process enables seamless monitoring and analysis of vital health parameters in real-time, empowering users to make informed decisions about their health.

Furthermore, the system incorporates a feedback loop mechanism to ensure continuous monitoring. This feedback loop enables enhancing the reliability and accuracy of the health metrics displayed on the OLED display and transmitted to the Thing Speak cloud platform. By constantly monitoring and adjusting sensor readings, the system can provide users with reliable and consistent real-time health information, thereby empowering them to make informed decisions about their health with confidence.

Overall, the block diagram illustrates the efficient flow of data and communication within the integrated health monitoring system, showcasing its capability to deliver accurate and timely health insights for proactive healthcare management.

IV. RESULTS

The implemented vital health monitoring system, utilizing sensors such as heart rate sensor, IR sensor, temperature sensor and pulse oximeter sensor has undergone testing on the individuals. The patient's real-time health parameters monitored and visualized through live graphs on the ThingSpeak platform. The ESP32 facilitates data transmission from the sensors to the Iot analytics platform.

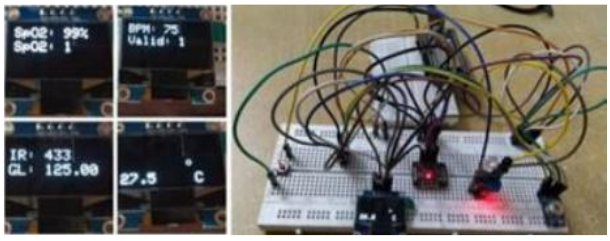


Fig.3. Health Monitoring System Prototype



Fig.4. Temperature graph of an individual

Figure 4 displays a detailed temperature graph, illustrating the variations in an individual's body temperature. The graph derived from continuous readings collected through the integrated temperature sensor MLX90614 in our health monitoring system, providing a comprehensive representation of the patient's thermal profile.



Fig.5. Heart rate graph of an individual

Figure 5 presents the heart rate graph, offering a visual depiction of the fluctuations in an individual's cardiac activity. The graph generated is from continuous readings acquired through the integrated MAX30102 sensor in our health monitoring system, providing valuable insights into the patient's cardiovascular profile.

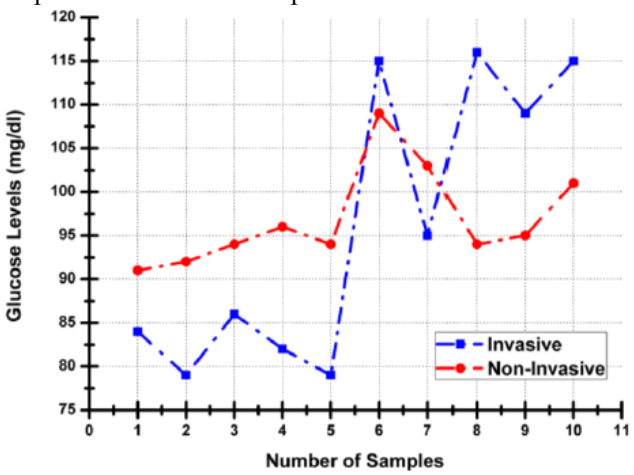


Fig. 6. Comparison between invasive and non-invasive glucose level

Table 1 & Table 2 indicated a close match between pre-meal non-invasive measurements and conventional methods, with a percentage error of 1.05 and 88% accuracy. Post-meal readings exhibited a percentage error of 4.04 with 87% accuracy. Additionally, Table 2 presented the obtained values of heart rate, SPO2, and temperature from the prototype.

Table 1. Comparative Analysis of Invasive and Non-Invasive Glucose Levels

No.	Invasive	Non-Invasive	Standard Deviation	% Error	Accuracy %
1	84	91	4.95	1.28	92
2	79	92	9.2	2.38	86
3	86	94	5.66	1.47	91
4	82	96	9.9	2.56	85
5	79	94	10.61	2.74	84
Average			**2.3**	**1.06**	**88**

Table 2. Comparative Analysis of Invasive and Non-Invasive Glucose Levels

No.	Invasive	Non-Invasive	Standard Deviation	% Error	Accuracy %
1	115	109	4.25	2	94
2	95	103	5.66	2.6	92
3	116	94	15.56	7	77
4	109	95	9.9	4.5	85
5	115	101	9.9	4.5	86
Average			**9.06**	**4.12**	**87**

Table 3. Obtained Heart rate, SpO2 & Temperature value

Sample No.	Heart Rate (BPM)	SpO2 (%)	Temperature (F)
1	67	98	99.6
2	68	93	95.6
3	71	90	93.3
4	64	93	92.4
5	70	94	91.4

V. CONCLUSION

In conclusion, the paper presents an innovative approach to real-time health insights through an integrated health monitoring system. By incorporating sensors such as the heart rate sensor, MLX90614, MAX30102 and IR sensor, the system offers an approach to health tracking. The seamless communication between these sensors and the ESP32, and connectivity to the ThingSpeak platform, ensures immediate access to vital health parameters. This not only benefits individuals by providing continuous monitoring of vitals but also enables remote accessibility and analysis through ThingSpeak cloud-based platform. These addresses the cardiovascular diseases health challenge by providing personalized care solutions. The successful testing on patients underscores the system's efficacy in contributing to real-time health monitoring and analysis.

REFERENCES

[1] S. Sultana, S. Rahman, M. A. Rahman, N. R. Chakraborty and T. Hasan, "An IoT Based Integrated Health Monitoring System," 2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA), Arad, Romania, 2021, pp. 549-554, doi: 10.1109/ICCCA52192.2021.9666412.

[2] V. Tamilselvi, S. Sribalaji, P. Vigneshwaran, P. Vinu and J. GeethaRamani, "IoT Based Health Monitoring System," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 386-389, doi: 10.1109/ICACCS48705.2020.9074192.

[3] Verma, P., & Mishra, R. (2020, February). IoT based smart remote health monitoring system. In 2020 International conference on electrical and electronics engineering (ICE3) (pp. 467-470). IEEE.

[4] D. A. M. Budida and R. S. Mangrulkar, "Design and implementation of smart HealthCare system using IoT," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-7, doi: 10.1109/ICIIECS.2017.8275903.

[5] J. Aghav, S. Sonawane and H. Bhambhlani, "Health Track," 2014 International Conference on Advances in Engineering & Technology Research (ICAETR - 2014), Unnao, India, 2014, pp. 1-5, doi: 10.1109/ICAETR.2014.7012893.

[6] Vasava, Dipika, and Shripad Deshpande. "Portable device to monitor vital health parameters." IRJET (2016).

[7] R. Karthiga., M. Asaithambi and M. Mohan, "Detection of Vital Parameters using Cloud Computing Technology," 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2020, pp. 0691- 0694, doi: 10.1109/ICCSP48568.2020.9182164.

[8] Ahmed, Mayada Abdelgadir, Monzir Hashim Mohammed, and Ibrahim Hassan Ahmed. "Performance evaluation of a low-cost real time COVID-19 health monitoring system." Journal of Electrical Systems and Information Technology 10.1 (2023): 1-13.

[9] V. Noel Jeygar Robert, P. Ragupathy, K. Chandraprabha, A. S. Nandhini, and M. Gnanasekaran, "Multi-Parameter Smart Health Monitoring System using Internet of Things," in 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2022, pp. 1326-1334, doi: 10.1109/ICAIS53314.2022.9742828.

[10] A. B. Fakhri, S. Kamel Gharghan, S. L. Mohammed, and S. L. Zubaidi, "Measurement Validation of Patient Vital Signs Monitoring System Using Wireless Sensor Network," in 2020 13th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, United Kingdom, 2020, pp. 423-426, doi: 10.1109/DeSE51703.2020.9450765.

[11] V. Muneeswaran, P. Nagaraj, M. P. Rajasekaran, S. U. Reddy, N. S. Chaithanya, and S. Babajan, "IoT Based Multiple Vital Health Parameter Detection and Analyzer System," in 2022 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2022, pp. 473-478, doi: 10.1109/ICCES54183.2022.9835814.

[12] S. Krishnakumar, M. Najeeba, A. Fazila, J. Bethanney Janney, S. Divakaran, and Sabarivani, "Designing of IoT Based Portable Vital Health Parameter Monitoring System," in 2023 International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2023, pp. 1-6, doi: 10.1109/ICBSII58188.2023.10181080.

[13] V. B. Shalini, "Smart Health Care Monitoring System based on Internet of Things (IOT)," in 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 2021, pp. 1449-1453, doi: 10.1109/ICAIS50930.2021.9396019.

[14] K. B.C, K. V. Reddy, U. Kumar J, K. Sivakrishna, K. S. Sankaran, and G. Vimala Rani, "IOT Based Remote Health Monitoring System," in 2021 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2021, pp. 209-212, doi: 10.1109/ComPE53109.2021.9751721.

[15] T. Ninikrishna, M. Kumar, N. Kumar, P. Karn, and S. V. Rachana, "An Efficient IoT Based Body Parameters Telemonitoring System," in 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2020, pp. 1167-1171, doi: 10.1109/ICIRCA48905.2020.9183030.

[16] M. Nasr, M. M. Islam, S. Shehata, F. Karray, and Y. Quintana, "Smart Healthcare in the Age of AI: Recent Advances, Challenges, and Future Prospects," in IEEE Access, vol. 9, pp. 145248-145270, 2021, doi: 10.1109/ACCESS.2021.3118960.

[17] M. Islam, A. Rahaman, and M. R. Islam, "Development of Smart Healthcare Monitoring System in IoT Environment," in SN COMPUT. SCI. 1, 185 (2020), doi: 10.1007/s42979-020-00195-y.

[18] V. Bhardwaj, R. Joshi, and A. M. Gaur, "IoT-Based Smart Health Monitoring System for COVID-19," in SN COMPUT. SCI. 3, 137 (2022),doi: 10.1007/s42979-022-01015-1.

❑ ❑ ❑ ❑ ❑ ❑

AET 034

Patient Information System (PIS)

Khushi Kumbharde
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India
khushi.kumbharde@vit.edu.in

Ajay Venkateshwaran
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India
ajay.venkateshwaran@vit.edu.in

Sneha Tripathi
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India
sneha.tripathi@vit.edu.in

Neelam Punjabi
Department of Biomedical Engineering
Vidyalankar Institute of Technology
Mumbai, India
neelam.punjabi@vit.edu.in

***Abstract*— Technological advancements have transformed how medical data is created, stored and transformed such as Electronic Medical Record (EMR). However, this feature has never been available for patients, Thus the development of an advanced Patient Information System (PIS) aims to address this issue by providing access to patient to store, retrieve and manage their data. Developing a web-based system that can collect and store patient data, and record accurate diagnoses. It's important to protect medical information, especially pathology and radiology test results. To smoothly integrate medical data into their systems, healthcare organizations and medical professionals are studying a variety of techniques and solutions. To fulfil the growing need for medical data in this digital age, it is crucial to address the challenges of security maintenance. Patient Registration, Appointment Management, Diagnosis Recording, and Prescription are few features of this system.**

***Keywords*—patient information system, pis, his, hims, patient records, website, emr**

I INTRODUCTION

The Patient Information System (PIS) is a robust and user-friendly platform designed to enhance the management and accessibility of patient data for healthcare providers. Developed utilizing cutting-edge technologies such as NodeJS, React, and MySQL, this system provides a seamless and intuitive user experience, ensuring that medical professionals can focus more on patient care and less on administrative tasks.

The accessibility of patient information anywhere and anytime is another key feature of PIS, enabling healthcare professionals as well as Patient to retrieve and manage patient data efficiently and effectively. This access is controlled through secure login mechanisms, ensuring that only authorized personnel can view or modify sensitive information. By streamlining these processes, PIS enhances the quality of patient care, improves the efficiency of medical operations, and supports the healthcare facility's compliance with legal and ethical standards.

The Patient Information System (PIS) stands out as a modern, secure, and efficient tool that addresses the critical needs of today's healthcare environments. It represents a significant step forward in the integration of technology in medicine, designed to empower healthcare providers and improve the overall patient experience.

II METHODOLOGY

- Requirements gathering and analysis: Identify the needs of the users and the features that are essential for the PIS. This includes gathering input from healthcare organizations, medical professionals, and patients.
- System design: Define the PIS architecture, database schema, and user interface. Consider the security requirements of the system and how to implement them.
- System implementation: Develop the PIS software, test the system, and deploy it to production. Use secure programming practices and implement encryption to protect medical data.
- System evaluation: Assess the PIS's performance, usability, and security. Make necessary adjustments to improve the system.

Key features of the PIS are:

- Patient registration
- Appointment management
- Diagnosis recording
- Medical data encryption

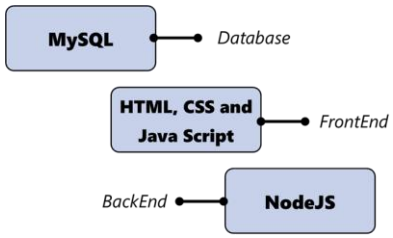


Fig 6.1: Methodology

Implementation technologies used are JavaScript programming language and MySQL for database. This methodology can be used to develop a secure and user-friendly PIS that meets the needs of healthcare organizations, medical professionals, and patients.

III. DESIGN & IMPLEMENTATION

The Patient Information System (PIS) is developed using JavaScript programming language and MySQL database management system. The system design will encompass patient registration, appointment management, prescription recording, and medical data encryption. The implementation process involves developing the PIS website, testing the system, and deploying it to production, ensuring secure programming practices and encryption to protect medical

data. The system aims to streamline administrative processes, enhance patient care quality, and ensure data security and accessibility in healthcare facilities.

A. Proposed Sytem

This system (PIS) is a software application that helps healthcare providers manage their patients' medical records, appointments, and other administrative tasks. A proposed PIS is a web-based or mobile app that is accessible to all authorized users, including doctors, nurses, and administrative staff.

The proposed PIS includes the following features:

- Patient management: This module allows users to create and manage patient profiles, including demographic information, medical history, and current medications.
- Appointment scheduling: This module allows users to schedule and manage patient appointments, including sending reminders and tracking attendance.
- Medical records management: This module allows users to store and manage patient medical records, including clinical notes, lab results, and imaging reports.
- Reporting and analytics: This module generates reports and analytics on patient data, such as appointment rates, missed appointments, and common diagnoses.

The system includes features such as patient management, appointment scheduling, and medical records management. The software requirements include MySQL for database management and Express.js

B. System Design

The system includes features such as patient management, appointment scheduling, and medical records management. The software requirements include MySQL for database management and Express.js for backend development. Visual Studio Code is used for working with Node.js, MySQL, HTML, and CSS in developing the system. The PIS is a comprehensive solution for healthcare facilities, with future benefits extending to health insurance companies, researchers, pharmaceutical companies, smart health tracking systems, and billing/payment processes.

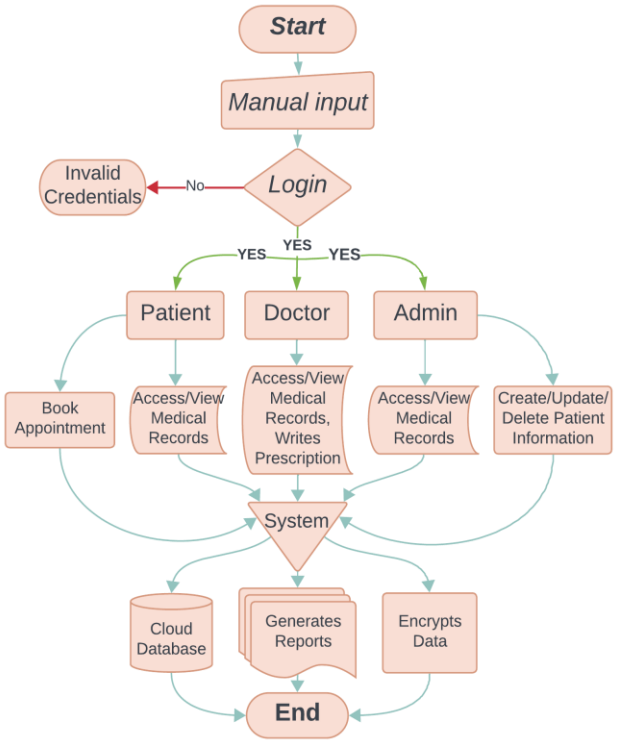


Fig. 1. **System Flow**

The above Flowchart describes the working of the system starting from its input as a user to getting the desired output.

IV. Conclusion

In conclusion, the Patient Information System (PIS) represents a comprehensive solution tailored to the specific needs of modern healthcare facilities. By prioritizing user-centric design, cutting-edge technologies, robust data security, efficient data management, and accessibility, PIS streamlines administrative processes while enhancing the quality of patient care. Leveraging advanced encryption protocols and stringent access controls, PIS ensures the confidentiality and integrity of patient data, fostering trust among both healthcare providers and patients. Ultimately, PIS stands as a testament to the transformative potential of technology in healthcare, promising improved patient outcomes, enhanced operational effectiveness, and a more patient-centered approach to healthcare delivery.

V. Future Scope

- Health Insurance Companies: Access to up-to-date patient information streamlines claims processing, fraud detection, and trend analysis, enabling tailored insurance plans and enhancing customer satisfaction.
- Researchers and Public Health Officials: Anonymized data aids in population health studies, disease trend identification, and assessment of healthcare interventions, facilitating evidence-based policies and advancing medical research.
- Pharmaceutical Companies: Real-world patient data supports clinical trials, medication adherence monitoring, and post-market surveillance, informing targeted marketing strategies and aiding in drug development efforts.
- Smart Health Tracking System: Demographics can be utilized for tracking a patient's vitals. It can also be used to

illustrate how one's state of health has changed during the years.

- Billing and Payment: Patient's expense throughout its ADT can be recorded with the help of this system, which later can be used as to claim the Medical Insurance.

ACKNOWLEDGMENT

We are grateful to Prof. Neelam Punjabi for her guidance and constant follow-up, as well as for giving us the essential information about the project and assistance in completing it. The project would not have been possible without the kind support and assistance of several individuals and all lab assistants at Vidyalankar Institute of Technology for their kind cooperation and encouragement in completing this research. We would like to express our heartfelt gratitude to every one of them. We would like to thank our parents for their encouragement and support.

REFERENCES

[1] Kotapati Saimanoj, Grandhi Poojitha, Khushbu Devendra Dixit, Laxmi Jayannavar, "Hospital Management System using Web Technology," TEST engineering and management, vol. 83, pp. 4493-4496, June 2020.

[2] Mwase Ali, Twinamatsiko Jerald, Mukuuma Kassim, "A Web-Based Hospital Management System For Iganga District Hospital, Uganda," IJARIIE, vol. 6, ISSN(O)-2395-4396, May 2020.

[3] Catherine Creighton, "A Literature Review on Communication Between Picture Archiving and Communication Systems and Radiology Information Systems and/or Hospital Information Systems," Journal of Digital Imaging, vol. 12, pp. 138-143, August 1999.

[4] PengWei Wang, ZhiJun Ding, ChangJun Jiang, and MengChu Zhou, Fellow, "Design and Implementation of a Web-Service-Based Public-Oriented Personalized Health Care Platform," IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, vol. 43, pp. 4, July 2013.

[5] Gunjan Yadav, Parth Lad, Parul Pandey, Tejaswi Kolla, "Advanced Hospital Database Management System," IJARCCE, vol. 5, Issue 4, April 2016.

[6] Ruchi Dumbre, Purva Raut, Bhagyshree mahamuni, Priyanka Khose, Prof.Jagruti Wagh, "Healthcare management system and domain search of nearest Medical services," IJISET, vol. 3 Issue 3, March 2016.

[7] Dr. Anupama R, Monica Pahwa, "Hospital Information Management System (HIMS)- A Study of Efficiency in Indian Scenario" Amity Management Review, 3(1), 40-49, 2013.

[8] Sejong Lee, Jaehyeon Kim, Yongseok Kwon, Teasung Kim, Sunghyun Cho, "Privacy Preservation in Patient Information Exchange Systems Based on Blockchain: System Design Study", Journal of Medical Internet Research, March 2022.

[9] Ashema Hasti, Esha Bisht, Akansha Rathi, Monika Chaudhary, "Hospital Management system", MATA SUNDRI COLLEGE FOR WOMEN.

[10] Sura abed Sarab, "Hospital Management System", University of Baghdad, College of Science, June 2020.

[11] Zhihong Liu, "Design and Implementation of Hospital Emergency Nursing Information Management System", ICSCSE, November 2016.

[12] Baki Koyuncu, Hakan Koyuncu, "Intelligent Hospital Management System (IHMS)", CICN, December 2015.

❑ ❑ ❑ ❑ ❑ ❑

AET 035

Skin Cancer Detection using TensorFlow in Python

Durva Detke, Madhushree Thatte, Sanika Nalawade
Guided by Prof. Suvarna Udgire

***Abstract*—The most prevalent disease in humans, skin cancer, is mostly diagnosed visually. A clinical screening is performed initially, followed by a dermoscopic examination, biopsy, and histological analysis may be necessary. Because skin lesions can vary widely in their appearance, automatically classifying them from images is a difficult challenge [36].**

Currently, skin cancer is among most frequent diseases in humans. This methodology-based Fourier spectral analysis uses filters such as classic, inverse and k-law nonlinear. The sample images are obtained by a specialist as a replacement spectral technique is developed and quantitative measurement in the complex pattern found cancerous skin spots. Finally, which spectral index calculated gets a variety of spectral indexes defined for carcinoma. Our results show a confidence of level 95.4%. carcinoma mainly occurs thanks to exposure to sunlight. Ozone is depletion and maintained chemical exposures in other factors involved in precipitating carcinoma. Mutations of the p53 gene involved UV-induced carcinogenesis. The P53 gene acts in vital development in SCC. The need for early diagnosis of the skin cancer is due sharp increase in cases of melanoma skin cancer the high death rate and the exorbitant cost of treatment are two major causes of the rise in skin cancer cases. the majority of the time it takes time to heal when cancer cells are manually found using image processing and machine learning techniques this project proposed a man-made carcinoma detection system after the images are segmented the features of the impacted skin cells are extracted using the features extraction technique the extracted features are stratified using a convolutional neural network classifier a deep learning-based technique skin cancer needs to be identified as soon as possible because it is a warning sign it takes a lot of time and money to perform the diagnostic manually nonetheless modern science has become advanced by using artificial intelligence an machine learning to make it easy to detect cancerous cells, specially convolution neural network is employed to detect cancerous cells more quickly and efficiently.

***Keywords*— melanoma, ABCD, CAD, TensorFlow, Confusion matrix**

I. INTRODUCTION (HEADING 1)

Skin Cancer is the most active type of cancer in the present decade [1]. Uncontrolled growth of abnormal skin cells is known as skin cancer. It happens when skin cells sustain nrepaired DNA damage, which sets off mutations or genetic flaws that cause the skin cells to multiply readily and form malignant tumours [2]. These gene mutations quickly and can cause skin cells to proliferate malignant tumours to form. DNA genetic changes in epidermal cells caused by UV light account for most cases of skin cancer., with early tumours often controlled by the immune system. However, compromised immune surveillance can allow the progression of malignant cell masses into tumours [3]. Melanoma is a malignancy of melanocytes, melanin (pigment) producing cells in the basal layer of the epidermis. Melanocytes are of neural crest origin and therefore express many signaling molecules and factors that promote migration and metastasis after malignant transformation. Despite representing only 1% of skin cancers, melanoma accounts for over 80% of skin cancer deaths. Melanoma accounts for 1.7% of global cancer diagnoses and is the fifth most prevalent cancer in the USA. Melanoma incidence is rising in developed, predominantly fair-skinned countries, growing over 320% in the USA since 1975 [4].



Figure 1. Medical Image of Melanoma Skin Cancer

II. LITERATURE SURVEY

Skin cancer killing millions of people across the world every day, must be diagnosed in its initial stages. The reason why people die of skin cancer is mostly due to misdiagnosis or rather ignorance, however not limited to. The paper represents the classification and detection of a skin anomaly as malignant or benign.

A. Background and research

There are several techniques used mainly for the identification and classification of the images. Md Ashraful

Alam Milton used digital image processing to classify malignant and benign skin lesions from dermoscopy images and a deep neural network to increase the dermatologist's efficacy [3]. Similar preprocessing image processing techniques of Dermoscopic images followed using neural network and fuzzy inference systems are used by B. Salah et al., 2011 [12].

Wavelet Transform along with different techniques for identification and classification of skin lesions. Most of the examples of the techniques are, texture analysis by wavelet transforms applied to RBIR/ Region Based Image Retrieval used in a paper presented by Hiremath, P. S. et al., 2006 [8].

Discrete Wave Transform, power spectral density graphics, PCA and ANN are used to elevate the classification accuracy by Kara, S. et al., 2007 [11]. On another note, K. Karibasappa et al., 2004 used ANN using Wavelet Transform Coefficients for facial recognition [9].

Kasmi, R et al., 2016 produced a paper in which the dermoscopic images are pre-processed using the dull Razor method and Gaussian filter for Image enhancement, median filter for noise control, colour-based K-means clustering for image segmentation, ABCD method (Asymmetry Border Colour Diameter) and GLCM (Gray Level Co-occurrence Matrix) for feature extraction, along with MSVM (Multiple Kernel Learning Support Vector Machine) for high accuracy [23].

ABCD (Asymmetry, Border, Colour, Diameter) based Computer Aided Design is used for Automated melanoma recognition by Ganster, H. et al., 2001 [24]. Also, ABCD is used with Artificial Neural Networks in Computer-aided melanoma skin cancer detection presented by Kanimozhi, T. et al., 2016 [2]. CAD mainly uses hair detection and exclusion inspired by the ABCD method for aiding the dermatologist in the accurate and efficient diagnosis of melanocytic skin lesions a method demonstrated in research conducted by S. Pathan et al., 2017 [31]. F. Gorunescu et al., 2007 used data mining with computer-aided design for an enhancement in a non-invasive diagnosis of cancer. Additionally, it is applied in EUSE (Endoscopic Ultrasound Elastography) [10].

El - Gamal, M. A. 2013 used classical Deep Neural Networks i.e., Artificial Neural Networks (ANN), Convoluted Neural Networks (CNN), Kohnen selforganizing Neural Networks (KNN), Generative adversarial Neural Networks (GAN) for skin cancer image classification [21].

Guergueb, T. et al., implement a clustering of deep models and an ensemble of deep learning to detect skin cancer [14]. Another such example of an ensemble-based Convolution Neural Network is also given by Qureshi, A. et al., 2022 [16]. A deep transfer learning technique based on the MobileNetV2 model proposed by Rashid, J. et al., 2022 performs better than the current techniques in terms of accuracy and computational efficiency [29].

Another approach using a Convolutional Neural Network

for accurate melanoma detection with an average precision of 0.88 is employed by Codella, N. C. F et al., 2017 [25]. Shaikh, J. et al., 2022 use deep learning with the CNN approach to improve accuracy in the detection of skin cancer in the paper Improved skin cancer detection using CNN [7]. Kwiatkowska, D. et al., 2021 in their published model, explore the use of ensemble models, a combination of multiple CNNs further improving the sensitivity and specificity, thus showing a significant improvement in melanoma detection [30].

Guo, L. et al., 2020 in their published paper Effective Melanoma Recognition Using Deep Convolutional Neural Network with Covariance Discriminant Loss, tested on the International Symposium on Biomedical Imaging (ISBI) 2018 Skin Lesion Analysis dataset with a sensitivity of 0.942 [33]. MBDCNN (Mutual Bootstrapping Deep Convolutional Neural Networks) used by Xie, Y. et al., 2020 facilitates simultaneous segmentation and classification of dermoscopic images of skin legions [5].

A paper presented by Ashraf R. et al., 2020 proposed an

approach that utilizes a Transfer learning-based Convolutional Neural Network (CNN) architecture. Specifically, it employs a pre-trained CNN model with data augmentation, focusing on Region of Interest (ROI) images extracted from the DermIS and DermQuest datasets. This method aims to enhance skin cancer detection [1]. Using a combination of advanced AlexNet Convolutional Neural

Network and morphological techniques achieves a significant pixel-level accuracy and sensitivity of 0.942 and 0.923 in skin lesion classification suggested by Bharathi, G. et al., 2022 [13].

In this project, we use TensorFlow, Keras and other such libraries to preprocess the dataset as well as classify the dataset into benign (nevus and seborrheic keratosis) and malignant (melanoma). The dataset retrieved is a small part

ISIC archive dataset. The training samples are 2000 and the validation samples are 150. We use an InceptionV3 model which returns an image classification model. We make use of the TensorFlow Hub library for transfer learning. with pretrained weights. Given that it is a binary classification problem, the model is built with binary cross-entropy loss and accuracy as a metric. To evaluate the model, a confusion matrix is drawn due to the dataset's high level of imbalance.

III. PROPOSED METHODOLOGY

To correctly classify the input image into benign or malignant class and further into nevus and seborrheic and melanoma subclass, identifying markers, colours, texture, extracting features, etc is essential. The primary care of the image is the preprocessing which essentially means making use of digital image processing.

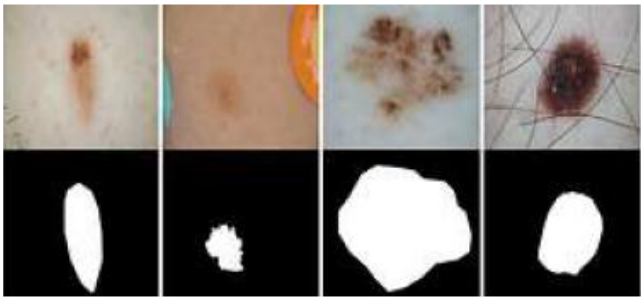
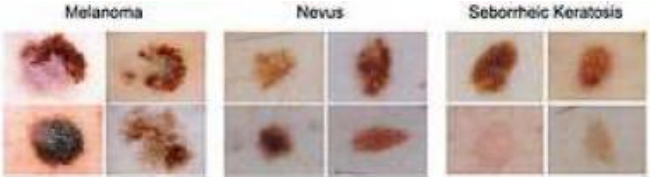


Figure 2. Legion Mask Segmentation



Attribute Detection Four Classes



Diagnostic Classification 3 Classes

A. Data Collection

The data is collected from a small part of the ISIC archive dataset. The data set contains 2000 testing samples and 150 validation samples in all. It contains images from both malignant (indicating cancerous lesions) and benign noncancerous lesions) skin lesions. These images are crucial for building and fine-tuning the skin cancer detection model.

B. Data Loading and Processing

Python libraries simplify data handling and complex tasks. The libraries used are Pandas, Numpy, Matplotlib, Scit-learn (Sklearn), and TensorFlow. The data is loaded and organized using a data frame, the labels are based on file paths. A binary model is used for which the labels are converted into binary values (0 for benign, 1 for malignant).

C. Data Exploration

A data imbalance is observed with less than 2% of malignant skin cancer images. Random images of both categories are visualized using Matplotlib.

D. Model Architecture

The proposed system utilizes EfficienNet architecture as a backbone. With the help of Transfer learning with pretrained weights tinstructed on the ImageNet dataset successively fine-tuned. To improve the model's robustness and prevent overfitting data is augmented(rotation, zoom, etc.)

E. Model Training

The dataset of 2000 training and 600 testing datasets are split into. training and validations set. The model is trained using the training data and performance metrics(e.g., accuracy, loss, specificity, precision, etc) is monitored during training. The model is evaluated using the validation set. For the same, metrics such as the ROC (Receiver Operating Characteristic) curve, AUC (Area Under the Curve), Confusion matrix, sensitivity and specificity are calculated. To optimize the model's performance, Hyperparameter Tuning i.e., experiment with hyperparameters (e.g., learning rate, batch size).

F. Use in Real Time

The program is integrated into the backend and the user interface is made simple for the ease of use of layman. The

interface also redirects based on the user's choice to "Practo" for a physician consult.

IV. PROCESSED MODEL USED FOR THE SUBJECT

1. Pre-process and collect the data: Collect a substantial and varied dataset of patient records, including diagnosis, symptoms, lab test results, medical history, and demographic data. Clean the information by eliminating copies, filling missing qualities, and revising mistakes. Encoding categorical variables, scaling numerical variables, and handling outliers are all part of the preprocessing of the data.
2. Visualize and investigate the data: Investigate the dataset utilizing elucidating measurements, representations, and component determination strategies. Determine the relevant characteristics that can assist in disease prediction. Dissect the dissemination and relationships of the elements concerning the sicknesses.
3. Selection and instruction of models: Select proper model architecture based on dataset size, complexity and available computational resources. Utilizing the appropriate hyperparameters and evaluation metrics, train the models on the engineered and pre-processed dataset. To evaluate the model's performance and avoid overfitting, employ cross-validation methods.
4. Model Assessment and Tuning: Utilize performance metrics like accuracy, precision, recall, and the F1 score to evaluate the trained models. Recognize the best-performing model and tune its hyperparameters utilizing lattice search or irregular inquiry procedures. Addressing class imbalance by oversampling the minority classes or using techniques like synthetic Minority Oversampling Technique (SMOTE).
5. Send the model to appery.io: When the model is prepared and assessed, convey it in an appery.io application that takes in client input thermoscopic image and returns the diagnosis of the lesion along with an accuracy rate of the model. Create a user friendly, interactive app that can handle a variety of situations and edge cases by making use of the user interface components in appery's cTell!!!.
6. Verify and test the app: Test the application utilizing image inputs, either from a picture gallery or as a real-time image. Utilize user feedback and ground truth labels to verify the app's results. Check to see that the app meets the project's requirements and constraints.
7. Keep up with and update the application: To find any flaws, errors, or enhancements, keep an eye on the app's performance and user feedback. Regularly update the app and model in response to new data, research findings, and user requirements. Make certain that the app remains useful and relevant to stakeholders. Test the application utilizing unit tests, incorporation tests, and client acknowledgment tests.

A. Feasibility Study

The technical, financial, and operational viability of the project would typically be evaluated in a feasibility study for skin cancer detection using TensorFlow in Python.

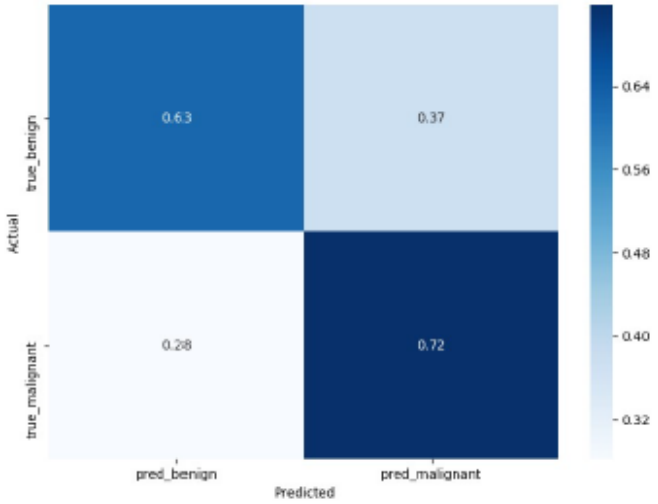
Technical adequacy: The technical viability of the project is contingent on the quality and availability of the data, the suitability and accuracy of the deep learning algorithms, and the app's framework's compatibility and scalability. With sufficient samples for each disease, the data ought to be comprehensive, unbiased, and representative.

Deep learning algorithms ought to be able to solve classification problems involving multiple classes with high accuracy and interpretability. The input and output interface for the appery framework ought to be user-friendly and compatible with the deep learning models.

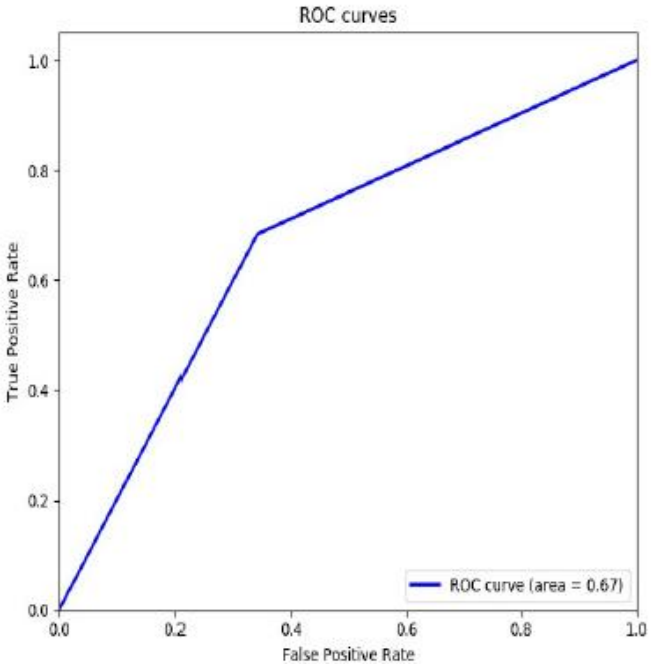
Functional achievability: The Web usability, dependability, and security, as well as the legal and ethical aspects of handling patient data, are critical to the project's operational viability.

V. EXPERIMENTAL RESULT AND DISCUSSION

Detecting skin cancer with deep learning using Tensor Flow in Python can produce stimulating outcomes in improving disease diagnosis and treatment. However, it is needed to address the limitations and challenges associated with the use of machine learning in healthcare. When the patient provides the parameter according to the disease, it will show whether the patient has the disease based on the disease identified. The parameters will show the range of the necessary values, and if any of the values is provided are out of that range, are invalid, or missing, an error message will be displayed, requesting the user to add the correct value.



Confusion Matrix



CONCLUSION

Skin Cancer is one of the most common and serious type of cancer. We demonstrated that a dermoscopic image classification using a set of deep neural network-based techniques can achieve competitive classification performance in need to identify skin lesions. This might develop into a highly automated and precise system for classifying dermoscopic images that dermatologists with experience could use in conjunction with one another.

REFERENCES

[1] Ashraf R., Afzal S., Rehman A.U., Gul S., Baber J., Bakhtyar M., Mehmood I., Song O.Y., Maqsood M. Region-of-Interest Based Transfer Learning Assisted Framework for Skin Cancer Detection. IEEE Access. 2020;8:147858–147871. doi: 10.1109/ACCESS.2020.3014701.

[2] Kanimozhi, T., & Murthi, A. (2016). Computer aided melanoma skin cancer detection using artificial neural network classifier. Singaporean Journal of Scientific Research (SJSR), Journal of Selected Areas in Microelectronics (JSAM), 8(2), 35-42. doi:10.1007/s11063-020-10364-y

[3] Md Ashraful Alam Milton, Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection Challenge, Autonomous Unversity of Barcelona, Campus de la UAB, Plaça Cívica, 08193 Bellaterra,Barcelona,Spainmdashrafulalam.milton@ecampus.uab.cat https://doi.org/10.48550/arXiv.1901.10802

[4] National Cancer Institute Melanoma of the Skin-Cancer Stat Facts. [(accessed on 10 May 2021)]; Available online: https://seer.cancer.gov/statfacts/html/melan.html

[5] Xie, Y., Zhang, J., Xia, Y., & Shen, C. (2020). A mutual bootstrapping model for automated skin lesion segmentation and classification. Ieee Transactions on Medical Imaging, 39(7), 2482-2493. https://doi.org/10.1109/tmi.2020.2972964

[6] Chicho, B. and Sallow, A. (2021). A comprehensive survey of deep learning models based on keras framework. Journal of Soft Computing & Data ining, 2(2). https://doi.org/10.30880/jscdm.2021.02.02.005

[7] Shaikh, J., Khan, R., Ingle, Y., & Shaikh, N. (2022). Improved skin cancer detection using CNN. International Journal of Health Sciences, 14347-14360. https://doi.org/10.53730/ijhs.v6ns2.8762

[8] Hiremath, P. S., & Shivashankar, S. (2008). Wavelet based co-occurrence histogram features for texture classification with an application to script identification in a document image. Pattern Recognition Letters, 9(9), 1182–1189. https://doi.org/10.1016/j.patrec.2008.01.012

[9] K. Karibasappa, S. Patnaik, Face Recognition by ANN using Wavelet Transform Coefficients, IE(I) Journal- P, 85,(2004), pp. 17-23.https://irjet.com/archives/V7/i5/IRJET-V7I5195.pdf

[10] Florin Gorunescu. (2007). Data Mining Techniques in Computer-Aided Diagnosis: Non-Invasive Cancer Detection. https://doi.org/10.5281/zenodo.1059721

[11] Kara, S., & Dirgenali, F. (2007). A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks. Expert Systems With Applications, 32(2), 632–640. https://doi.org/10.1016/j.eswa.2006.01.043

[12] Salah, B., Alshraideh, M., Beidas, R., & Hayajneh, F. (2011). Skin Cancer Recognition by Using a Neuro- Fuzzy System. Cancer Informatics, 10, 1–11. doi: 10.4137/CIN.S5950

[13] Bharathi, G. and Malleswaran, M. (2022). Detection and segmentation of melanoma skin cancer in dermoscopy images using modified Alexnet convolutional neural network-morphological methodology. Concurrency and Computation Practice and Experience, 34(25). https://doi.org/10.1002/cpe.7266

[14] Guergueb, T. and Akhloufi, M. (2022). Skin cancer detection using ensemble learning and grouping of deep models.. https://doi.org/10.1145/3549555.3549584

[15] Kutzner, H., Jutzi, T., Krahl, D., Krieghoff-Henning, E., Heppt, M., Hekler, A., … & Brinker, T. (2020). Overdiagnosis of melanoma – causes, consequences and solutions. JDDG Journal Der Deutschen Dermatologischen Gesellschaft, 18(11), 1236-1243. https://doi.org/10.1111/ddg.14233

[16] Qureshi, A. and Roos, T. (2022). Transfer learning with ensembles of deep neural networks for skin cancer detection in imbalanced data sets. Neural Processing Letters, 55(4), 4461-4479. https://doi.org/10.1007/s11063-022-11049-4

[17] Shaikh, J., Khan, R., Ingle, Y., & Shaikh, N. (2022). Improved skin cancer detection using CNN. International Journal of Health Sciences, 14347-14360. https://doi.org/10.53730/ijhs.v6ns2.8762

[18] Tan, T., Zhang, L., Neoh, S., & Lim, C. (2018). Intelligent skin cancer detection using enhanced particle swarm optimization. Knowledge-Based Systems, 158, 118-135. https://doi.org/10.1016/j.knosys.2018.05.042

[19] Trager, M., Geskin, L., & Liu, L. (2020). Biomarkers in melanoma and non-melanoma skin cancer prevention and risk stratification. Experimental Dermatology, 1(1), 4-12. https://doi.org/10.1111/exd.14114

[20] Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., Alsaiari, S. A., Saeed, A. H. M., Alraddadi, M. O., & Mahnashi, M. H. (2021). Skin Cancer Detection: A review using Deep learning techniques. International Journal of Environmental Research and Public Health, 18(10), 5479. https://doi.org/10.3390/ijerph18105479

[21] El-Gamal, M. A. (2013). Automatic skin cancer images classification. International Journal of Advanced Computer Science and Applications, 4(3). doi:10.14569/IJACSA. 2013.040342

[22] Monika, M. K., Vignesh, N. A., Kumari, C. U., Kumar, M., & Lydia, E. L. (2020). Skin cancer detection and classification using machine learning. Materials Today: Proceedings, 33, 4266–4270. https://doi.org/10.1016/j.matpr.2020.07.366

[23] Kasmi, R., & Mokrani, K. (2016). Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule. Iet Image Processing, 10(6), 448–455. https://doi.org/10.1049/iet-ipr. 2015.0385

[24] Ganster, H., Pinz, P., Röhrer, R., Wildling, E., Binder,

M., & Kittler, H. (2001). Automated melanoma recognition. IEEE Transactions on Medical Imaging, 20(3), 233–239. https://doi.org/10.1109/42.918473

[25] Codella, N. C. F., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., & Halpern, A. (2017, October 13). Skin lesion analysis toward Melanoma Detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC). arXiv.org. https://arxiv.org/abs/1710.05006

[26] Celebi, M. E., Kingravi, H. A., Uddin, B., Iyatomi, H., Aslandogan, Y. A., Stoecker, W. V., & Moss, R. H. (2007). A methodological approach to the classification of dermoscopy images. Computerized Medical Imaging and Graphics, 31(6), 362–373.https://doi.org/10.1016/j.compmedimag.2007.01.003

[27] R. J. Stanley, W. F. Stoecker, and R. H. Moss, "A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images," Skin. Res. Technol., vol. 13, no. 1, pp. 62–72, 2007.

[28] Khan, M. Q., Hussain, A., Rehman, S. U., Khan, U., Maqsood, M., Mehmood, K., & Khan, M. A. (2019). Classification of melanoma and nevus in digital images for diagnosis of skin cancer. IEEE Access, 7, 90132– 90144. https://doi.org/10.1109/access.2019.2926837

[29] Rashid, J., Ishfaq, M., Ali, G., Saeed, M. R., Hussain, M., Alkhalifah, T., Alturise, F., & Samand, N. (2022). Skin cancer disease detection using transfer learning technique. Applied Sciences, 12(11), 5714.

https://doi.org/10.3390/app12115714

[30] Kwiatkowska, D., Kluska, P., & Reich, A. (2021). Convolutional neural networks for the detection of malignant melanoma in dermoscopy images. Postępy Dermatologii I Alergologii, 38(3), 412–420.

https://doi.org/10.5114/ada.2021.107927

[31] S. Pathan, P. C. Siddalingaswamy, L. Lakshmi and K. G. Prabhu, "Classification of benign and malignant melanocytic lesions: A CAD tool," 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, India, 2017, pp. 1308-1312, doi:

10.1109/ICACCI.2017.8126022.

[32] Dabral, I., Singh, M., & Kumar, K. (2021). Cancer Detection Using Convolutional Neural Network. In Conference Proceedings of ICDLAIR2019 (pp. 290–298).

[33] Guo, L., Xie, G., Xu, X., & Ren, J. (2020). Effective Melanoma Recognition Using Deep Convolutional Neural Network with Covariance Discriminant Loss. Sensors, 20(20), 5786. https://doi.org/10.3390/s20205786

[34] GeeksforGeeks. (2023, May 24). Skin Cancer Detection using TensorFlow. GeeksforGeeks. https://www.geeksforgeeks.org/skin-cancer-detectionusing-tensorflow/

[35] Skin Cancer Detection using TensorFlow in Python.

(n.d.). https://www.tutorialspoint.com/skin-cancerdetection-using-tensorflow-in-python

[36] Esteva, A., Kuprel, B., Novoa, R. et al. Dermatologistlevel classification of skin cancer with deep neural networks. Nature 542, 115–118 (2017). https://doi.org/10.1038/nature21056

❑ ❑ ❑ ❑ ❑ ❑

AET 036

Automated Inventory Analyzer

Shreesh Nagral, Mallika Chourasia, Satyen Chavan, Dhruva Sahani, Vaibhav Kshirsagar
Department of Electronics and Telecommunication Engineering, University of Mumbai, India.

***Abstract* – This project introduces an innovative inventory management system for a business or organization. This is a software-based project which is helpful for the businesses that operate general stores, where the store owner keeps the records of sales and purchase for efficient management. This project eliminates the paperwork, human faults, manual delays and speeds up the process. We have created a user hierarchy with roles for both admin and employee. This project includes various features like the purchase order mail features that includes the automated mailing of purchase orders in case the inventory drops below a specific threshold. The inventory includes groceries, accessories, stationary items, electronics, etc. We have designed an enhanced dashboard which will help in the sales time series analysis and will also provide the monthly/yearly report KPIs, sales by category and inventory turnover. This project presents a machine learning approach for sales prediction in a retail setting. Leveraging historical sales data encompassing product attributes, store information, and temporal features, the model forecasts future sales trends. System that is less costly and suitable for a supermarket.**

***Keywords*– inventory, sales prediction, machine learning**

1. Introduction

In this digital era, there is a huge amount of information waiting to be processed and transmitted, which makes the further development and use of the database particularly urgent. With the increase in market competition, all businesses require efficient handling and management methods. With the development of technology, computer operations and management are increasingly simplified. At the same time, the rapid change of market economy and intense competition makes it inevitable for the supermarket owners to digitalize their process. But, sometimes, demand forecasts are not accurate, and it becomes challenging to determine the correct levels of inventory to maintain and establish appropriate safety stock levels that would ensure stock availability for customers. The case of overstocking a particular product led to wasted resources, expired products and takes up valuable warehouse storage space. The case of understocking leads to running out of products and therefore, lost sales and dissatisfied customers. This will hinder business growth because as your sales volume grows and your inventory expands, it will also be increasingly difficult to keep track of inventory levels on paper. To solve the above problems and improve the economic benefit, this project is aimed at designing an efficient inventory management system for a supermarket [3]. This project presents a machine learning approach for sales prediction in a retail setting. dynamics. To mitigate overfitting, various techniques are employed, including cross validation, regularization, and dropout. By providing accurate sales predictions, the project aims to empower retailers with valuable insights for optimizing inventory management and strategic decision-making [9].

2. Materials and methods

2.1 Methodology

Understanding System Architecture: The inventory management system architecture is designed for scalability, flexibility, and efficiency. Next.js forms the foundation for both frontend and backend, facilitating seamless communication. MongoDB is chosen for its document-oriented nature and scalability, making it ideal for diverse inventory datasets.

Efficient Data Storage: MongoDB efficiently stores and retrieves inventory data, regardless of its structure. Next.js API routes handle backend logic and API endpoints, ensuring consistency and efficiency in data manipulation and retrieval.

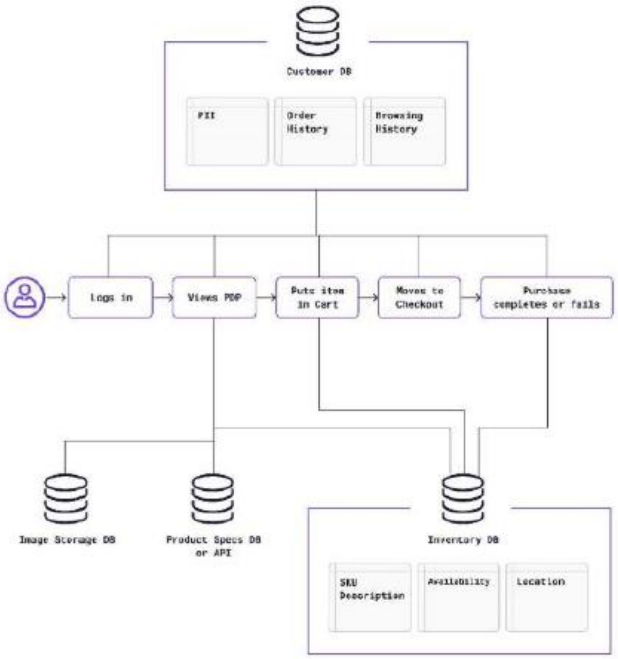


Fig 2.1 Inventory management system architecture

User-Friendly Frontend: The frontend, developed using Next.js, offers a modern and intuitive platform for managing inventory and accessing analytics. Server-side rendering enables fast page loads and smooth navigation, enhancing the user experience.

Predictive Analytics: The system employs predictive analytics to forecast future inventory needs accurately. Various techniques, including regression analysis and machine learning models, are used to optimize stock levels and reduce costs.

Graphical Insights: The Nivo library is integrated into the front end to provide users with graphical and analytical insights into inventory trends.

Customizable charts and visualizations enrich the user experience, enabling informed decision-making.

Efficient API Design: Dedicated API routes ensure consistency and efficiency in data retrieval and manipulation. By following best practices in API design, developers ensure smooth communication between client and server components.

2.2 Workflow

The flowchart represents a high-level view of a software system that could be used for retail management, e-commerce, or any inventory-related business. The system seems designed to track critical

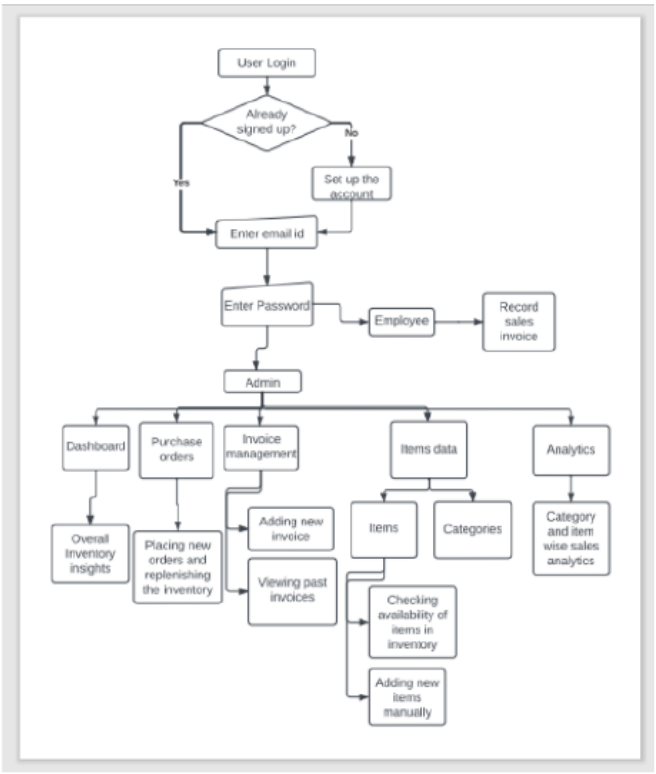


Fig 2.2 Inventory management user flow

aspects of the business workflow, from sales to inventory, and to provide decision-makers with actionable insights through analytics. The starting point for accessing the system, which could be a portal or an application interface. The decision point bifurcates the process based on whether the user has an account. Employee accounts gain access to sales invoicing, while administrators oversee various modules, including stock management and financial tracking. Users, such as store managers, access the inventory management dashboard upon login. They can add new products, update inventory upon receipt, and record sales transactions. Return and exchange requests are handled within the system, along with routine inventory monitoring and replenishment. Administrators receive inventory deliveries, verify them against purchase orders, and update the database. They prioritize orders, coordinate promotions, and evaluate supplier performance. Reports on inventory turnover and financial performance inform optimization efforts, driven by feedback from staff and customers.

3. Results

In this project, we have conducted a comparative analysis of four popular regression algorithms for sales forecasting: XGBoost, Support Vector Regressor (SVR), Random Forest Regressor and Artificial Neural Network (ANN) The objective of this analysis was to identify the most effective model for predicting sales data, considering factors such as predictive accuracy, generalization performance, and computational efficiency.

XGBoost stands out as a leading-edge machine learning algorithm prized for its speed, scalability, and accuracy in predictive modeling tasks. It leverages gradient boosting to construct a robust ensemble of decision trees, refining model performance iteratively by minimizing the loss function.

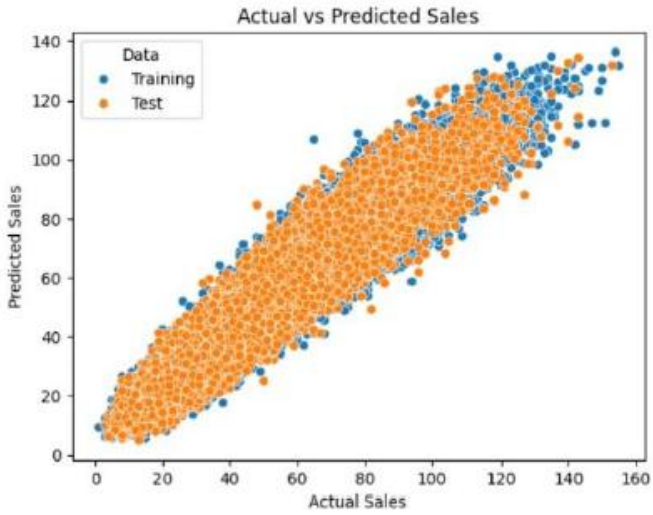


Fig 3.1 XGBoost training vs test

This iterative approach enables XGBoost models to uncover subtle patterns and interactions within the data, resulting in highly precise sales projections.

Support Vector Regression (SVR) is a favored method for sales forecasting, especially when faced with nonlinear or non-monotonic relationships between input variables and sales.

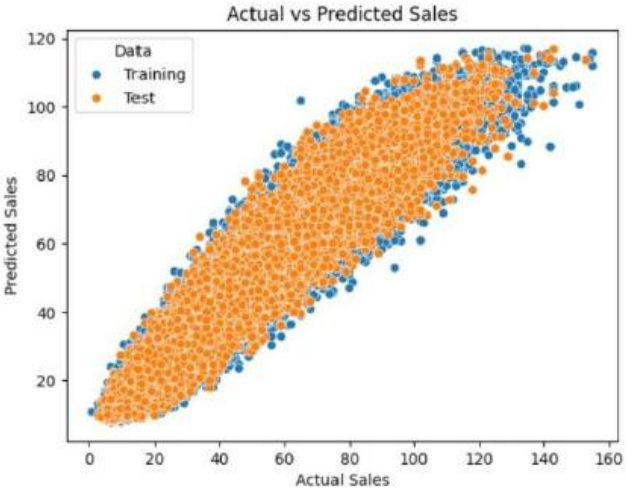


Fig 3.2 SVM training vs test

By leveraging support vectors to identify the optimal hyperplane, SVR models excel at capturing intricate patterns within the data and generalizing to new data points. With its capability to handle high-dimensional data and nonlinear relationships, SVR proves to be a valuable tool across diverse sales forecasting tasks.

Artificial Neural Networks (ANN) are a powerful tool for sales forecasting due to their ability to capture complex nonlinear relationships in the data. Artificial neural network (ANN) models have the capability to detect patterns and trends in vast amounts of historical sales data that conventional statistical techniques might miss.

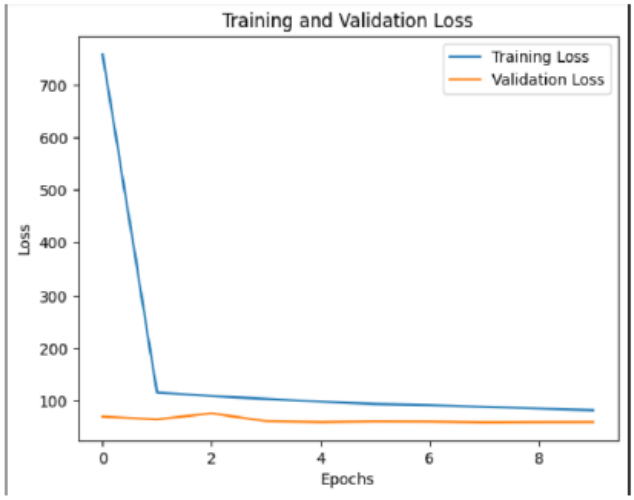


Fig 3.3 ANN training loss and validation loss

To increase forecasting accuracy, ANN models can also include a variety of features, including time-series data, product attributes, and market conditions.

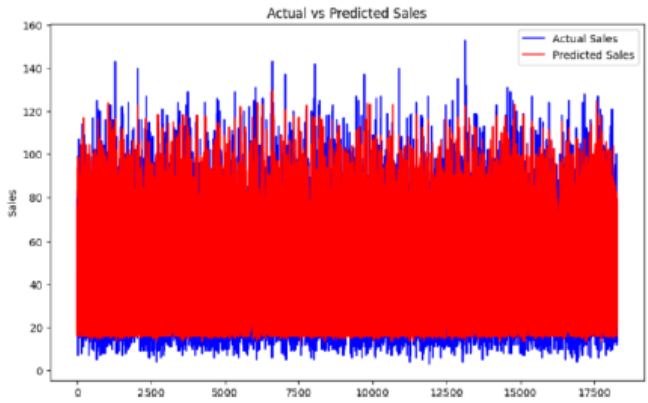


Fig 3.4 ANN actual sales vs predicted sales

Random Forest Regression is a versatile machine learning method for predicting numerical values. By combining predictions from multiple decision trees, it minimizes overfitting and improves accuracy.



Fig 3.5 Random Forest Regression training vs test

This technique is adept at capturing complex relationships between input variables and is well-suited for large datasets with diverse features and nonlinear relationships, making it ideal for sales forecasting.

4. Conclusion

This project implements a better inventory management system that solves all the problems that occur in inventory and helps the company to face the problems by following proper techniques and controlling. All around the world, inventory managers are faced with the complexities of simultaneously striving to retain constant manufacturing operation, render quality and adequate services to customer and keep goods at optimum level as well as having proper inventory management policy in place. Due to local and global competition, any company that must survive in the competitive world must take an advantage of effective management of its inventory seriously therefore the use of digital approach for effective and efficient inventory management system has become indispensable. This project includes various features like the purchase order mail features that includes the automated mailing of purchase orders in case the inventory drops below a specific threshold. This project has fostered a culture of collaboration, innovation, and continuous improvement. Cross-functional teamwork, stakeholder engagement, and a shared commitment to excellence have been instrumental in overcoming challenges, driving progress, and achieving our project goals. Ultimately, a successful inventory management system project delivers tangible benefits such as reduced carrying costs, improved inventory accuracy, increased productivity, and enhanced customer satisfaction. By harnessing the power of technology and best practices in inventory management, organizations can gain a competitive edge in today's dynamic and demanding business environment.

5. Acknowledgement

The authors take this opportunity to express our profound gratitude and deep regards to our guide Professor Vaibhav Kshirsagar for his guidance, monitoring and constant encouragement throughout the course of this project work. We also take this opportunity to express a deep sense of gratitude to Prof Dr. Varsha Turkar, HOD of E.X.T.C. Dept. for her cordial support, valuable information, and guidance, which helped us in completing this task through various stages.

REFERENCES

[1] "Inventory Management Techniques for Grocery Shops in Bangladesh: A Study on a Grocery Shop in Khulna City" by Md Moslah Uddin, Mhd. Jahangir Alam in September 2023 from International Journal of Science and Business 28

[2] "Strategic Inventory Management Practices and the Performance of Supermarkets" in Nairobi County, Kenya by Robert Arasa, Achuora John Odoyo in April 2020 from European Journal of Business Management and Research

[3] "Supermarket Sales and Inventory Management System: A Complete Case Study" by Renz Victor D. Cleofe, Richard Joven Jefferson B. Delos Reyes, John Ruzzene

A. Baylon, Keira Maria S. Sarmiento, Ashley Nidyn A. Garachico, Keith Crystal C. Maligaya from Good Shepherd Academy, Philippines

[4] "Study of Inventory Techniques in E-commerce Grocery to Minimize Out of Stock and Increase Customer Shopping Experience" by Soum Suvra Das in International Journal of Management

[5] "Simulation of inventory management systems in retail stores: A Case Study" by Puppala Sridhar, C.R. Vishnu, R Sridharan

[6] "Inventory management and forecasting in retail: A review" by R. J. Hyndman and G. Athanasopoulos.

[7] "A study of inventory management techniques applied by organized retailers' and its effect on customer satisfaction and retailer's financial performance with special reference to Bhopal and Indore City" by Dr. Pratyush Tripathi and Deepak Tiwari in International Journal of Innovation and Applied Studies.

[8] "Warehouse inventory management system using IoT and open-source framework" by B. Sai Subrahmanya Tejesh, S. Neeraja in Alexandria Engineering Journal.

[9] "Predictive Analysis for Big Mart Sales Using Machine Learning Algorithms" by Ranjitha P and Spandana M. in ICICCS

❑ ❑ ❑ ❑ ❑ ❑

AET 037

Deep Learning-based Plant Identification App: Leveraging Efficient Net Model and Modern Web Technologies

Dr. Sheetal Mapare
Vidyalankar Institute of Technology Mumbai, India
sheetal.mapare@vit.edu.in

Arnav Gonge
Vidyalankar Institute of Technology Mumbai, India
arnav.gonge@vit.edu.in

Siddhi Manche
Vidyalankar Institute of Technology Mumbai, India
siddhi.manche@vit.edu.in

Vedant Patwardhan
Vidyalankar Institute of Technology Mumbai, India
vedant.patwardhan@vit.edu.in

***Abstract*— To slow down the rate of decreasing biodiversity, the process of automated plant identification plays a crucial role. It is also one of the keys factors in agriculture, research, and ecology. In the following paper, we propose a recent approach to tackle plant identification engaging deep learning techniques and state-of-the-art web and mobile development frameworks. Our application leverages the Efficient Net model trained on the Plant CLEF 2016 dataset along with an Indian medicinal plant image dataset. The backend was created using Fast API and containerized using Docker, while the frontend interface utilized Android Studio for development. After thorough testing and validation, we demonstrate the effectiveness of our application in plant identification tasks resulting in an 82.2% Top-1 Accuracy. Our findings provide a pillar in using the Efficient Net model for plant identification tasks.**

Keywords—Plant Identification, Deep Learning, Efficient Net

I. INTRODUCTION

With about 382,000 species of plants in the world, plant identification is a monotonous process requiring accurate domain knowledge. This process is of utmost importance as many plants provide medicinal properties that are vital for treatment as well as drug synthesis. Advancements in computer-vision, ML and DL have opened a pathway to automate the identification of plant species. Be that as it may, the effectiveness of these solutions varies due to the massive amount of plant species to identify. An effective solution would require huge amounts of high-quality training data and intensive compute resources.

Recent research has shown significant advancements in plant leaf disease detection methodologies. Guo et al. [1] proposed a comprehensive approach, integrating a deep learning algorithm—Region Proposal Network for leaf localization—and the Chan-Vese algorithm for symptom segmentation, alongside transfer learning for classification. Their work highlights the effectiveness of this combined approach in achieving accurate plant leaf disease detection. Building upon this, M.H. Saleem et al. [2] demonstrated that deep learning architectures such as ResNet, GoogLeNet, and LeNet consistently outperform traditional ML methods like SVM and Random Forest in this domain. Moreover, Mohanty et al. [3] showcased the practical applicability of DL models in plant disease detection through a mobile application-based solution. Collectively, these studies underscore the growing prominence of deep learning techniques in revolutionizing image-based plant disease detection.

The identification of medicinal plants holds significant importance in traditional medicinal practices, with over 80% of the global population relying on herbal remedies. Reference [4] proposes a methodology for the identification of medicinal plants based on the visual characteristics of leaves and flowers, as suggested by Jayalath et al. The methodology comprises four main steps: image enhancement, segmentation, feature extraction, and classification using Back Propagation Artificial Neural Networks (ANN). Notably, the study emphasizes the use of high-quality images collected through interviews with Ayurvedic doctors and botanical gardens, resulting in a dataset of 5000 images with uniform resolution and size. However, while the inclusion of flowers facilitates plant identification, the reliance solely on high-quality images may limit the scalability of the approach, considering the challenges associated with obtaining such images for all plant species.

On the other hand, reference [5] focuses on the utilization of deep learning, specifically convolutional neural networks (CNNs) like Densenet121, for the automated identification of medicinal plant leaves, as demonstrated by Rao et al. The study highlights the importance of TensorFlow and Keras in implementing the Densenet121 architecture, proving the significance of deep learning frameworks in modern plant identification research. By leveraging deep learning techniques, this approach offers a scalable solution for plant identification, potentially overcoming the limitations of image quality and availability encountered in traditional methods.

Combining the strengths of both approaches—visual feature extraction and deep learning—holds promise for advancing the identification of medicinal plants, crucial for traditional medicinal practices and biodiversity conservation efforts. However, despite the abundance of medicinal plants worldwide, their accurate identification and understanding of their uses remain challenging, amplified by the scarcity of high-quality data. Acknowledging these challenges, our contributions aim to address these limitations and enhance the efficiency of plant identification methods. We have developed a model capable of accurately predicting over 588 species of plants, utilizing both visual features and DL techniques. Furthermore, to facilitate accessibility and usability, we have created an intuitive application that enables users to run predictions on images directly from their own devices. In addition to predictive capabilities, our application incorporates a valuable feedback feature. This feature allows users to suggest the correct identification of a plant if initially misidentified, contributing to the continuous improvement and refinement of our model. By integrating user feedback and continuously updating our dataset, we strive to enhance the accuracy and reliability of our plant species recognition system, thereby supporting the sustainable utilization of plants having medicinal properties and the preservation of biodiversity.

II. METHODOLOGY

A. System Architecture

The system architecture involves deploying a plant species prediction model on an AWS EC2 instance, facilitated by Docker containerization [6]. The DL model is packaged into a Docker container for efficient deployment and management on the EC2 backend. Users interact with an Android app frontend to upload plant images for identification. These images are processed by the Dockerized model running on the EC2 instance, which provides prediction services via API endpoints. Predictions, along with user history and plant information, are stored in an AWS RDS database. Additionally, AWS S3 is utilized for image storage. This architecture ensures seamless integration, scalability, and ease of deployment, allowing for efficient identification.

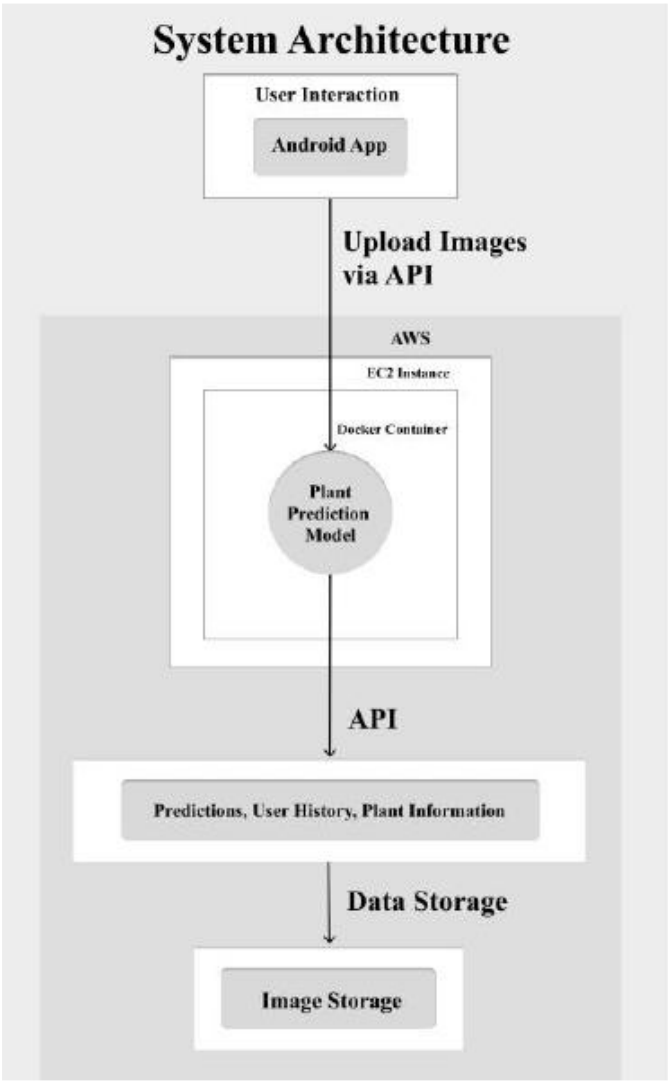


Fig. 1. Proposed System Architecture

B. Datasets

1) PlantCLEF 2015:

Aiming to develop an image-based plant identification system, PlantCLEF 2015 [7] is a dataset designed to test systems and methodologies on a massive scale that can be adjusted to real-world scenarios. The dataset was created using a community platform that allowed participation in 2011 and had thousands of participants. PlantCLEF 2015 is a collection of carefully chosen photos from various cameras, contributors, locations, seasons, and unique plants. More specifically, the PlantCLEF 2015 dataset consists of 113,205 images from 41,794 observations of 1000 different kinds of ferns, plants, and trees that are found throughout Western Europe. Every image has a binomen linked with it and corresponds to one of the seven viewing categories ("whole plant", "fruit", "leaf", "flower", "stem", "branch", and "leaf scan") in the meta-data. For our model, we utilized the PlantCLEF 2015 dataset as the primary dataset. Specifically, we focused on images labeled with the leaf and leaf scan tags from a smaller set of over 553 plant species. This subset was chosen to ensure a diverse representation of plants suitable for training our model. The dataset comprised a total of 19,305 training images and 4,956 validation images. By selecting images from these specific categories and species, we aimed to train our model to accurately identify plant species on the basis of leaf characteristics.

TABLE I. IMAGES EXTRACTED FROM PLANTCLEF 2015

Images	Images	Images	Images
Train	19305	9664	9641
Validation	4956	2414	2542

2) Indian Medicinal Plant Dataset:

To support the development of our mobile application tailored for the Indian Subcontinent, the availability of a comprehensive dataset of Indian medicinal plant species was important. This dataset [8] comprising of 5,945 images, sourced from various regions of Karnataka and Kerala, was utilized in our research. These images capture the diverse characteristics of forty plant species, including variations in resolution, illumination, background, and seasonal conditions Captured in real-time using smartphones, these images provide valuable insights for researchers interested in algorithmic models based on image processing, machine learning, and deep learning.

TABLE II. IMAGES EXTRACTED FROM INDIAN MEDICINAL PLANT DATASET

Images	Total
Train	4756
Validation	1189

C. Model Training

We employed the EfficientNetV2S model [9] for our image classification task, utilizing pre-trained weights from the ImageNet dataset [10] as the base for transfer learning. To adapt the model to our specific task, we added a dropout layer with a dropout rate of 0.2 to prevent overfitting. Subsequently, we appended a new softmax layer to the model to perform classification. During training, we employed a staged approach to fine-tune the model. Initially, we froze all layers except the newly added top layer and trained the model for 50 epochs. Following this, we gradually unfroze the last 156 layers of the network while keeping the batch normalization layers frozen. This fine-tuning strategy allows the model to learn task-specific features while retaining the learned representations from the pre-trained weights. The model was trained for an additional 50 epochs after unfreezing the specified layers. This approach aims to strike a balance between leveraging pre-trained features and adapting the model to our specific classification task.

1) Focal Loss:

The focal loss function, pioneered by Lin et al. [11], has emerged as a potent solution for handling class imbalance in classification tasks, particularly in computer vision applications. This loss function dynamically adjusts the contribution of individual training examples based on their predicted probabilities, effectively down weighting the influence of well-classified instances, and focusing more on those that are harder to classify. By introducing a modulating factor controlled by a parameter known as gamma, the focal loss provides a flexible mechanism to mitigate the effects of class imbalance and improve the performance of machine learning models.

In our research, we employed the focal loss, specifically the Categorical Focal Crossentropy variant, to address class imbalance in our dataset. Notably, we calculated class weights (alpha values) based on the class frequencies in the dataset. Each class was assigned a unique alpha value proportional to its class count, allowing the model to prioritize learning from minority classes effectively. This approach aimed to rectify the imbalance in class representation and improve the model's ability to generalize across all classes.

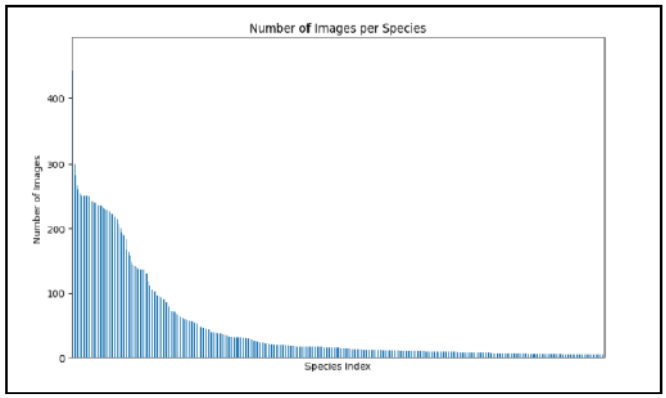


Fig. 2. Class counts for our merged datasets. Class indices are not displayed on the x-axis for conciseness.

2) Optimizer and Learning Rate:

In our research, we implemented a Cyclical Learning Rate (CLR) scheduler to optimize the training process of our model. CLR, proposed by Smith [12], is a learning rate policy that cyclically adjusts the learning rate between two boundaries over a predefined number of iterations. This dynamic adjustment enables the model to explore a wider range of learning rates during training, potentially leading to faster convergence and improved generalization performance.

The Triangular Cyclical Learning Rate scheduler, inspired by CLR, operates by oscillating the learning rate between an initial value and a maximal value in a triangular fashion over a specified number of steps. This cyclic variation in the learning rate allows the model to effectively navigate the optimization landscape and discover better solutions. For our experiments, we configured the scheduler with a starting learning rate of 10-3 and a maximum learning rate of 10-2 for training only the softmax layer. We set the step size to 1504, determining the frequency of iterations within each cycle. After unfreezing the layers, we set the initial and maximal learning rates to 10-4 and 10-5 respectively. For this iteration the step size was set to 6016. By utilizing this cyclical learning rate strategy, we aimed to enhance the training dynamics of our model and potentially achieve better convergence and generalization.

D. Backend Server

Our system is built using FastAPI, a modern web framework for building APIs in Python. It utilizes TensorFlow for machine learning functionalities, including model loading and predictions. The system architecture includes AWS services such as S3 for image storage and RDS for database management.

We implement endpoints for user registration, login, image classification, feedback submission, and retrieval of user history. These endpoints handle requests from clients and interact with the database and model for processing. Our backend system is containerized using Docker for easy deployment and scalability. The Docker container encapsulates the entire backend application along with its dependencies, ensuring consistency across different environments. Subsequently, the Dockerized backend is deployed on an AWS EC2 instance, offering a robust and scalable infrastructure for serving API requests. EC2 instances provide flexibility in terms of computing power and storage capacity, enabling our system to adeptly handle varying workloads. Additionally, the images from user history are stored in S3, while all user-related information resides in the RDS, ensuring efficient management and retrieval of data. Overall, our system provides a user-friendly interface for users to classify medicinal plant species using images captured from smartphones, with the capability to store user interactions for further analysis and improvement.

E. Database

The MySQL database schema facilitates various functionalities essential for the application's operation. It includes user management capabilities, allowing for the registration and authentication of users. Additionally, the database stores user interaction history, enabling the tracking of user activities and predictions made by the application. Feedback collection from users is also supported, providing a mechanism for users to submit their opinions or comments regarding the application's performance or content.

Moreover, the schema includes tables for storing information about plant species, such as their names, common names, and uses, which is integral for the application's core functionality of plant species identification and information retrieval. When the model runs a prediction on the image, the class index is queried from the database to retrieve additional information about the predicted plant species, enhancing the user's experience by providing comprehensive details about the identified plants.

F. Mobile Frontend

For the mobile frontend of our plant identification application, we utilized Flutter, a framework for building cross-platform mobile applications using the Dart programming language. Flutter offers a set of pre-built widgets and tools that allows to develop visually appealing and responsive user interfaces. The user interface (UI) of our mobile application was designed to be interactive and user-friendly, ensuring a seamless experience for users.

We employed Flutter's widget library to provide easy access

to key functionalities such as image upload, prediction interface and feedback submission. The feature allows users to either capture real-time photos using their mobile camera or select an image from their gallery. This functionality leverages Flutter's image picker plugin to provide a seamless image selection experience.

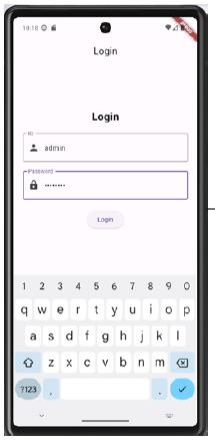


Fig. 3. Login Page



Fig. 4. Image Selection Page

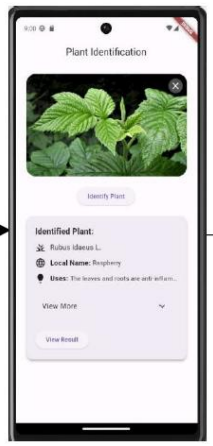


Fig. 5. Plant Identification

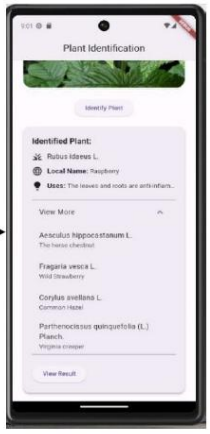


Fig. 6. Detailed information of identified plants

Once users upload an image, the mobile application communicates with the backend server to process the image using our trained model. The prediction results, including the top predicted plant species and their corresponding confidence scores, are then displayed to the user within the application's interface. We utilized Flutter's text and layout widgets to present this information in a clear and visually appealing manner.

In addition to identifying plants, our mobile application includes a feedback feature that allows users to provide corrections or additional information if a plant is misidentified. Users can easily submit feedback directly from the application, contributing to the continuous improvement of our plant identification model. This functionality was implemented using Flutter's form widgets and HTTP client for communication with the backend server.

To enhance user engagement and provide a personalized experience, our mobile application includes a user history feature that maintains a record of past plant identification requests and their corresponding results. This feature enables users to review their previous interactions with the application. Flutter's local storage capabilities were utilized to store and retrieve user history data within the application.

Overall, the mobile frontend of our plant identification application offers a seamless and intuitive user experience, leveraging Flutter's capabilities to deliver a responsive and visually appealing interface. By combining efficient backend processing with a user-friendly frontend, our application empowers users to identify plant species accurately and contribute to the refinement of our model through feedback submissions.

III. EXPERIMENTS AND RESULTS

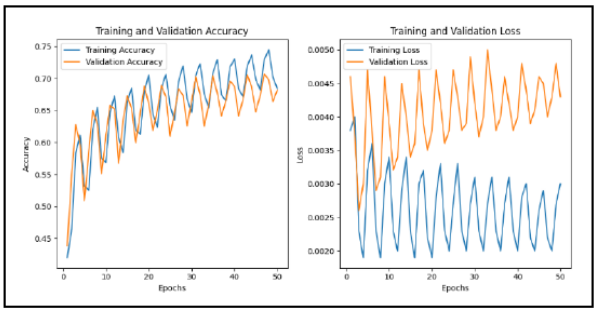


Fig. 7. Training and Validation Loss Trends (Softmax Layer)

A. Initial Training Phase

During the initial training phase, focusing on training only the softmax layer for the first 50 epochs, the model exhibited a noticeable improvement in performance. The training accuracy steadily increased to approximately 74%, indicating effective learning from the training dataset. Simultaneously, the validation accuracy, although slightly lower, reached a commendable 71%, demonstrating the model's ability to generalize well to unseen data. The corresponding loss metrics showed consistent convergence, with the training loss decreasing to around 0.0022 and the validation loss stabilizing at approximately 0.0040.

B. Unfreezing Layers

Subsequently, as the last 156 layers were unfrozen while keeping batch normalization layers frozen, the model underwent further refinement to enhance its predictive capabilities. Over the subsequent 50 epochs, the model's accuracy experienced a notable boost, reaching approximately 92% on the training dataset and 82% on the validation dataset. This improvement suggests that unfreezing additional layers facilitated the model in capturing more intricate patterns and nuances present in the data. Moreover, the loss metrics continued to decline, with the training loss diminishing to approximately 0.0001 and the validation loss converging around 0.0013.

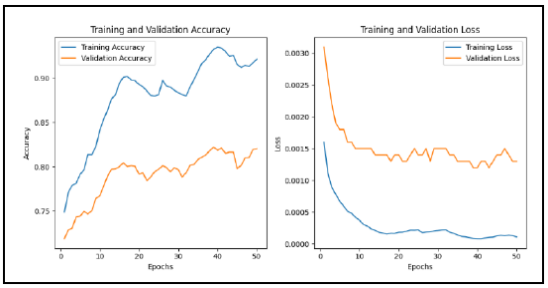


Fig. 8. Training and Validation Loss Trends (After Unfreezing 156 Layers)

C. Precision, Recall, F1 Score, and Support:

Evaluation of precision, recall, F1 score, and support across all classes revealed promising results. The weighted average

precision (1), recall (2), and F1 score (3) were computed to be 0.84, 0.84, and 0.83, respectively, indicating a balanced performance across different classes. The cumulative support (4), representing the number of samples, was found to be 6145, affirming the model's robustness and reliability across a diverse range of medicinal plant species.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (1)$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (2)$$

$$F1\ Score = 2 \times \frac{Precision\ x\ Recall}{Precision + Recall} \quad (3)$$

$$Support = Total\ no.\ samples \quad (4)$$

IV. CONCLUSION

In conclusion, our research presents a comprehensive approach to leveraging machine learning for the classification of medicinal plants using mobile-captured images. Through the creation and implementation of a sophisticated model architecture, we have demonstrated promising results in accurately identifying various plant species, which holds major implications for fields such as agriculture, pharmacology, and botanical research. Our system, built on Fast API and Tensor Flow, offers a user-friendly interface for seamless interaction, while AWS services provide robust backend infrastructure for scalability and reliability.

The evaluation of our model's performance showcases commendable accuracy and generalization capabilities, particularly evident during the initial training phase focusing on the softmax layer and subsequent fine-tuning of additional layers. Precision, recall, and F1 score metrics further attest to the balanced performance across different plant species, underscoring the model's robustness and reliability.

V. FUTURE WORK

Moving forward, our research suggests several areas for future exploration and enhancement:

A. Expansion of Dataset:
Acquiring a more extensive and diverse dataset, including images spanning various plant ages and growth stages, is essential for improving model generalization and accuracy.

B. Integration of Plant Age Information:
Incorporating metadata on plant age into the dataset and training process can facilitate better classification accuracy, enabling the model to adapt to variations in plant morphology over time.

C. Exploration of Transfer Learning Techniques:
Further investigation into transfer learning methods, particularly utilizing pre-trained models on larger datasets like ImageNet, could accelerate model training and enhance classification performance.

D. Crowdsourced Data Collection:
Engaging citizen scientists in crowdsourced data collection initiatives can help gather a broader range of images from different geographical regions, enriching the dataset and improving model robustness.

E. Integration of Multi-Sensor Data:
Exploring the integration of multi-sensor data, such as spectral or environmental information, alongside image data, could provide valuable context for plant classification, enhancing model interpretability and resilience to environmental variability.

REFERENCES

[1] Y. Guo, J. Zhang, C. Yin, X. Hu, Y. Zou, Z. Xue, and W. Wang, "Plant Disease Identification Based on Deep Learning Algorithm in Smart Farming," Discrete Dynamics in Nature and Society, vol. 2020, p. 2479172, Aug. 2020, doi: 10.1155/2020/2479172.

[2] M. H. Saleem, J. Potgieter, and K. Mahmood Arif, "Plant Disease Detection and Classification by Deep Learning," Plants (Basel), vol. 8, no. 11, p. 468, Oct. 2019, doi: 10.3390/plants8110468.

[3] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," Front. Plant Sci., vol. 7, pp. 1419, May 2016, doi: 10.3389/fpls.2016.01419.

[4] D. Jayalath, D. Nadeeshan, G. Amarawansha, H. Jayasuriya, and D. Nawinna, Identification of Medicinal Plants by Visual Characteristics of Leaves and Flowers, Dec. 2019. doi: 10.1109/ICIIS47346.2019.9063275.

[5] R. Upendar Rao, M. Sai Lahari, K. Pavana Sri, K. Yaminee Srujana, and D. Yaswanth, "Paper Id: IJRASET41190," IJRASET, Apr. 03, 2022. doi: 10.22214/ijraset.2022.41190.

[6] M. Openja, F. Majidi, F. Khomh, B. Chembakottu, and H. Li, "Studying the Practices of Deploying Machine Learning Projects on Docker," in Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering (EASE '22). Association for Computing Machinery, New York, NY, USA, 2022, pp. 190–200. doi: 10.1145/3530019.3530039.

[7] A. Joly et al., "LifeCLEF 2015: Multimedia Life Species Identification Challenges," in Experimental IR Meets Multilinguality, Multimodality, and Interaction, J. Mothe et al. (Eds.), Springer International Publishing, Cham, 2015, pp. 462-483. DOI: 10.1007/978-3-319-24027-5_46.

[8] B. R. Pushpa and Shobha Rani, "Indian Medicinal Leaves Image Datasets," Mendeley Data, Version 3, 2023, doi: 10.17632/748f8jkphb.3.

[9] M. Tan and Q. V. Le, "EfficientNetV2: Smaller Models and Faster Training," arXiv:2104.00298 [cs.CV], 2021.

[10] J. Deng et al., "ImageNet: A large-scale hierarchical image database," in Proceedings of the Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009, IEEE, Piscatway, NJ, USA, 2009, pp. 248–255.

[11] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2980-2988. doi:10.1109/ICCV.2017.324

[12] Smith, L. N. (2017). Cyclical Learning Rates for Training Neural Networks. Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), 464-472. doi:10.1109/WACV.2017.58

❑ ❑ ❑ ❑ ❑ ❑

AET 038

EcoFlow: Enhanced Water Efficient Farming

Sahil Baikar, Shubham Bochare, Bhavika Pawar, Riddhi Diwan
Department of Electronics and Telecommunication Engineering, University of Mumbai, India.

Abstract– **Agriculture is a critical component of global food production, but it is also a resource-intensive endeavour, particularly in terms of water usage. With increasing concerns about water scarcity and environmental sustainability, it is essential to develop innovative approaches that enhance water efficiency in farming. This paper explores Ecoflow, a comprehensive system for using water more efficiently in farming. It monitors soil and automates water release to address wastage, increase yields, and conserve water. This system combines precision irrigation techniques, real-time weather monitoring, and sensor-based data collection to optimize water distribution on agricultural fields. It leverages the Internet of Things (IoT) to provide farmers with valuable insights into soil moisture levels, crop health, and weather conditions. This data-driven approach enables farmers to make informed decisions about when and where to irrigate, ultimately reducing water wastage and improving crop yields. Ecoflow represents a significant step forward in sustainable agriculture, addressing the pressing need to reduce water consumption in farming while ensuring food security for a growing global population. By adopting this integrated approach, farmers can play a pivotal role in mitigating the environmental challenges associated with water use in agriculture, making Ecoflow a promising solution for the future of water-efficient farming.**

Keywords– **water-efficient, irrigation, IOT.**

1. Introduction

With agriculture being a foundation of sustenance worldwide, the efficient utilization of water resources has emerged as a critical imperative. As irrigation remains a dominant consumer of freshwater, optimizing its usage is important in mitigating water scarcity concerns and fostering agricultural sustainability.

The Ecoflow system represents a pioneering approach towards optimizing water usage for agricultural irrigation. By amalgamating soil monitoring capabilities with automated water release mechanisms, Ecoflow endeavors to reduce wastage, increase yields, and conserve water resources. This system operates on the premise of accurately finding crop requirements through continuous soil moisture level monitoring. Through precision water delivery tailored to the crops' needs, Ecoflow not only minimizes wastage but also fosters optimal crop growth, facilitating year-round cultivation and elevating farm yields.

Innovative research endeavors have emphasized the success of employing smart soil moisture monitoring systems in improving irrigation practices. Calibration and validation studies [2-3], have shed light on the feasibility of integrating low-cost capacitive moisture sensors into automated soil moisture monitoring systems, thereby facilitating precise water management. Additionally, evaluation studies [4-7], have provided insights into the performance characteristics and applicability of soil moisture sensing technologies across diverse soil types and environmental conditions.

Furthermore, the convergence of machine learning and IoT technologies has furnished novel avenues for enhancing soil monitoring capabilities and facilitating informed decision-making in agricultural practices [5]. Leveraging IoT-enabled systems for real-time monitoring of air temperature and soil moisture, as illustrated in reference [6], holds promise in augmenting irrigation efficiency and optimizing resource allocation in agricultural ecosystems.

Against this setting, our research endeavors to obtain the transformative potential of the Ecoflow system in revolutionizing water management practices in agricultural irrigation. By synthesizing insights from existing literature and leveraging advancements in sensor technologies and IoT, our contributions seek to address the multifaceted challenges confronting water-efficient irrigation. Through the development of an intuitive and adaptive irrigation system, we aspire to foster a more sustainable and environmentally responsible future, characterized by prudent water utilization practices and enhanced agricultural productivity.

2. Materials and methods

2.1 Programmable Logic Controller

The adopted Programmable Logic Controller featured an ESP32 microcontroller, renowned for its versatility in various electronic applications. Boasting an array of around 36 GPIO pins, including roughly 18 analog input pins with 12-bit resolution (approximately 0.81 mV/bit), the ESP32 provided ample interfacing capabilities. Its flash memory, available in configurations such as 4 MB or 8 MB, facilitated robust data storage. Operating at 3.3V, this microcontroller supported diverse projects with its integrated Wi-Fi and Bluetooth capabilities, enabling seamless wireless connectivity. With dual-core or single-core processing options, the ESP32 stood as a powerful and flexible solution for programmable logic control systems.

2.2 Sensors and modules

In this, three soil moisture sensors of the FC-28 model, equipped with four pins each (digital communication, analog communication, VCC, and ground), were utilized. The FC-28 sensors consist of a probe with two electrodes in direct contact with the soil and a compact module housing an LM393 comparator chip. The LM393 chip receives, processes, and transmits data from the probe to the microcontroller. Electric potential difference, resulting from soil resistance to current flow, is measured by the electrodes and presented as analog values. The sensor operates by applying an electrical signal through one electrode into the soil, with the second electrode receiving the signal. The resistance of the soil between the electrodes determines the potential difference, indicating soil moisture. Lower resistance corresponds to wetter soil and vice versa. The study used soil relative humidity data for sensor calibration and accuracy evaluation.

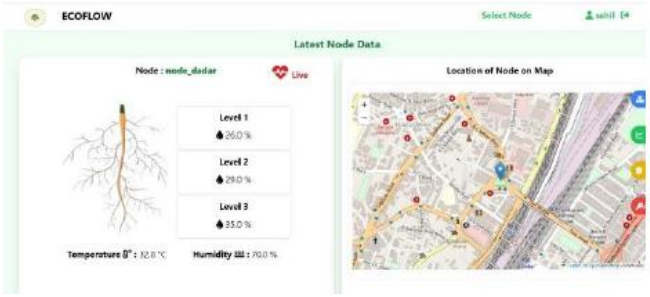
Additionally, a DHT-11 sensor was employed in the study. The DHT-11 is a digital temperature and humidity sensor featuring a single-wire communication interface. It measures temperature and humidity and communicates the data to the microcontroller. The DHT-11 enhances the comprehensive analysis of environmental conditions, contributing valuable insights to the study. The collected data were effectively displayed in real-time using a GUI, facilitating convenient monitoring of the measured soil moisture content, temperature, and humidity.

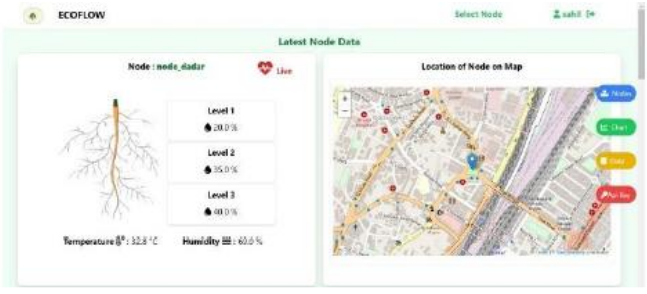
2.3 Experimental Proceedings

A systematic experimental approach was employed to harness data from soil moisture sensors and a temperature and humidity sensor. This essential dataset served as a primary input for the microcontroller, initiating a series of intricate processes within the system. The transmitted data found its destination in both a dedicated website and a relay module, crucial components governing the regulation of water flow in the irrigation system. Adopting a rigorous experimental design, the website was conceived as an interactive platform, allowing end-users to access real-time data through a Graphical User Interface (GUI). The data, rich with information on soil moisture and environmental conditions, was not merely presented in raw form but meticulously organized into charts and tables. This thoughtful presentation strategy aimed to provide nuanced insights for each individual node within the system and simultaneously offer a broader, cluster-level overview.

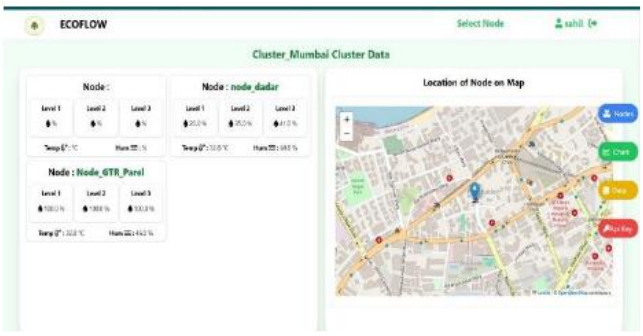
3. Results

Arduino IDE is the main development environment that we use to control the Ecoflow system. The microcontroller involved is programmed to display values of the soil moisture as well as the temperature so as to give insights to the crop health and build a more capable and accurate system.





In general, the values read by the sensors differ significantly depending on the soil texture. Therefore, in order to use this sensor properly, the user would need to have prior knowledge of the soil type in which the sensor will be installed. The results concerning the influence of soil texture on the readings performed by the sensors.



Clusters of nodes representing different places. Nodes are labeled with place names, and clusters are visually differentiated. This provides insight into the spatial distribution and grouping of places. And shows how moisture levels vary with depth over time at different places.



4. Discussion

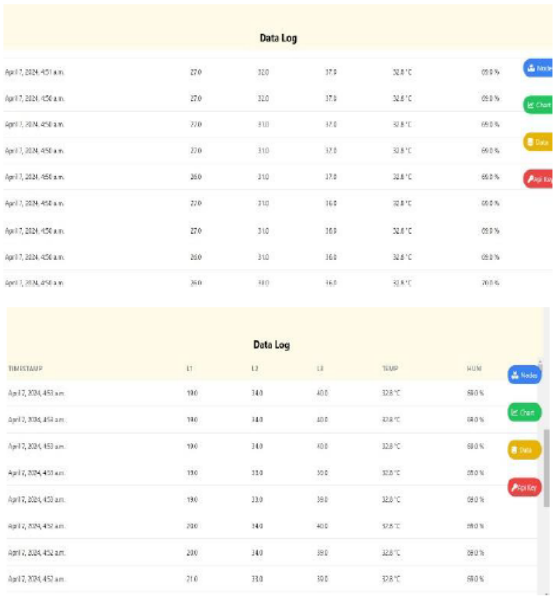
The output of sensors at each level of the crop provides valuable information that is instrumental in optimizing water use efficiency and crop management practices. In this discussion, we will delve into the significance of sensor data at different crop levels and its implications for enhancing agricultural productivity.



The picture depicts a chart correlating moisture levels with date/time, where each data point represents a specific time and moisture reading. Additionally, the colour gradient on the chart corresponds to different soil depths, providing insight into how moisture levels vary with depth over time.



SQL Lite database stores moisture level data correlated with date and time, where each entry includes the moisture level, soil depth, and corresponding date and time. The structure can be represented through an Entity-Relationship Diagram (ERD) or a tabular format showcasing the tables and their respective columns.



The output of sensors at each level of the crop plays a pivotal role in optimizing agricultural management practices, improving resource use efficiency, and enhancing crop productivity and resilience. By leveraging sensor data to inform decision-making processes, farmers can adopt precision agriculture approaches that maximize yields while minimizing inputs and environmental impacts, ultimately contributing to the long-term sustainability and profitability of farming operations.

5. Conclusions

In conclusion, the implementation of water-efficient farming practices holds immense promise for addressing contemporary agricultural challenges. Through the integration of innovative techniques such as drip irrigation, precision agriculture, and soil moisture conservation methods, significant strides can be made towards mitigating water scarcity while simultaneously enhancing agricultural productivity and sustainability. Our research has underscored the critical role of crop selection, soil health management, and technological innovations in optimizing water usage within farming systems. By adopting these practices, farmers can not only reduce water consumption but also improve soil health, increase crop resilience to drought, and minimize environmental impacts associated with conventional agricultural practices. However, it is crucial to acknowledge the existing barriers to widespread adoption, including financial constraints, limited access to technology, and socio-economic factors. Addressing these challenges will require concerted efforts from policymakers, agricultural stakeholders, and researchers to provide necessary support, incentives, and knowledge dissemination mechanisms. Looking ahead, the continued advancement of water-efficient farming practices remains imperative in the face of escalating water scarcity and climate variability. Future research endeavors should focus on refining existing techniques, developing cost-effective solutions, and fostering interdisciplinary collaborations to accelerate the transition towards more sustainable and resilient agricultural systems. In conclusion, by embracing water-efficient farming practices, we can not only safeguard our precious water resources but also cultivate a more resilient, equitable, and environmentally sustainable agricultural sector for future generations.

6. Acknowledgement

The authors take this opportunity to express our profound gratitude and deep regards to our guide Professor Dattatray Bade for his guidance, monitoring and constant encouragement throughout the course of this project work. We also take this opportunity to express a deep sense of gratitude to Prof Dr. Varsha Turkar HOD of E.X.T.C. Dept. for her cordial support, valuable information and guidance, which helped us in completing this task through various stages.

REFERENCES

[1] Les Levidow, Daniele Zaccaria, Rodrigo Maia, Eduardo Vivas, Mladen Todorovic, Alessandra Scardigno, Improving water-efficient irrigation: Prospects and difficulties of innovative practices, Agricultural Water Management, Volume 146, 2014, Pages 84-94.

[2] Nagahage, E.A.A.D.; Nagahage, I.S.P.; Fujino, T. Calibration and Validation of a Low-Cost Capacitive Moisture Sensor to Integrate the Automated Soil Moisture Monitoring System. Agriculture 2019, 9, 141.

[3] Santos, Fernando & Cardoso Madureira Tavares, Leticia & De Moura Araújo, Guilherme & de Lima Casseres dos Santos, Lucas & Nogueira, Caio & Bachini, Matheus & Teixeira, Marcos. (2021). confidence analysis and calibration of a fc-28 soil moisture sensor mounted on a microcontroller platform. Nativa. 9. 123-128. 10.31413/nativa.v9i1.9152.

[4] Sharma, K., Irmak, S., Kukal, M. S., Vuran, M. C., Jhala, A., & Qiao, X. (2020). Evaluation of soil moisture sensing technologies in silt loam and loamy sand soils: Assessment of performance, temperature sensitivity, and site- And sensor-specific calibration functions. Transactions of the ASABE, 64(4), 1123-1139.

[5] Md Reazul Islam, Khondokar Oliullah, Md Mohsin Kabir, Munzirul Alom, M.F. Mridha, Machine learning enabled IoT system for soil nutrients monitoring and crop recommendation, Journal of Agriculture and Food Research, Volume 14, 2023, 100880, ISSN 2666-1543.

[6] Development of air temperature and soil moisture monitoring systems with LoRA technology, Aji Purnomo, Fendi, Maulana Yoeseph, Nanang, Alim Tri Bawono, Sahirul, Hartono, Rudi, T2 - Journal of Physics: Conference Series, 2021, IOP Publishing.

[7] Khaled Obaideen, Bashria A.A. Yousef, Maryam Nooman AlMallahi, Yong Chai Tan, Montaser Mahmoud, Hadi Jaber, Mohamad Ramadan, An overview of smart irrigation systems using IoT, Energy Nexus, Volume 7, 2022, 100124, ISSN 2772-4271

[8] D. Hatanaka, A. Ahrary and D. Ludena, "Research on Soil Moisture Measurement Using Moisture Sensor," 2015 IIAI 4th International Congress on Advanced Applied Informatics, Okayama, Japan, 2015, pp. 663-668, do i: 10.1109/IIAI-AAI.2015.289.

[9] G. Patrizi, A. Bartolini, L. Ciani, V. Gallo, P. Sommella and M. Carratù, "A Virtual Soil Moisture Sensor for Smart Farming Using Deep Learning," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-11, 2022, Art no. 2515411, doi: 10.1109/TIM.2022.3196446.

[10] K. Bindra and A. Mishra, "A detailed study of clustering algorithms," 2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2017, pp. 371-376, do i: 10.1109/ICRITO.2017.8342454.

❑ ❑ ❑ ❑ ❑ ❑

AET 039

Verification & Tracking of Humans in Campus

Atul Oak
Professor (Project Guide), Department of Electronics and Telecommunication Engineering, Vidyalankar Institute of Technology.

Alok Tripathi
Department of Electronics and Telecommunication Engineering, Vidyalankar Institute of Technology.

Shivank Saxena
Department of Electronics and Telecommunication Engineering, Vidyalankar Institute of Technology.

Krishnakumar Yadav
Department of Electronics and Telecommunication Engineering, Vidyalankar Institute of Technology.

Ayush Zode
Department of Electronics and Telecommunication Engineering, Vidyalankar Institute of Technology.

***Abstract* – The aim of the proposed system is to create a reliable and affordable system that streamlines human verification, real-time tracking, and attendance management within the educational and corporate campuses. The purpose of this proposed system is to provide a robust and practical solution for small-scale campuses seeking to bolster their security infrastructure and attendance management processes. By enabling facial registration, real-time face verification, access timestamp recording, web-based access log queries, our system addresses the multifaceted needs of campus administrators. The goal is to simplify security management, automate attendance tracking, and offer a user-friendly interface for accessing essential information on campus access activities.**

***Index Terms* – Attendance System, Smart Attendance, Face Recognition, Tracking of humans.**

INTRODUCTION

The project represents an innovative initiative that harnesses facial recognition technology's capabilities to bolster security and streamline attendance management within campus environments. Tailored for implementation in diverse settings, including educational institutions and corporate campuses, the system integrates essential components such as a Raspberry Pi, a camera, and advanced deep learning algorithms. This collaborative configuration forms a robust and efficient facial recognition system, with the Raspberry Pi serving as the central processing unit responsible for executing deep learning algorithms and managing overall system operations.

The camera captures images of individuals seeking access to the campus, and these images undergo processing by deep learning algorithms for identity verification. Key features of the system encompass facial registration, real-time face verification, and the recording of access timestamps. Furthermore, the system offers a web-based interface for accessing and querying access logs.

PROBLEM STATEMENT

In today's dynamic and security-conscious environments, traditional attendance management and access control systems often prove inadequate in ensuring both efficiency and security. Educational institutions and corporate campuses face the challenge of maintaining secure premises while efficiently managing attendance records. To address these issues, the project aims to develop an innovative facial recognition system tailored for campus environments, leveraging cutting-edge technology to enhance security and streamline attendance management processes.

The current absence of a comprehensive, integrated solution for attendance management and access control within campus environments poses several challenges:

1. **Security Vulnerabilities:** Conventional methods of access control, such as swipe cards or PIN codes, are susceptible to security breaches through theft, loss, or unauthorized sharing. These vulnerabilities compromise the integrity of campus security protocols.
2. **Inefficient Attendance Tracking:** Manual attendance recording systems are time-consuming, prone to errors, and lack real-time monitoring capabilities. This inefficiency leads to inaccuracies in attendance data and hampers timely intervention in the event of security breaches or emergencies.
3. **Limited Scalability:** Many existing attendance management systems lack scalability, making them unsuitable for accommodating the diverse needs of large educational institutions or corporate campuses. As the size of the campus grows, the inadequacy of traditional systems becomes increasingly apparent.
4. **Lack of Integration:** The absence of a unified system for attendance management and access control results in disjointed operations, requiring separate solutions for attendance tracking and security management. This lack of integration complicates system administration and increases the likelihood of errors.

The proposed facial recognition system seeks to address these challenges by offering a comprehensive solution that integrates seamlessly into campus environments. However, several key issues must be addressed to ensure the success and effectiveness of the system:

1. **Accuracy and Reliability:** The system must demonstrate high accuracy in facial recognition and identity verification to minimize false positives and negatives. Robust testing and refinement of deep learning algorithms are essential to achieve reliable performance under varying lighting conditions and facial expressions.
2. **Privacy and Data Security:** Facial recognition technology raises concerns regarding privacy and data security, particularly regarding the collection and storage of biometric data. The system must adhere to stringent privacy regulations and employ robust encryption and access controls to safeguard sensitive information.

3. **User Acceptance and Accessibility:** The system's user interface should be intuitive and user-friendly, catering to the diverse needs and technological competencies of campus stakeholders, including students, faculty, and administrative staff. Training and support resources must be provided to ensure widespread adoption and acceptance of the system.
4. **Scalability and Compatibility:** The system should be scalable to accommodate the evolving needs and growth of campus environments, supporting seamless integration with existing infrastructure and future expansion initiatives. Compatibility with diverse hardware and software platforms is crucial to ensure interoperability and ease of deployment.

By addressing these challenges, the proposed facial recognition system has the potential to revolutionize attendance management and access control within campus environments, enhancing security, efficiency, and convenience for all stakeholders.

LITERATURE SURVEY

In developing a robust facial recognition system for our project, we conducted an exhaustive examination of various algorithms and methodologies, with the goal of identifying the most suitable approach. Since the project is split into two parts- Face Enrolment and Face Recognition, we needed two algorithms, one for feature extraction and the other for face verification.

We first started our search for the feature extraction algorithm and discovered the concept of Eigen faces. Eigen faces refer to an appearance-based approach to face recognition that seeks to capture the variation in a collection of face images and utilize this data to encode and analyze images of individual faces in a holistic manner. By this method we can take automatic attendance. Feature Extraction and Detection is done by projecting new image onto A low dimensional linear “face space” defined by the Eigen faces. Eigen faces is probably one of the simplest face recognition methods and rather old. We were introduced to this new concept by “Automated Attendance System through Eigen Faces Using Image Processing” by M.Sree Vidya, K.Arul (2016).

Along with its merits the Eigen Faces method for image processing also has several limitations. One of the main limitations is its sensitivity to variations in lighting. Different lighting conditions can cause image saturation and significantly affect the accuracy of recognition. Another limitation is related to pose differences. Changes in the orientation of the head can distort the 2D feature distances, which can lead to inaccuracies in the recognition process.

Lastly, variations in facial expressions can alter the location and shape of features, impacting the recognition process. These limitations highlight the challenges in using the Eigen Faces method for image processing and suggest the need for more robust methods or additional preprocessing steps to handle these variations.

The second algorithm we referred was LMFD: Lightweight Multi-Feature Descriptors for Image Stitching from “LMFD: Lightweight Multi-Feature Descriptors for Image Stitching” is authored by Yingbo Fan, Shanjun Mao, Mei Li, Jitong Kang, and Ben Li (2023). The paper discusses the challenges in image stitching, particularly the limitations of existing feature descriptors such as poor robustness to noise and rotational transformations. To overcome these challenges, the paper introduces Lightweight Multi-Feature Descriptors (LMFD), which enhance image stitching accuracy and efficiency by using binary matrices for computational simplicity. LMFD employs gradients, means, and global information around feature points to create robust descriptors with rotational invariance and noise resistance. The effectiveness of LMFD is demonstrated through experiments on Hpatches and 2D-HeLa datasets, showing superior performance in image matching accuracy.

One of the main limitations of LMFD is the redundant calculations involved in the construction of multiple feature descriptors. This process involves repetitive computations for points surrounding the same feature point, which can lead to inefficiencies. Another limitation is related to parameter optimization. The parameters used in the algorithm, such as the radius of feature point detection and the size of the convolution kernel for descriptor calculation, could potentially be further optimized. This optimization could possibly be achieved through neural network training. While not explicitly stated as a limitation in the paper, the optimization of computation time is mentioned as an area for future work.

Due to above demerits these Algorithms were not suitable for our project. Instead, we decided to use another algorithm which can prevent the above issues while delivering the required performance - Histogram Oriented Gradients (HOG).

We learnt about this Algorithm from the research paper titled – “Histograms of Oriented Gradients for Human Detection” by Navneet Dalal and Bill Triggs (2005). The paper introduces the Histogram of Oriented Gradients (HOG) descriptor as a feature representation for human detection. It describes how HOG captures the local gradient orientation information in an image, providing a concise yet discriminative representation of object appearance. Feature Extraction: We saw detailed the process of extracting HOG features from image patches. They explain how gradient magnitude and orientation are computed and how histograms of gradient orientations are constructed within local cells. Block Normalization: The paper discusses the importance of block normalization in HOG, emphasizing its role in enhancing the descriptor's robustness to lighting variations and contrast changes. Different normalization methods, including L1 and L2 normalization, are explored and evaluated. Comparison with Existing Approaches: The paper compares HOG-based human detection with other state-of-the-art methods at the time, demonstrating its superior performance in terms of accuracy and computational efficiency. Parameter Tuning and Optimization: we came across the impact of various parameters, such as cell size, block size, and the number of histogram bins, on the performance of HOG based human detection. They provide insights into parameter selection and optimization strategies based on experimental results. Real-World Applications: The authors discussed potential real-world applications of HOG-based human detection, such as pedestrian detection in surveillance systems and driver assistance systems. They emphasize the scalability and adaptability of the approach to different domains and scenarios.

After referring the paper, we found that HOG would be most appropriate for our project. We then moved onto finding the algorithm for Facial Recognition and with this we have also came across following algorithms: Haar Cascade Classifiers, Local Binary Patterns (LBP), and Convolution Neural Networks (CNN). Haar Cascade is a machine learning object detection method widely employed for recognizing objects in images or videos. It operates by analyzing patterns of pixel intensity, known as Haar-like features, which are used to train a classifier distinguishing between objects and the background. Notably fast and accurate, Haar Cascade is commonly utilized in face detection, as introduced in "Rapid Object Detection using a Boosted Cascade of Simple Features" by Viola and Jones (2001).

LBPH (Local Binary Pattern Histograms) is a facial recognition algorithm that extracts texture information by comparing the intensity of each pixel with its neighbors. Creating binary patterns and generating histograms for local regions, LBPH is valued for its simplicity, efficiency, and robustness to lighting variations. Its efficacy in texture-based face recognition was demonstrated in "Face Recognition with Local Binary Patterns" by Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen.

While Haar Cascade Classifiers excel in real-time detection, challenges may arise in handling variations in pose and lighting. They necessitate careful parameter tuning and may not adapt well to dynamic environments. LBP, robust to texture variations, may fall short in addressing more complex features and scenarios.

To find out more about this algorithm we referred to “FaceNet: A Unified Embedding for Face Recognition and Clustering” by Schroff et al.'s (2015). Renowned for its ability to learn complex facial features with minimal data, CNN excels in handling variations in pose, lighting, and expression. Its one-shot learning capability is particularly valuable where extensive training data collection is impractical. Moreover, its adaptability and continuous improvement with more data make it highly accurate over time. We selected CNN as the cornerstone for our facial recognition system due to its state-of the-art performance, real-world robustness, and adaptability. The deep learning approach aligns seamlessly with the intricate requirements of campus security and attendance management.

Table 3.1 Differences Between Haar Cascade Classifiers, LBP, and FaceNet:

Methods	Dimens-ionality Reduce	Handling Variations	Accuracy and Robustness	Training Data Requirements
Haar Cascade Classifiers	No	Limited	Moderate	Moderate
Local Binary Patterns (LBP)	No	Robust	Moderate	Moderate
FaceNet (Deep Learning)	Yes	Robust	High	Low (One-shot learning)

METHODOLOGY

The system operates in two modes: enrollment for registering individuals and recognition for verifying and marking attendance. The process involves capturing images, detecting, and preprocessing faces, and using a training set for recognition.

Enrollment Mode:

Image Capture: The system captures an image of the person’s face.

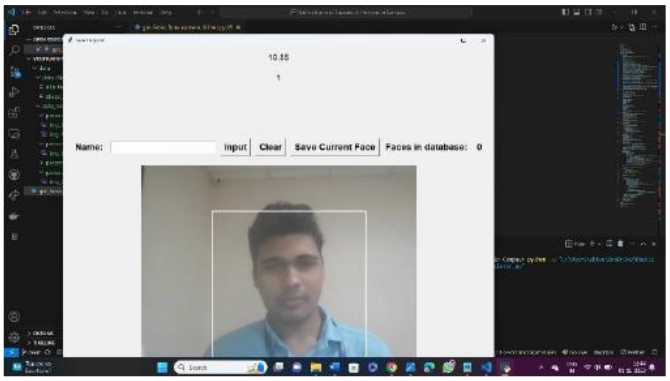


Figure 4.1: Image Face Detection

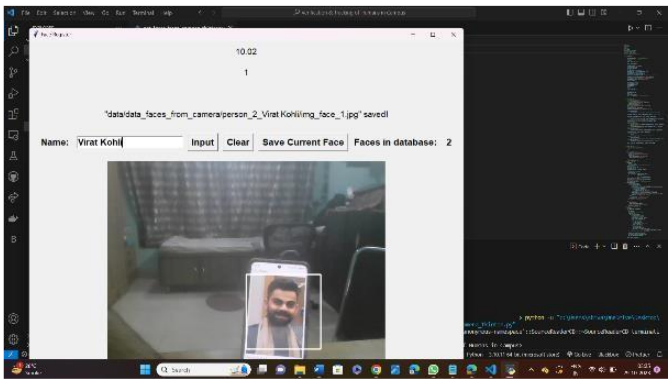


Figure 4.2: Details & Information Filling

Image Preprocessing: The system preprocesses the image to improve the quality for recognition.

Training Set: The pre-processed image is added to a training set of images for future recognition.

Recognition Mode:

Image Capture: The system captures an image of the person whose face is to be recognized.

Detected Face: The system detects the presence and location of a face in the captured image.

Preprocessing: The system preprocesses the captured image as in the enrollment mode.

Recognition: The pre-processed image is compared to the faces in the training set to recognize the person.

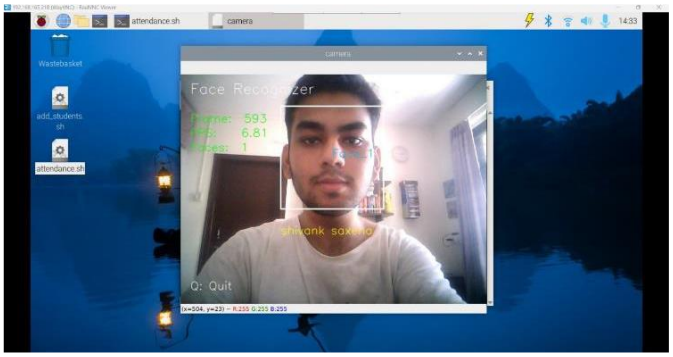


Figure 4.3: Face Recognition by matching the features present in the database

Attendance Marking: The system marks the person’s attendance based on successful recognition.



Figure 4.4: Attendance tracking window

IMPLEMENTATION

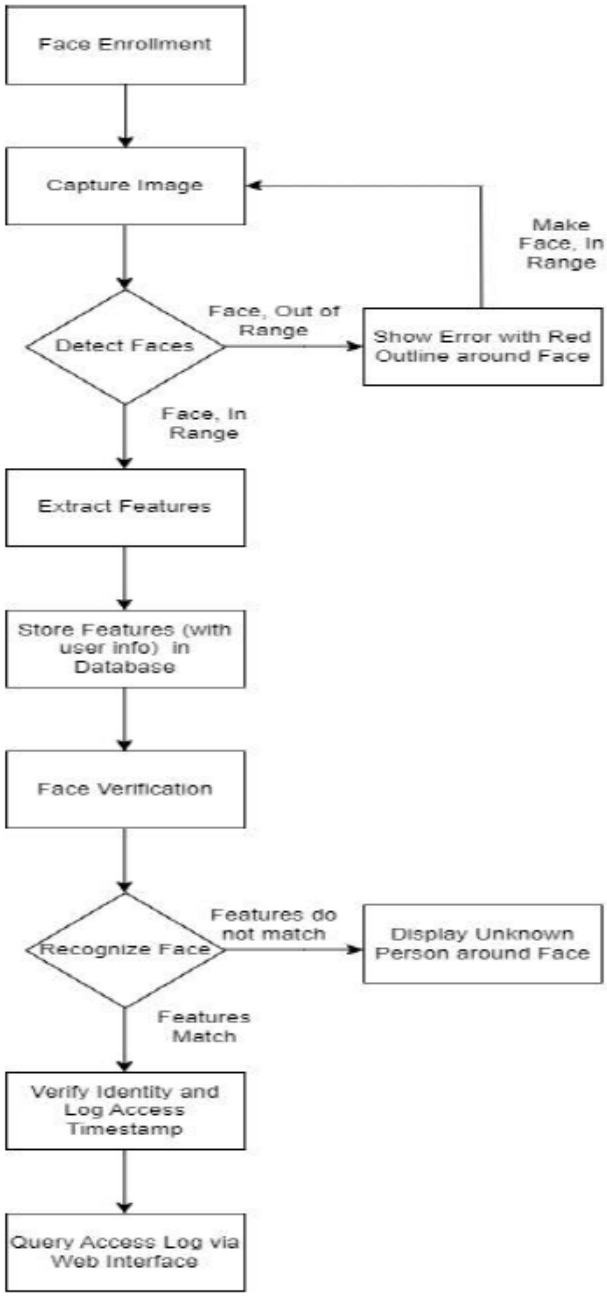


Figure 5.1: Flowchart

System Architecture

The system integrates a Raspberry Pi, camera, and deep learning algorithms to create a facial recognition system for security and attendance management in campus environments.

Feature Extraction Algorithm

Utilizes Dlib's Histogram of Oriented Gradients (HOG)-based face detection algorithm for identifying faces in images and video streams.

Face Verification Algorithm

Employs a deep learning model, FaceNet, to generate face embeddings and perform face recognition with high accuracy and efficiency.

Attendance Logging

A database is maintained to store attendance logs, recording the date, time, and identity of individuals, which can be accessed and queried through a web-based interface.

Testing and Validation

The system is rigorously tested in various conditions to ensure accuracy, reliability, and performance, with results indicating successful implementation of the initial phase.

RESULT

Testing the device in different conditions was a key aspect of the project. The test sites are chosen to represent realistic scenarios where the system can be deployed and tested. The campus environment offers a variety of challenges and opportunities for facial recognition, such as varying lighting conditions, occlusions, poses, expressions, and identities. The surrounding areas provide additional diversity and complexity for testing the system's robustness and adaptability. The test sites include the Vidyalankar Institute of Technology campus and the surrounding areas and in an artificially lit room. The datasets encompass images compiled from diverse origins, such as group members in real time and pictures of people shown to the camera from the phone.

This first testing was done using the laptop's built-in camera and we observed that because the camera having poor quality the faces were getting detected only in optimal lighting conditions and at specific angles. Due to the absence of mobility, it was also very inconvenient for the user to participate in the process. Hence to add mobility and portability and increase video capture quality it was decided to use the Raspberry Pi and an external camera. After implementing the system using the raspberry pi and the external camera, we tested whether the device could recognize multiple people at once or not. So, we first tested it with only one person in front of the camera which the system was able to recognize and then we tested it with multiple people in front of the camera and since we are using CNN for facial recognition, the device was able to recognize more than one person simultaneously and mark both of their attendance simultaneously along with time.

The system continuously captures frames from the camera. For each captured frame it performs face detection to locate faces within the image. It extracts facial features from the detected faces and compares the extracted features with the features stored in the database. Simultaneously, it determines the identity of the recognized individual based on similarity of features is determined. If a match is found, the attendance for the recognized individual is marked. This process for each frame captured by the camera is repeated.

A separate table in the database to store attendance logs is created, whenever attendance is marked for an individual. It records the date, time, and identity of the individual in the attendance database. The database schema ensures the attendance logs are indexed and organized for efficient retrieval and analysis. Essentially an SQLite Database is created from here a Flask application provides two routes - one for rendering the initial page and another for handling form submissions to retrieve and display attendance data for a selected date from an SQLite database. Attendance statistics are displayed on a web page.

CONCLUSION

The project's inception was marked by an exhaustive exploration of existing literature, unravelling a diverse array of face detection and recognition methodologies. This comprehensive survey, spanning scholarly articles, research papers, and technical publications, fortified the theoretical foundations and laid the groundwork for subsequent stages. The implementation of a smart attendance system offers a modern and efficient solution to traditional attendance tracking methods. By using facial recognition technology, this system streamlines the attendance marking process, reducing administrative burdens, and minimizing errors. Through the integration of face registration, feature extraction, and database management, the system ensures accurate attendance records while providing a seamless user experience. The selected face registration and recognition method paved the road for project implementation. Meticulously curated components were chosen for compatibility, considering factors such as processing prowess, scalability, mobility, and integration capabilities. The project's unwavering goal is to deliver a cutting-edge, efficient, and intelligent monitoring and tracking system. This ambition reflects a dedication to technological advancement, addressing contemporary needs and paving the way for future innovations in surveillance and tracking. The fusion of avant-garde methods with selected components positions the project as a notable contribution to security, surveillance, access control, and beyond. With its ability to accurately identify individuals and log attendance data in the present time, the system not only improves efficiency of operation but also paves the way for potential developments in attendance tracking and management. As such, its implementation represents a forward-thinking approach to addressing the evolving needs of modern-day institutions, ensuring they remain equipped with the tools necessary to thrive in an increasingly digital and data-driven world.

FUTURE SCOPE

The future scope of the project can be vast and diverse, as this technology continues to evolve and find applications in various fields. The potential avenues for future development:

Enhanced Accuracy and Speed: Continuous improvement in algorithms and hardware can lead to quicker and more accurate face detection systems. This could involve the development of more effective deep learning forms or optimization techniques to streamline processing.

- Robustness to Variations: Future face detection systems may become more robust to variations in lighting conditions, facial expressions, occlusions, and other environmental factors. This could involve training algorithms on more diverse datasets or incorporating techniques for data augmentation and normalization.
- Privacy-Preserving Solutions: As concerns over privacy and data security grow, there is a must for face recognition devices that can operate while preserving the anonymity of individuals. Future research may focus on developing techniques such as federated learning or differential privacy to address these concerns.
- Person Tracking for Proctored Environments: Face detection systems enhanced with person tracking can enable more sophisticated monitoring and surveillance in proctored areas such as college, offices, hospitals etc. By continuously tracking individuals' movements within the designated area, these devices can make sure that people remain within the prescribed boundaries and detect any suspicious behaviour or unauthorized activities.

REFERENCES

[1] M. S. Vidya and K. Arul, "Automated Attendance System through Eigen Faces Using Image Processing," IEEE Transactions on Image Processing, vol. 25, no. 9, pp. 4226-4235, Sep. 2016.

[2] Y. Fan, S. Mao, M. Li, J. Kang, and B. Li, "LMFD: Lightweight Multi-Feature Descriptors for Image Stitching," IEEE Transactions on Image Processing, vol. 32, no. 6, pp. 1200-1210, Jun. 2023.

[3] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 2005, pp. 886-893.

[4] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 2001, pp. I-511-I-518.

[5] T. Ahonen, A. Hadid, and M. Pietikäinen, "Local Binary Patterns for Face Recognition," in 2004 8th European Conference on Computer Vision, 2006, vol. 2, pp. 469-481.

[6] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 815-823

[7] Dhairya Kumar, "Introduction to FaceNet: A Unified Embedding for Face Recognition and Clustering," in Analytics Vidhya, https://medium.com/analytics-vidhya/introduction-tofacenet-a-unified-embedding-for-face-recognition-and-clustering-dbdac8e6f02

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional

Neural Networks," in Advances in Neural Information Processing Systems 25, 2012, pp. 1097-1105.

[9] Panagiotis Antoniadis, "Introduction to Triplet Loss," in Baeldung, https://www.baeldung.com/cs/triplet-loss

[10]Raspberry Pi Foundation, "Raspberry Pi," https://www.raspberrypi.org/

[11] github.com/davisking/dlib-models

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "YOLO: Real-Time Object Detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788

[14] OpenCV Library: https://opencv.org/

[15] Szeliski, Richard. Computer Vision: Algorithms and Applications. Springer, 2011.

[16] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT press, 2016.

[17] Bao, Marian, et al. Face Recognition by Convolutional Neural Networks. Springer, 2018.

[18] Li, S. Z., & Jain, A. K. (Eds.). (2011). Handbook of face recognition. Springer.

[19] Brownlee, J. (2019). Deep learning for computer vision. Machine Learning Mastery.

[20] https://github.com/Arijit1080/Face-Recognition-Based-Attendance-System

❑ ❑ ❑ ❑ ❑ ❑

AET 040

Automatic Medicine Dispenser using QR code

Chinmay Prashant Tawde
Vidyalankar institute of Technology, Mumbai

Prof. Beena Ballal
Vidyalankar institute of Technology, Mumbai

Makarand Uday Nagvekar
Vidyalankar institute of Technology, Mumbai

Mohammed Ahtesham Shaikh
Vidyalankar institute of Technology, Mumbai

Vaishnavi Phatkare
Vidyalankar institute of Technology, Mumbai

***Abstract* – This project introduces a QR code-enabled Medicine Vending Machine, powered by an Arduino Mega 2560. Utilizing components such as a DC Gear Motor, IR sensor, and Bluetooth Module, the system offers a user-friendly interface for dispensing medications. Users scan QR codes for precise and automated dispensing. The report details the design and implementation, emphasizing the system's potential to enhance medication accessibility and efficiency in healthcare.**

***General Terms* – Medication Dispensing System: A user interface to automate the dispensing of medications, enhancing accuracy and efficiency.**

***Keywords* – Arduino Mega, IOT, Vending machines, QR code, Smart Health Technology, Wireless.**

INTRODUCTION

In response to the changing healthcare and technology landscape, our project developed a vending machine using QR codetechnology. The system is powered by Arduino Mega 2560with important components such as DC gear motor, L298N motor driver, infrared sensor, HC05 Bluetooth moduleand 12V power supply for drug delivery.The design uses a spring and a motormounted wheel to ensure accurate and reliable dispensing. Integration of infrared sensors increases user safety by detecting problems during delivery.To trigger the automatic dispensing process, users only need to scan the QR code associated with the drug. This report describes the design, implementation, and potential use of our new system to improve medication utilization and clinical efficiency. The goal of the project is to eliminate the long and stressful queues in Hospitals and Pharmacies.

LITERATURE SURVEY:

The concept of an Automatic Medicine Dispenser Using QR Code represents a convergence of healthcare and technology, aimed at streamlining medication distribution and enhancing patient care. The project's core ideas find resonance in existing literature, revealing a growing interest in similar innovations. [1] introduced a prototype medication dispenser. This dispenser involves caregivers inputting medications via QR Code scanning and patients triggering medicine dispensing through a touchscreen GUI. Importantly, caregivers receive SMS notifications if patients miss their medication doses.[3] presented a scalable and remotely manageable smart medication dispenser. It incorporates a real-time clock to monitor patients', and medications are dispensed when patients press a button at the prescribed time. The system uses infrared sensors to track medication quantities and employs LAN for periodic medicine status updates without patient intervention.[4] discussed a prototype of a smart medical dispenser that monitors patient medication dispensed or not and sends alerts through a line if patients fail to take their medications within the specified time frame.[5] detailed the design and implementation of an autonomous Arduino-controlled device, enabling dispenser to configure medication vending to patients.[6] designed a low-cost automated medicine dispenser for pharmacists, capable of counting and dispensing medicines into vials. Infrared sensors were used for accurate counting of dispensed medicines with a 90% accuracy rate.

Title and Authors

The title of the project is Automatic medicine dispenser using QR code. The authors affiliated with this project are Chinmay Prashant Tawde from Vidyalankar Institute of Technology, Prof Beena Ballal from Vidyalankar Institute of Technology, Makarand Uday Nagvekar from Vidyalankar Institute of Technology, Mohammed Ahtesham Shaikh from Vidyalankar Institute of Technology and Vaishnavi Phatkare from Vidyalankar Institute of Technology.

Methodology

Step 1: Prescription Generation

1.1. Doctor Generates Prescription: The process begins when a doctor generates a prescription for a patient using specialized software. This prescription includes essential patient information, medical instructions, and details of the required medications.

Step 2: Prescription QR Code

2.1. QR Code Generation: The software generates a unique QR code based on the information within the prescription. This QR code acts as a digital representation of the prescription.

Step 3: Scanning the Prescription

3.1. Patient Approaches the Medicine Vending Machine The patient approaches the Automatic Medicine Vending Machine (MVM), which is equipped with a scanner and a user interface.

3.2. QR Code Scanning:

The patient scans the QR code using the Medicine Vending Machine 's scanning interface.

Step 4: Data Processing

4.1. QR Code Interpretation: The Medicine Vending Machine 's software processes the scanned QR code, extracting all the data encoded within it.

4.2. Medication Identification: The software identifies the medications prescribed, their quantities, and any specific instructions.

Step 5: Medicine Selection

5.1. User Interaction: The Medicine Vending Machine 's user interface prompts the patient to confirm the selection. The patient verifies the list of medications to be dispensed.

5.2. Confirmation and Dispensing: Once the patient confirms, the Medicine Vending Machine software activates the dispensing process.

Step 6: Dispensing Medications

6.1. The Medicine Vending Machine's hardware components, including motors and sensors, are engaged.

6.2. Medication Rotation: The motors rotate a spiral mechanism that holds the medications. The rotation of the spiral dispenses the required medications into a collection area.

Step 7: Medication Collection

7.1. Collection Point: The dispensed medications are collected by the patient from a designated collection point within the Medicine Vending Machine.

Step 8: Completion

8.1. Notification: The Medicine Vending Machine's software sends a completion notification to the patient and records the transaction.

Step 9: Inventory Management

9.1. Inventory Tracking: The system updates the inventory of available medications to maintain stock levels.

FIGURE

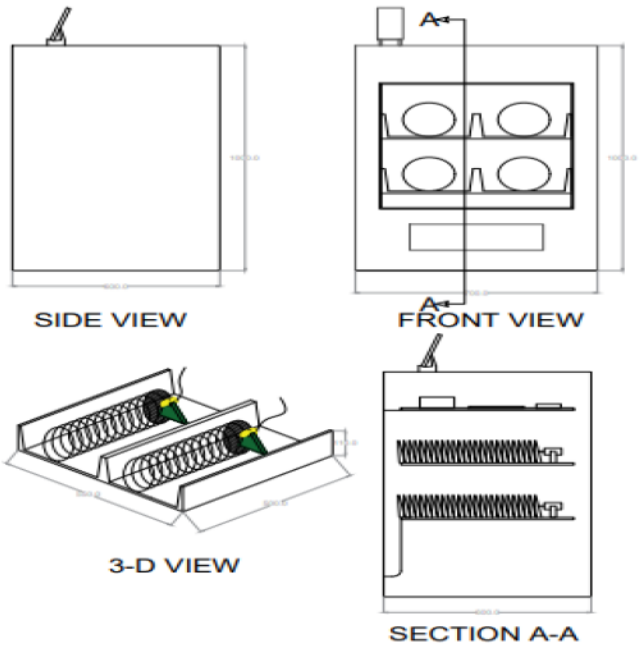


Fig 1: Layout of vending machine with measurements

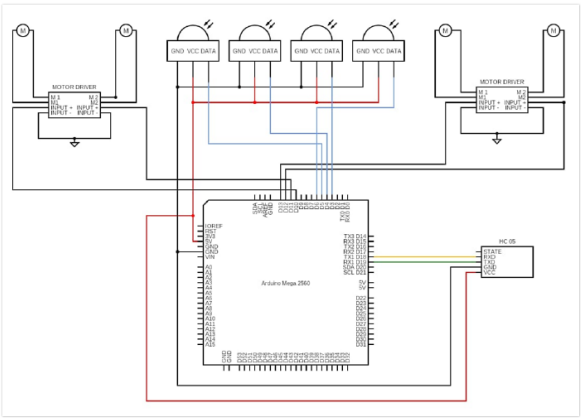


Fig 2: Circuit diagram

WORKING

The project's workflow begins when a user scans the QR code associated with a specific drug on the vending machine. The QR code triggers the Arduino Mega 2560, which acts as the system's brain. The Arduino processes the information and activates the DC gear motor, controlled by the L298N motor driver, to initiate the dispensing mechanism.

As the motor starts, the integrated infrared sensors play a crucial role in ensuring the safe and reliable delivery of the medication. These sensors constantly monitor the dispensing process, detecting any anomalies or issues, thereby enhancing user safety. If a problem is detected, the system can halt the process and alert users or administrators. The dispensing mechanism relies on a spring and a motor-mounted wheel, providing accuracy and reliability in delivering the prescribed medication. The HC05 Bluetooth module facilitates communication and connectivity, allowing for seamless interaction between the user's scanning action and the dispensing process. The entire system is powered by a 12V power supply, ensuring sufficient energy to drive the components effectively. The first step in designing a website interface is creating a digital prototype, often done using Figma. Figma enables the translation of concepts into interactive prototypes, allowing the structuring of screens, adding design elements, and simulating user interactions. For the front end, essential technologies include HTML for structure, CSS for presentation and layout, and JavaScript for interactivity. This innovative solution aims to streamline medication dispensing, reduce queues and enhancing clinical efficiency for a user-friendly healthcare experience.

CONCLUSION

In conclusion, our QR code-enabled medication dispensing system, powered by Arduino Mega 2560, represents a significant leap forward in healthcare technology. By seamlessly integrating components such as DC gear motor, L298N motor driver, infrared sensors, and HC05 Bluetooth module. It is an efficient and user-friendly solution to alleviate the challenges of long queues in hospitals and pharmacies. The design's reliance on a spring and motor-mounted wheel ensures precise and reliable medication delivery, while the integration of infrared sensors enhances user safety by promptly detecting and addressing any issues during the dispensing process.

REFERENCES

[1] Anusha M, Chandana J, Harshitha A U, Shruthi T V, International Journal for Research in Applied Science & Engineering Technology (IJRASET)

[2] A Brolin1, R Mithun, V Gokulnath and M Harivishanth, IOP Conference Series: Materials Science and Engineering, Volume 402, 2nd International conference on Advances in Mechanical Engineering (ICAME 2018) 22–24 March 2018, Kattankulathur, India

[3] Karthik B R,Rakshitha P,Ritesh M,Vinutha A, Dr. Kavitha K SDepartment of Computer Science and Engineering, Professor, Global Academy of Technology, Bengaluru, Karnataka, India

[4] Ijariit Journal, Sivasubramaniyan S, International Journal Of Advance Research, Ideas And Innovations In Technology

[5] Kulmukhanova, N., Daribay, A., Temirtayev, I., & Bassembek, U. (2018). ZhardEM Medicine Vending Machine. 2018 International Conference on Computing and Network Communications

[6] Desai, P., Pattnaik, B., Dey, S., Aditya, T., Rajaraman, K., & Aarthy, M. (2019). All Time Medicine and Health Device. 2019 5th International Conference on Advanced Computing & Communication Systems.

❑ ❑ ❑ ❑ ❑ ❑

AET 041

Identification of Chronic Wound & Prediction of Wound Size & Depth using Deep Learning Approach.

Guide: Dr. Sheetal Mhapare
Tanaya Kotapkar, Muskan Shaikh, Sakshi Kothawale, Yash Bhole
Department Of Electronics and Telecommunication
Vidyalankar Institute of Technology, Wadala
Mumbai University.

***Abstract*–Chronic wounds are injuries or sores that unable to heal within the expected timeframe, for wound recovery around three months. These types of wounds can be caused by factors, such as underlying health conditions, like diabetes, poor blood circulation or issues affecting the system. Chronic wounds have an impact, on a person's quality of life as they can deteriorate quickly requiring monitoring of the healing progress. In settings doctors traditionally use a cotton swab inserted into the wound to assess its depth, which can be uncomfortable for patients. Therefore, having access to wound depth information from images is crucial for doctors to gauge how well a wound is healing. Image based analysis provides a way to evaluate wounds by identifying features related to healing. However, the diversity in wound types and imaging conditions poses challenges for segmenting wound images. To address this problem, we are detecting and segmenting the wound using convolutional neural network(CNN) called U-NET that generates maps of wound segmentation with generalization capabilities. We have also created a basic algorithm for measuring the length and width of the impacted area. Lastly, we use Roboflow software to create a generic map of the wound, dividing it into three regions: shallow, medium, and deep. Studies indicate that our proposed approach can discern the layers of depth in a wound image, aiding in the examination of deep wounds and reducing patient suffering from invasive procedures.**

***General Terms*–Chronic Wound , Tissue Healing.**

***Keywords*– Chronic Wound, Convolution Neural Network, U-NET, Roboflow.**

INTRODUCTION

Chronic wounds have an impact, on a person's quality of life as they can deteriorate quickly requiring monitoring of the healing progress. Therefore, having access to wound depth information from images is crucial for doctors to gauge how well a wound is healing. Image based analysis provides a way to evaluate wounds by identifying features related to healing as a result identification of chronic wound and estimation of depth is important.

Chronic wounds represent a persistent and debilitating aspect of healthcare, often arising from underlying conditions such as diabetes, vascular diseases, or pressure sores [1][2]. Unlike acute wounds, which typically progress through the phases of healing within a reasonable timeframe, chronic wounds linger, posing significant challenges to patients and healthcare providers alike. The management and treatment of persistent wound revolves around wound care, a critical component aimed at facilitating healing, preventing infection, and improving overall patient well-being [2]. The determination of wound depth is crucial to appropriate wound care, which has traditionally been done using a combination of visual inspection and tactile assessment.[1].

In the realm of wound assessment, healthcare practitioners rely on a various ways to gauge the depth of a wound. Among these methods, Common methods include using a sterilized swab of cotton or probe. [1]. By gently probing the wound bed, doctors can assess the extent of tissue involvement and gain look into the severity of the wound. This traditional technique, although simple in nature, is painful and can lead to worsening of the wound [8]. As a result, depth determination holds paramount importance in the management of chronic wounds [1][2]. It serves as a crucial factor in determining the appropriate course of action for wound care. Understanding the depth of a wound allows healthcare providers to tailor treatment strategies to address the specific needs of the wound, such as selecting appropriate dressings, implementing debridement techniques, or considering surgical interventions [1][2]. Moreover, accurate depth assessment enables healthcare teams to monitor wound progression over time, facilitating timely adjustments to the treatment plan as needed.

The significance of depth determination extends beyond mere clinical considerations; it profoundly impacts patient outcomes and quality of life[3]. A thorough understanding of the wound's depth allows healthcare providers to intervene effectively, promoting optimal healing and mitigating the risk of complications such as infection or tissue necrosis. By addressing the underlying factors contributing to wound chronicity and implementing targeted therapies based on depth assessment, healthcare teams can significantly improve patient well-being and enhance overall healthcare outcomes.

The purposes of our research are to first detect chronic wounds and segment them from the background, i.e. normal skin, and then identify different layers of the wound, which are divided into three categories for convenience: deep, shallow, and medium regions, with the assistance of medical experts such as physicians and doctors. Finally, figure out the size of the wound by measuring its length and breadth. Wound depth in the wound photograph will be classified into three categories: shallow, medium, and deep. The layers were created using Roboflow program. We performed segmentation, which is the separating of wounds from normal skin. This segmentation is carried out utilizing a type of convolutional neural network known as UNET also We developed a simple Python program to estimate the length and breadth of a wound.

In conclusion, chronic wounds represent a complex and challenging aspect of healthcare, requiring meticulous attention and specialized interventions [1]. Wound care emerges as a pivotal component in managing chronic wounds, with depth determination playing a central role in guiding treatment decisions and optimizing patient outcomes [4].

LITERATURE REVIEW

Introduction to Chronic Wounds

Chronic wounds, like diabetic foot ulcers, venous ulcers, and pressure inflammation, are long-term injuries that fail to recover as predicted [1][2]. They frequently stem from underlying health issues including diabetes, poor circulation, or prolonged strain on the skin. Chronic wounds pose significant challenges in healthcare due to their prolonged healing times, susceptibility to infection, and potential for complications [1][2].

Adverse Effects of Chronic Wounds

Chronic wounds not only create physical discomfort but also have significant negative consequences on patients' quality of life.. They may give rise to pain, mobility limitations, social isolation, and psychological distress [1][2]. Moreover, chronic wounds increase the risk of secondary infections, osteomyelitis, and amputation, imposing a substantial burden on healthcare systems and society [1].

Traditional Methods for Wound Depth Assessment

Determining the depth of a wound is crucial for guiding treatment decisions and predicting healing outcomes [1][2]. Traditional methods for wound depth assessment include visual inspection, probing, and measurements using rulers or depth gauges. On the other hand, these procedures are generally subjective, error-prone, and lack precision, resulting in inconsistency in clinical practice and unsatisfactory patient care.[3].

Deep learning methodology to depth assessment

Recent advances in deep learning, notably neural networks using convolution (CNNs), have showed promise for increasing wound assessment accuracy. CNNs can automatically learn and identify characteristics in wound photos depending on depth [1][2]. In addition, Generative Adversarial Networks (GANs) were utilized to produce synthetic depth maps of wounds, improving the performance of depth detection techniques.[12].

CNN and U-Net for Wound Segmentation

CNNs were an form of deep learning model that has been shown to be extremely effective in image classification, object recognition, and segmentation tasks[11]. In the case of wound segmentation, a CNN can be developed to learn features from input wound photos and identify each pixel as either wound or background. The architecture typically consists of multiple convolutional layers followed by pooling layers to extract hierarchical features from the input images[1][11]. The final layers may involve upsampling and convolution operations to generate a segmentation mask at the output. CNNs can be trained end-to-end using annotated wound images, where the network learns to optimize a loss function that measures the pixel-wise difference between predicted and ground truth segmentation masks[6][11].

U-Net is a prominent CNN architecture that was created primarily for biomedical image segmentation applications, such as wound segmentation[11]. The U-Net architecture is composed of a contracting path (encoder) and an expansive path (decoder), with no links between appropriate layers in the encoder and decoder.[11] The contracting path captures contextual information from the input image through convolution and pooling operations, while the expansive path performs up sampling and convolution operations to generate a high-resolution segmentation mask. The skip connections allow the network to preserve spatial information and capture fine details during the up sampling process, making U-Net particularly effective for tasks where precise localization is crucial, such as wound segmentation [6].

In conclusion, chronic wounds represent a significant healthcare challenge due to their prolonged healing times and adverse effects on patients' well-being. Traditional methods for wound depth assessment often lack precision and reliability. Deep learning approaches, such as CNNs and U-Net, offer promising solutions for accurate depth detection and wound segmentation, potentially improving clinical decision-making and patient outcomes in wound care[11]. Continued studying and developing in this sector has the potential to transform chronic wound remediation and reduce the burden on individuals and medical facilities.

ARCHITECTURE

Segmentation

Algorithm used: - U-Net

The U-Net architecture presented in the provided code comprises a series of down-sampling blocks followed by up-sampling blocks, with a bottleneck layer in between. Each down-sampling block has two convolutional layers that was followed by max-pooling, whereas each up-sampling block combines bypass connections from the corresponding down-sampling block and then applies two convolutional layers. The final output is obtained through a convolutional layer with sigmoid activation, producing a segmentation map. This architecture is widely employed for semantic segmentation tasks, particularly in medical imaging and other domains necessitating pixel-level classification.

Length and width (breadth)

To calculate the length and width of the subdivided output image, we can use the contour bounding box information. The bounding box displays the coordinates of the top-left corner (x, y), as well as the width w and height h of the boundary rectangle that surrounds the segmented item.

Prediction of regions / Estimation of depth

The main aim is to classify the wound into three categories viz deep, medium, and shallow. We had trained the model with YOLOv8.

The dataset was divided in following manner –

Train images - 490 along with labels
Valid images – 48 along with labels

Test Images – 5

The model consists of 261 layers with a total of 3,264,201 parameters and 3,264,185 gradients. The model's computational complexity is measured at 12.1 GFLOPs (billions of floating-point operations per second). Convolutional layers are responsible for feature extraction, while the segmentation layer processes these features to generate pixel-wise segmentation predictions for the input image. The first layer is a convolutional layer (Conv) with 464 parameters that accepts three input channels and produces sixteen output channels, each with a kernel size of three and a stride of two. The AdamW optimizer is utilized for optimization.

RESULT OF SEGMENTATION

The primary aim is to segment the main wound from the background. For this purpose, we have used U-net segmentation model. The dataset was divided in following manner.

Train images - 810 along with labels
Valid images – 200 along with labels
Test Images – 200

The model consists of 28 layers with a total of 1962625 parameters. The model's computational accuracy at 98.09%. Convolutional layers extract features, and the segmentation layer uses these features to provide pixel-wise segmentation predictions on the input image. The first layer is a convolutional layer (Conv) with 448 parameters, input size of 256, output size of 256, and a total of 16 channels. The AdamW optimizer is utilized for optimization.

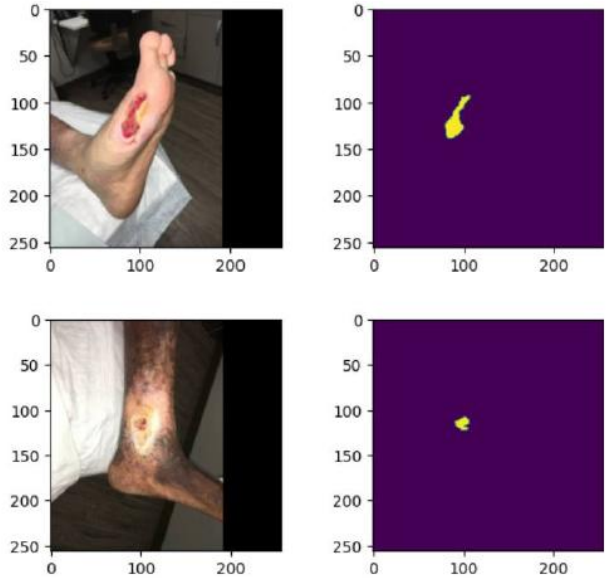


Fig 1 – Test Images used for segmentation.

Training Images



Fig 2 – Training Images used for segmentation.

RESULT OF PREDICTION OF REGION

The mean average precision (mAP) values are as follows:
For the background class (B):
mAP50(B): 0.46885
mAP50-95(B): 0.20725
For the foreground class (M):
mAP50(M): 0.21412
mAP50-95(M): 0.10933

Normalized confusion matrix

A normalized confusion matrix is a variant on the confusion matrix, a table used to assess the success of a classification method. In a confusion matrix, the rows represent the actual classes, while the columns represent the predicted classes. Each cell in the matrix represents the number of instances from the dataset that were classified into a particular combination of actual and predicted classes.

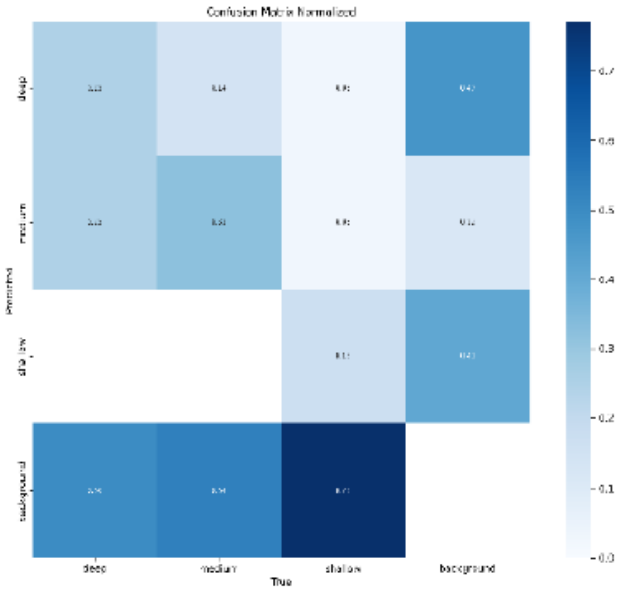


Fig 3– Normalized confusion matrix

Training batch result



Fig 4– Result of Training Batch

Graphs

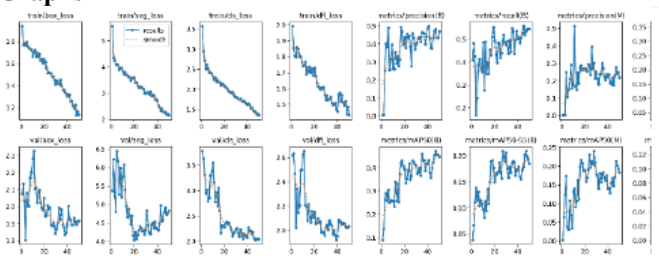
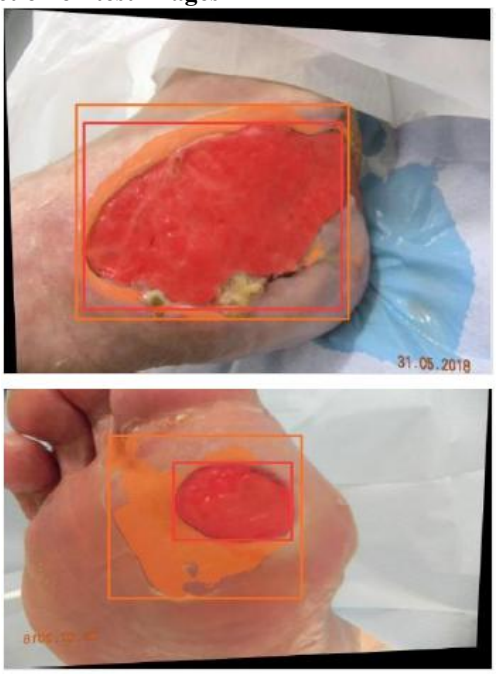


Fig 5– Result Graphs

Prediction on test images



FUTURE

Advanced methods such as Generative Adversarial Networks (GANs), CycleGANs, and other techniques for deep learning show great potential for improving wound depth diagnosis in healthcare [1][2]. GANs can create synthetic wound depth pictures to supplement datasets and address data shortage challenges. Conditional GANs allow for the generation of precise depth maps, increasing the accuracy of depth estimation models. CycleGANs provide domain adaptation by translating wound pictures between multiple imaging modalities or devices [1][2]. These algorithms provide solutions for data harmonization across several healthcare contexts, enabling model generalizability. Additionally, CNNs tailored for wound depth detection learn intricate features associated with varying depths from large-scale datasets. Integration of attention mechanisms enhances focus on relevant regions within wound images, improving depth estimation accuracy. Multi-task learning frameworks enable simultaneous The prediction of depth, coupled with other clinical factors, promotes thorough wound assessment. Ensemble methods for learning combine predictions from multiple models to improve robustness and dependability in real-world scenarios. Continued efforts to conduct research and development are required to maximize the possible benefits of these algorithms, advance wound care treatment, and improving patient outcomes.

CONCLUSION

The aims of our study are first to detect the chronic wound and segment the wound from the background i.e the normal skin. Second, identify the various layers of the wound, which have been separated into three categories for convenience: deep, shallow, and middle region, with the assistance of medical specialists such as physicians and doctors. Finally, compute the size of the wound. i.e. to measure its length and breadth. In this paper find that We tested our model on foot ulcer photos from the publicly accessible Medetec Wound Dataset to demonstrate its resilience. We used U-Net to do segmentation by alternating between down-sampling and up-sampling blocks. This implies 810 train images with labels, 200 valid images with labels, and 200 test images. In this article, we proposed a normalized confusion matrix, a modification of a confusion matrix, which is a table used to evaluate the effectiveness of a classification algorithm. We also provided an approach for calculating the length and breadth of an image using the bounding box method. Our research further explains how to use Yolov8 to divide the layers of wound pictures into shallow, deep, and medium.

REFRENCES

1) CWD2GAN: Generative Adversarial Network of Chronic Wound Depth Detection for Predicting Chronic Wound Depth Chiun-Li Chin*, Chieh-Yu Li, Yan-Ming Lai, Ting Chen,Tzu-Yu Sun and Jun-Cheng Lin http://dx.doi.org/10.1142/S0218001423510047

2) WG2AN: Synthetic wound image generation using generative adversarial network. The Journal of Engineering, 2021(5), 286-294. https://doi.org/10.1049/tje2.12033

3) The Medetec Wound Database (http://www.medetec.co.uk/¯les/medetec-image-databases.html) and Foot Ulcer Dataset,30 which are open-source datasets available online.

4) Chronic Wound Detection Using Machine Learning https://www.jetir.org/view?paper=JETIRCU06107

5) An Automated Wound System Empowered by Deep Leaning https://iopscience.iop.org/article/10.1088/1742-6596/2547/1/012005/meta.

6) Fully Automated Wound Segmentation with Deep Convolutional Neural Network https://www.nature.com/articles/s41598-020-78799-w

7) https://www.researchgate.net/publication/359177791_Automatic_wound_detection_and_size_estimation_using_deep_learning_algorithms Automatic Wound And Size Estimation Using Deep Learning.

8) https://ieeexplore.ieee.org/iel7/6287639/9668973/09843957.pdf A Survey of Wound Image Analysis Using Deep Learning

9) Yolo Based Deep Learning Model For Pressure Ulcer Detection And Segmentation https://www.mdpi.com/2227-9032/11/9/1222

10) https://www.researchgate.net/figure/Custom-designed-end-to-end-deep-learning-based-image-analysis-pipeline- The-pipeline_fig1_359177791 Custom Design End -To-End Deep Learning Based Image Analysis.

11) https://www.sciencedirect.com/science/article/pii/S235291 4822000375 Detect-And-Segment: A Deep Learning Approach To Automate Wound Image Segmentation.

12) https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/tje2.12033 WG2AN: Synthetic wound image generation using generative adversarial network.

❑❑❑❑❑❑

AET 042

Smart Cradle for Baby

Avani Shinde
Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
avani.shinde@vit.edu.in

Prof. Beena Ballal
Project Guide
Vidyalankar Institute of Technology
beena.ballal@vit.edu.in

Mrunal Jagtap
Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
mrunal.jagtap20@vit.edu.in

Sanika Karade
Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
Sanika.karade@vit.edu.in

Omkar Potare
Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
omkar.potare@vit.edu.in

Abstract **- In an era characterized by the demanding schedules of working parents, the task of nurturing an infant assumes an extraordinary level of complexity. The arduous juggling act between career obligations and childcare responsibilities can strain the very limits of human capacity. In such circumstances, where the relentless march of time leaves scant room for a full 24-hour devotion to the well-being of a precious newborn, the options have traditionally boiled down to entrusting the care of the child to either elderly family members or hiring a dedicated baby caregiver. However, the innate anxieties of parental love persist, and doubts linger like shadows in the minds of these devoted guardians. The persistent worries of parents for their children's safety and the comfort of their beloved offspring propel them to prompt frequent, often sometimes disruptive, visits during precious breaks in their workday. Breaks. A Smart Cradle System can resolve this issue in offer real-time. Resolution to this concern. A Smart Infant Care System consists of a microphone. Two sound sensors are used to identify the baby's crying activity. In this project, we will use one dataset i.e. voice of a baby crying. The lower-frequency crying voice will lead to the movement of the toy and on higher frequency music will play which also includes Mother's voice. The Temperature sensor notifies the parent about the environmental temperature near the baby and switches on the fan automatically. The speed of the mini fan remains constant as it is not adjustable to different speed settings. The Web Application also includes reminders for scheduled vaccinations and medicines in an application.**

Keywords**– Internet of Things (IoT), Wireless Communication, ESP 32**

1. INTRODUCTION

Ensuring your child's safety, comfort, and well-being is a top priority during the parenting journey. In this project an innovative "Smart Cradle for Baby," a significant invention that aims to give parents peace of mind and the best possible care and comfort for their infants. The typical infant cradle is redefined by this state-of-the-art technology, which includes numerous innovative features that make parenting simpler, safer, and more pleasurable. The Smart Cradle for Baby is more than just a crib—it's a companion, a guardian, and a caregiver all in one. It prioritizes the sensitive needs of newborns while meeting the changing needs of today's parents. The 'Smart Cradle for Baby' – an innovative solution in the domain of IoT and wireless communication, with a focus on Childcare Technology. In today's fast-paced world, many working parents face the daunting challenge of balancing their professional lives together with nurturing their infants. This delicate juggling act often results in increased stress and worry. Our project showcases a cutting-edge system designed to alleviate these concerns and provide parents with real-time support. The 'Smart Cradle for Baby' is a comprehensive solution that incorporates advanced technology to enhance infant care. Through the integration of components such as microphones, noise sensors, and thermal sensors, this system ensures the well-being of infants while giving parents peace of mind.

In recent years, many healthcare and health management systems have been designed to cater to the needs of adults and the elderly, offering various monitoring capabilities, automatic alerts, and additional features. However, these systems are not suitable for infants due to the need for cautious handling. Unlike adults, infants require a distinct approach to healthcare as they rely entirely on their parents. Infants cannot provide feedback on their health conditions, expressing discomfort solely through crying. Therefore, a specially designed healthcare system is essential for infant care, aiming to alleviate the burden on parents, particularly mothers. Supporting this notion, authors have developed a system based on a commercial GSM network. The microcontroller gathers filtered data from sensors such as the DHT11 for room temperature and a sound sensor to detect infant crying. Authors suggest mounting a camera on the cradle to provide live images of the infant to parents when they are away. Additionally, a speaker communication system allows infants to hear their parents' voices to calm them, along with a moving toy in the sleeping area to comfort them. A remote subsystem with a GSM module receives and forwards data to the microcontroller for further processing. [1] In a distinct research paper, the author devised a system harnessing the commercial GSM network. This system aims to sense essential parameters such as body temperature (measured via LM35), heart rate (detected through IR Transmitter and Receiver), respiratory rate (captured through a Piezo film sensor on the patient's chest), and blood pressure. These parameters are sensed, amplified with variable gain, filtered, and then transmitted to a microcontroller. A remote subsystem equipped with a GSM module receives the data, subsequently transferring it to a server via a USB port. The server stores and remotely displays the data on a website. Additionally, a separate SMS-based telemedicine system is described, wherein the patient's temperature is measured using an Infrared temperature sensor (MLX90614), and ECG signals are acquired via electrodes interfaced with the PIC16F877 microcontroller. Additionally, a wearable hardware device is developed to monitor the baby's biological status by integrating motion,

temperature, and heart rate sensors (both optical and pressure), all controlled by the microcontroller and connected to a Bluetooth module for wireless communication. Furthermore, the paper addresses the monitoring of temperature and humidity parameters by utilizing skin temperature and air-temperature probes to monitor the baby's surroundings. The humidity level within the incubator is measured using a humidity sensor from the SYHS2XX series. These signals are interfaced with the PIC microcontroller 18F4550, with communication facilitated through a GSM modem. [2] The author has devised an efficient and cost-effective Baby Monitoring System based on IoT technology, operating in real-time. A novel algorithm is pivotal in ensuring optimal baby care. The system utilizes a Node Microcontroller Unit to collect and transmit data via Wi-Fi to the Adafruit MQTT server. Various sensors are employed to monitor the baby's vital parameters, including ambient temperature, moisture levels, and crying. Using Nx Siemens software, a prototype has been created, with red Meranti wood selected as the material for the cradle. The system architecture encompasses an automatic cradle swing mechanism activated by a motor in response to the baby's cries. Furthermore, parents can remotely monitor their baby via a webcam and activate a lullaby toy through the MQTT sensor for the baby's amusement. This demonstrates the effectiveness of the baby monitoring system in monitoring both the baby's well-being and the surrounding environmental conditions, as evidenced by the prototype. [4]

2. METHODOLOGY

2.1 Hardware Development

The Smart Cradle System is equipped with advanced features to ensure the comfort and well-being of infants, primarily focusing on temperature regulation and sound recognition. Firstly, the system incorporates the Temperature Sensor DHT11, which continuously monitors the ambient temperature. If the temperature exceeds the threshold of 24°C, indicating a potentially uncomfortable environment for the baby, the system automatically activates a fan to provide cooling relief. The fan will switch off if the temperature falls below 24°C in the environment. Additionally, the system integrates the LM393 Sound Sensor, which is designed to detect the frequency of the infant's cries. This sensor operates in two distinct modes to address different crying patterns. If the cry frequency falls within the range of 500-600Hz, indicative of a relatively low-intensity cry, the system triggers a motor-assisted toy rotation to help soothe the infant. On the other hand, if the cry frequency registers between 900-1000Hz, suggesting a higher intensity cry, the system initiates the playback of preloaded songs accompanied by the comforting voice of the mother through a built-in speaker. In both scenarios, the cradle swings gently using a 12V wiper motor, providing a calming motion to aid in the baby's relaxation. By seamlessly combining temperature regulation with intelligent sound recognition capabilities, the Smart Cradle System offers a comprehensive solution to support infant comfort and promote peaceful sleep patterns. Additionally, the cradle is constructed with wood, ensuring the infant's safety as it does not conduct electricity, even with various components present.

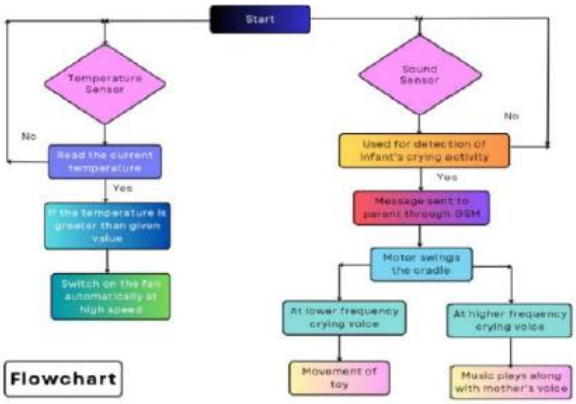


Figure 1:- Flowchart

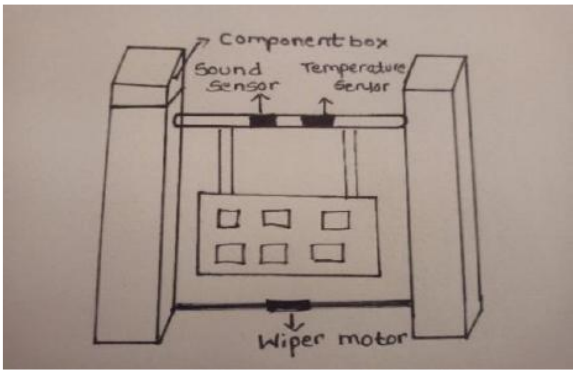


Figure 2:- Prototype design of Cradle

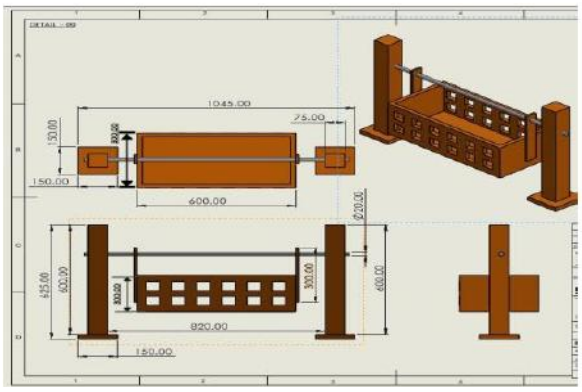


Figure 3:- Prototype designing using Solid Works software



Figure 4:- Physical structure of the Cradle made from MDF wood as wood is not a conducting material

Algorithm for Temperature Sensor:

Step 1: start
Step 2: Record the template
Step 3: temperature exceeds 24° C.
Step 4: fan's start.

Algorithm for Sound Sensor :
Step 1: start
Step 2: Sound detects the range between 500-600 Hz
Step 3: toy rotates
Step 4: Sound detects the range between 900-1000 Hz
Step 5: music plays

2.2 Software Development

The Website application is designed in a more user-friendly manner. Graphics and pictures are used to represent the functionality/options of the smart cradle system. This will deliver the content to the user and they can easily navigate to their requirements without much frustration. Smart cradles for babies can be easily controlled by the web application through any kind of device. The following features are available in it.

- Playing Music
- Swing the cradle
- Turn on/off the fan
- Checking the room temperature
- Get notifications about vaccinations

To start with the application first and use the features available in the application, you need to have an account, if the user already has an account, the user can login. Otherwise, create a new account. After once when the login and sign-up process is done a dashboard web page gets open. The dashboard contains several buttons/options through which we can control the cradle. The features are mainly.

1) **Swing of the Cradle -** When the button gets on the cradle starts to swing in a specific pattern which is mentioned in the hardware part of the project.
2) **Temperature & Fan -** When the temperature exceeds more than 24°C it indicates that the temperature is hot for the baby and the fan should be turned on which one can do with the help of the application.
3) **Playing Music -** When the frequency of the crying baby is detected by the sound sensor and comes to a level of 900-1000Hz the music which is stored in the application can be played there will be different music's stored which also can include the mother's voice. Music can be a great tool to make a baby stop crying or just a fun way to bond. So one can make use of this application to play some music to soothe your baby.
4) **Rotating of a toy-** When the frequency of the crying baby is detected by the sound sensor and comes to a level of 500-600Hz the toy starts to rotate in a particular direction and speed.
5) **Notifications about vaccinations-** The user has to set the date and time of the vaccination and on that day of vaccination a reminder will be generated on the application and there is no irresponsiveness regarding vaccinations.

The web application is done using 2 main parts frontend and the backend development. The Frontend development is done through HTML, CSS, and JavaScript & the Backend development is done using Node.js. The data is stored by the SQL query language. This all completes the features of the smart cradle for baby and how the process is to be done through a web application.

Figure 5 shows the prototype image of the login page of the smart cradle for baby which includes username and password. Figure 6 shows the prototype image of the dashboard page of the smart cradle for baby which includes the various features that can be controlled through the website.



Figure 4:- Prototype of the home page of the website.

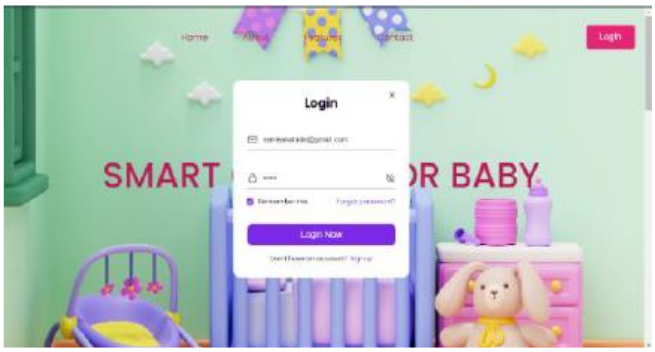


Figure 5:- Prototype of the login of the website.



Figure 6:- Prototype of the dashboared of the website.

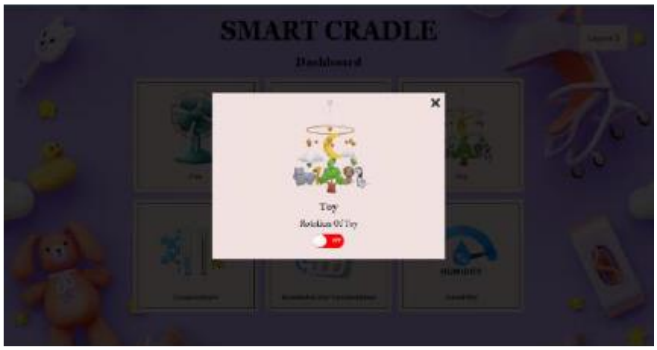


Figure 7:- Prototype of the toy rotation of the website.



Figure 8:- Control of smart cradle through Blynk IOT app.

3. CONCLUSION

In conclusion, the smart baby cradle offers parents valuable peace of mind by combining cutting-edge technology with thoughtful design. With a built-in temperature sensor, it ensures the optimal room temperature for the baby's comfort, adapting as needed. The sound sensor, attuned to the baby's cries, allows swift responses to their needs. Moreover, the inclusion of scheduled vaccination reminders through the accompanying app ensures parents never miss important healthcare milestones. This holistic approach to infant care not only enhances the well-being of the child but also lessens the anxiety of new parents. The Smart Cradle for Baby represents a promising innovation in infant care, aiming to combine safety, comfort, and technology. This concept has the capability to enhance parental peace of mind through advanced systems, while also contributing to safe sleep practices. The integration of smart home technology and a focus on energy efficiency further amplifies its appeal. However, to fully realize its potential, collaboration with healthcare professionals, adherence to regulations, and a keen understanding of user preferences will be crucial. As the smart baby care product market expands, the Smart Cradle is positioned to play a crucial role in reshaping how parents attend to the needs of their infants

4. FUTURE SCOPE

1. **Health Monitoring:** A number of sensors may be included into smart cradles to track an infant's temperature, heart rate, and breathing patterns. By using advanced analytics, parents may receive timely alerts about any irregularities and gain insights into their baby's health state.
2. **Sleep Analysis and Optimization:** Smart cradles may provide tailored recommendations to enhance the baby's sleep quality, promoting healthy development, by gathering data on sleep patterns and environmental variables.
3. **AI-driven Advice and Insights:** By analyzing data gathered by smart cradles, artificial intelligence algorithms may offer tailored advice and recommendations for baby care, including feeding schedules, calming methods, and developmental milestones.
4. **Safety Features:** More sophisticated safety features like automatic swaddling mechanisms, anti-roll designs, and integrated air quality sensors may be included in future generations of smart cradles.

REFERENCES

[1] Aniruddha Patil; Nitesh Janardan Patil; Anjali Deepak Mishra; Yogita Deepak Mane; “Smart Baby cradle” in 2018 International Conference on Smart City and Emerging Technology (ICSCET) -January 2018

[2] Ms. Amruta H. Muradnar; Ms. Komal P. Lahamge; Ms. Ujwala I. Borse; Mr. D.S.Thosar; “Smart Baby Cradle Monitoring System” in International Research Journal of Engineering and Technology (IRJET) -May 2020

[3] N Lakshman Pratap; K Anuroop; P Nirmala Devi; A Sandeep; Sunanda Nalajala; “IoT based Smart Cradle for Baby Monitoring System” in 2021 6th International Conference on Inventive Computation Technologies (ICICT) -20-22 Jan. 2021

[4] Waheb A. Jabbar; Hiew Kuet Shang; Saidatul N. I. S. Hamid; Akram A. Almohammedi; Roshahliza M. Ramli; Mohammed A. H. Ali; “IoT-BBMS: Internet of Things-Based Baby Monitoring System for Smart Cradle” in Journals and magazines of IEEE-July 2019.

[5] https://cepdnaclk.github.io/e17-3yp-Smart-Cradle/.

[6] https://www.twinschip.com/ESP32_WROOM-32_Development_Board

[7] https://www.mouser.com/datasheet/2/758/DHT11-Technical-Data-Sheet-Translated-Version-143054.pdf

[8] https://components101.com/modules/lm393-sound-detection-sensor-module

[9] SIM800A GSM Module | CE Store (cestore-mm.com)

[10] Wiper Left Angle Reversible Electric Worm Motor 12V 70W DC 35 50 RPM High Torque | eBay

[11] 12V High Quality 775 Brushed DC Motor | Microdcmotors.com

❑ ❑ ❑ ❑ ❑ ❑

AET 043

IoT based Chemiprobe for Surfactant Performance Modelling

D. Rodda, A. Vijayan, N. Rajpurkar, A. Jadhav, V. Kshirsagar
Author Affiliation(s)
(Final Year Students of Electronics & Telecommunication Engineering, VIT, Mumbai, Faculty Dept. of Electronics & Telecommunication Engineering, VIT, Mumbai)

***Abstract*– This study introduces a low-cost, user-friendly prototype for accurately measuring the Critical Micelle Concentration (CMC) of solid surfactants in detergents. Employing conductometric analysis with simple sensors and a microcontroller, the system leverages an automated Minimum Slope Detection Algorithm (MSD) for reliable CMC estimation. This innovative approach offers a portable and cost-effective alternative to existing methods. Testing with various commercially available surfactants demonstrates its effectiveness. This user-friendly tool holds promise for ongoing surfactant performance evaluation in research and industry.**

***Keywords*– Solid surfactants, Detergents, Critical Micelle Concentration (CMC), Cost-effective measurement, User-friendly technique, Conductometric analysis, Microcontroller, Minimum Slope Detection Algorithm (MSD), Surfactant performance evaluation, Internet of Things (IoT).**

1. INTRODUCTION

The Challenge of Surfactant Performance Evaluation: Surfactants are essential in detergents for effective cleaning, with their performance pivoting around the Critical Micelle Concentration (CMC). Traditional methods for determining CMC aret costly, intricate, and skill-intensive, which restricts the optimization of detergent formulations and comprehensive performance evaluation.

Introducing the IoT-Enabled Chemiprobe: Addressing these challenges, this study introduces a novel, cost-effective, and user-friendly IoT-enabled chemiprobe system. It employs conductometric analysis using sensors like Total Dissolved Solids (TDS) and Dallas temperature sensors integrated with an Arduino Nano microcontroller, simplifying the CMC measurement by bypassing the need for expensive specialized equipment.

Real-Time Data and Advanced Analysis: The system transmits data to a cloud platform (ThingSpeak), where Python scripts analyze the data to identify CMC, generate concentration vs. TDS graphs, and facilitate advanced statistical analyses for a comprehensive evaluation of different detergents.

Significance and Innovation: This IoT-enabled chemiprobe system not only simplifies real-time data acquisition and allows immediate experimental adjustments but also makes the results accessible remotely to stakeholders and researchers. This fosters collaboration and knowledge dissemination. The system not only enhances understanding of surfactant behavior in detergents through CMC determination but also employs a variety of performance metrics and statistical analyses for a more holistic evaluation.

Overall, this research presents a transformative, cost-effective, and user-friendly approach to surfactant performance evaluation, leveraging an IoT-enabled chemiprobe system. This innovation paves the way for future advancements in detergent formulation and scientific exploration in surfactant performance.

2. LITERATURE REVIEW

Problem and Existing Knowledge: Surfactants are crucial components in cleaning products like detergents, influencing their effectiveness. Several factors affect surfactant performance, including chemical composition, temperature, and solvent properties. The "gold standard" for evaluating surfactant efficiency is measuring the Critical Micelle Concentration (CMC), which indicates the minimum concentration at which micelles form in solution. Traditionally, advanced laboratory techniques are used for CMC measurement. However, these methods can be expensive, complex, and require skilled personnel.

Gaps and Need for Further Research: The existing literature review provided highlights the limitations of current CMC measurement techniques:

Miller et al. (2008): While this study focuses on comparing surfactant efficiency through emulsification and surface tension reduction, it emphasizes the need for a comprehensive approach to surfactant performance evaluation, which aligns with this research.

Das et al. (2012): This research underscores the importance of accurate CMC measurement but also highlights the reliance on specialized equipment, limiting broader application.The proposed chemiprobe system aims to address this gap by offering a cost-effective and accessible method.

Chaudhary et al. (2020): This study explores the properties of commercial soaps, bridging the gap between fundamental research and real-world applications. This aligns with this project's goal of evaluating surfactants in a practical context.

Contribution: Research proposes a novel, low-cost, and user-friendly IoT-enabled chemiprobe system for surfactant performance evaluation. This system utilizes conductometric analysis with readily available sensors and a microcontroller for automated CMC estimation. This addresses the limitations of existing methods by: Cost-effectiveness: Employing simple sensors and open-source software makes the system affordable.

Accessibility: The compact design and user-friendly interface make it easy to operate without requiring specialized skills.

Real-time Data and Automation: Data is collected and analyzed in real-time using a microcontroller and Python scripts, enabling efficient data acquisition and analysis.

IoT Integration: Cloud connectivity allows for remote data access and facilitates collaboration.

Transform Surfactant Performance Evaluation: The chemiprobe system offers a simpler, more accessible alternative to traditional CMC measurement techniques.

Enhance Detergent Formulation: Real-time data can guide the optimization of detergent formulations for improved cleaning performance.

Advance Surfactant Research: This system can facilitate broader research efforts by making CMC measurement more accessible.

3. METHODOLOGY

This research investigates the efficacy of an IoT-based chemiprobe system for surfactant performance evaluation. This section details the methodology employed, adhering to established research methodology guidelines and emphasizing the user-centric design of the system.

3.1 Test Sites and Datasets

The success of the chemiprobe system relies on well-designed test sites and comprehensive datasets. Here's how we approached this aspect:

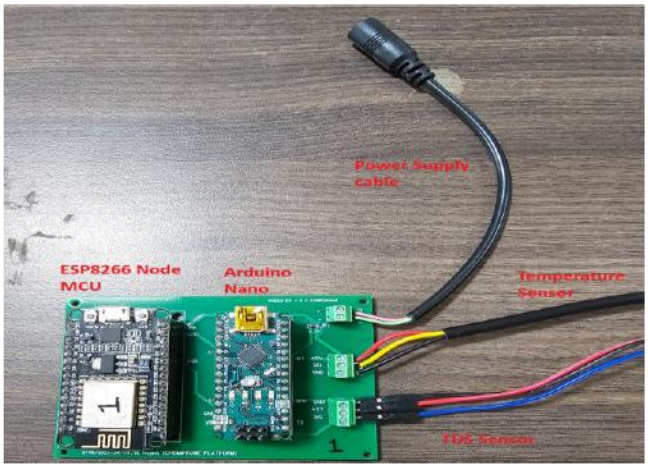


Figure 1. Data collector Hardware

Laboratory Setup:

Chemical Reagents: We designated a controlled laboratory environment for preparing detergent solutions with known concentrations using various surfactant types. This allows for capturing diverse surfactant behavior.

Sensor Calibration: The TDS and temperature sensors were meticulously calibrated in this controlled environment before experiments to ensure accurate readings throughout data acquisition.

Controlled Experiments:

Concentration Gradients: We implemented a systematic approach by varying detergent concentrations around and below the expected CMC. This ensures a comprehensive exploration of surfactant behavior at different concentration levels.

Repeated Trials: To minimize experimental variability and assess the reproducibility of results, experiments were conducted multiple times with each detergent formulation.

Environmental Monitoring: Exploring applications in environmental monitoring where surfactants play a role, such as water treatment, could provide valuable insights into real-world performance.

Thing Speak Cloud Integration:

Simulated Data: Prior to using real data, simulated data uploads to the ThingSpeak cloud platform were conducted to ensure seamless integration and data transmission. This helped identify and address any potential issues before real-world experimentation.

Security Protocols: Security measures were implemented to protect data integrity and user access while utilizing the Thing Speak cloud platform.

Detergent Selection: Commercially available detergents representing a variety of formulations were included in the experiments to assess the system's performance with common products.

Specialized surfactant formulations designed for specific industrial applications could also be tested.



Figure 2. Industrial Surfactant Samples

3.2 Hardware and Software Requirements

The chemiprobe system leverages readily available hardware and open-source software to achieve cost-effectiveness and accessibility.

Hardware:

Arduino Nano: The central microcontroller unit responsible for data acquisition and initial processing from the sensors.

NodeMCU: Facilitates communication between the Arduino Nano and the Thing Speak cloud platform for data transmission.

TDS Sensor: Measures the Total Dissolved Solids (TDS) in the detergent solutions, providing an indirect indicator of surfactant concentration.

Dallas Temperature Sensor: Measures the temperature of the detergent solutions, as temperature can influence surfactant behavior.

Power Supply: Ensures a stable source of power for the Arduino Nano and NodeMCU.

Software:

Arduino IDE: Integrated Development Environment used for programming the Arduino Nano to control sensor readings and data processing.

NodeMCU Flasher: Programs the NodeMCU for efficient data transmission to the cloud platform.

Python (with NumPy and SciPy): A versatile programming language used for data analysis tasks like graph plotting, statistical calculations, and algorithm implementation.

ThingSpeak: A cloud platform offering real-time data storage, visualization tools, and remote accessibility.

3.3 Workflow

The chemiprobe system follows a structured workflow encompassing experimental setup, data acquisition, transmission, analysis, and result presentation.

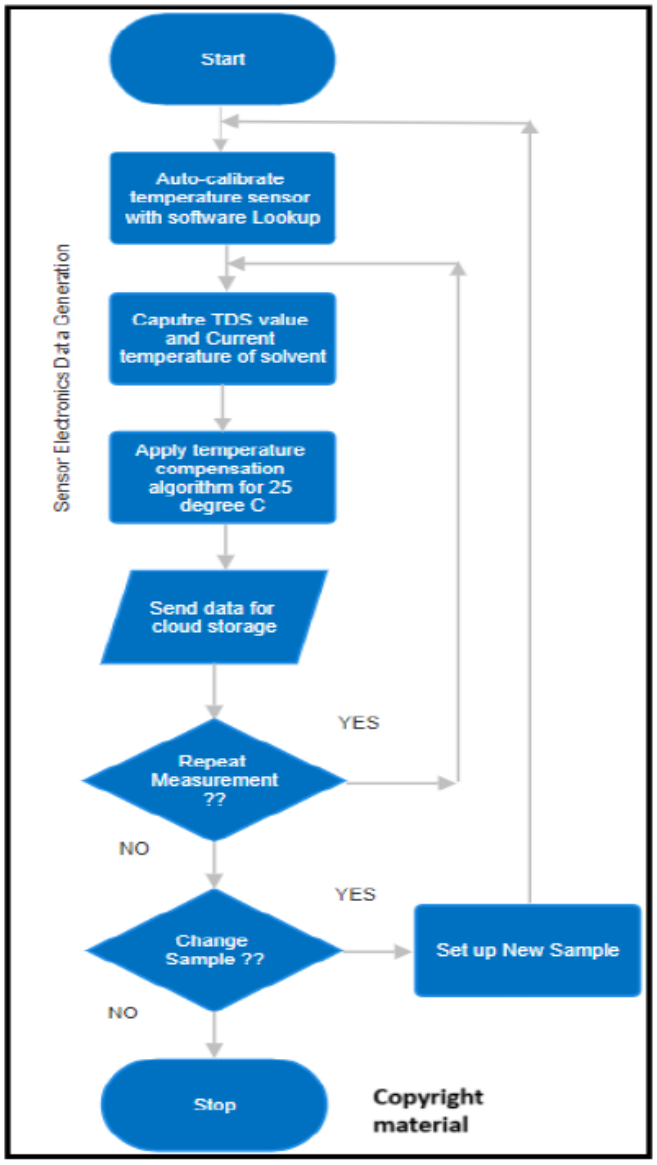
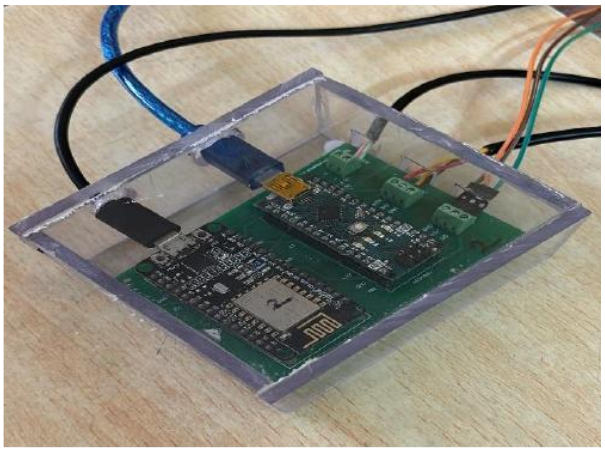


Figure 3. Simple Data Generation Chart

Experimental Setup: As mentioned in the test sites and datasets section, we established a controlled laboratory environment.



Figure 4. Precise measurements in mg



Sensor Integration: Integrate the calibrated TDS and Dallas temperature sensors into the detergent solutions, ensuring proper contact.

Data Acquisition using Arduino Nano: The Arduino Nano reads data from the sensors at a predetermined frequency, capturing both TDS and temperature measurements.

Data Transmission: The acquired data is then transmitted from the Arduino Nano to the NodeMCU for internet connectivity.

IoT Integration: The NodeMCU transmits the data to the Thing Speak cloud platform for real-time storage and accessibility.

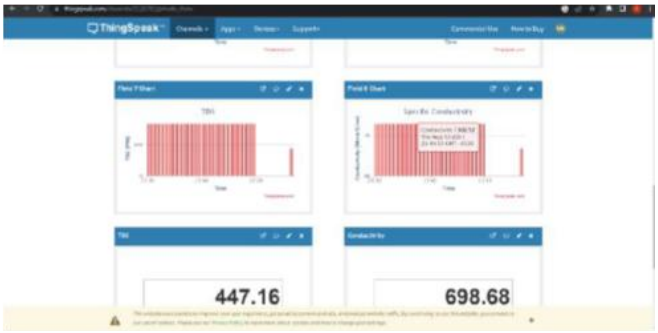


Figure 5. Thing Speak Realtime UI

Web Interface: A user-friendly web page is developed to visualize and disseminate the data stored on Thing Speak. This interface caters to researchers and potentially industry professionals, allowing for easy access to real-time and historical data.

3.4 Data Analysis

Implementing algorithms for CMC (Critical Micelle Concentration) detection from the concentration vs. TDS graphs. This will estimate the concentration at which micelle formation occurs for each detergent formulation.

3.5 User Interface

The chemiprobe system prioritizes a user-friendly experience. A web interface is developed to allow researchers and potentially industry professionals to interact with the system:

Data Visualization: Real-time and historical data can be visualized through graphs and charts depicting the relationship between concentration and CMC.

Data Access: Users can access experiment data and download reports for further analysis.

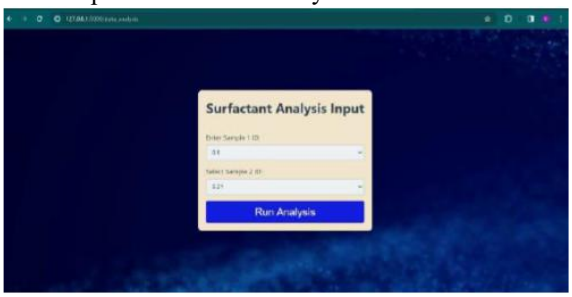


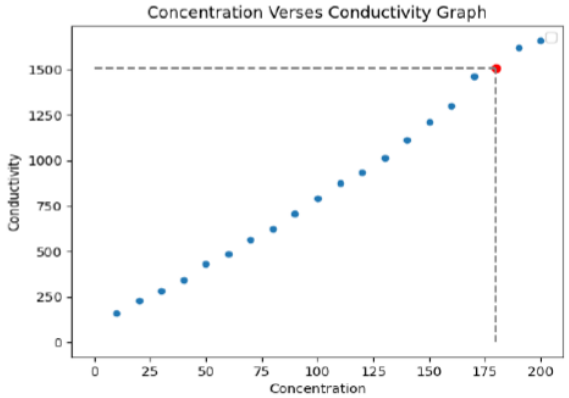
Figure 6. Web User-Interface

4. RESULTS

In this study, six different detergent samples were subjected to comprehensive analysis to evaluate their surfactant properties and performance characteristics. The selection of multiple samples allowed for a comparative assessment of surfactant behavior, enabling insights into the variability and efficacy of different formulations. By analyzing key parameters such as Critical Micelle Concentration (CMC), conductivity, and molar conductance, we gained valuable insights into the surfactant's ability to form micelles and facilitate ion-mediated processes. This multi-sample approach enhances the robustness and reliability of our findings, providing a comprehensive understanding of surfactant performance and informing decisions related to formulation optimization and application suitability.

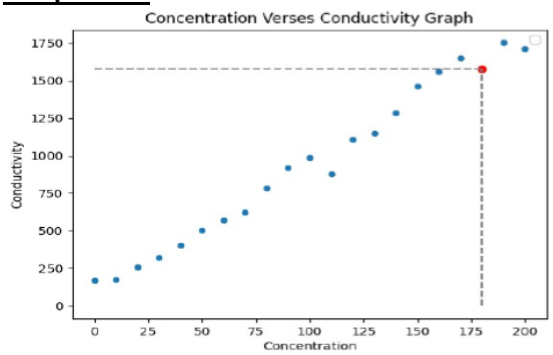
Results of Data Analysis:

Sample 1.21:



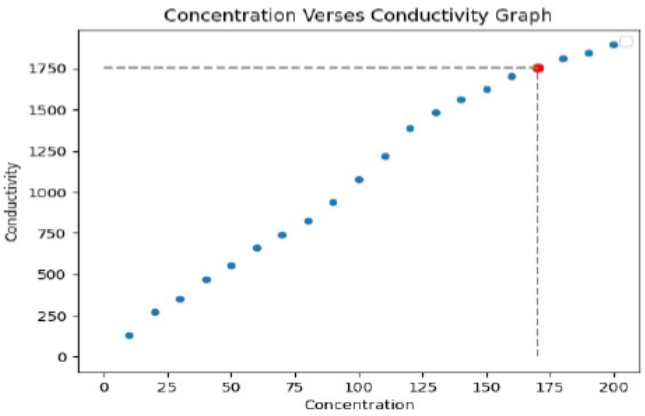
Parameter	Estimated Values
CMC	180 mg
Specific Conductivity	1509.89 micro-S/cm
CMC Moles	13.42
Molar Conductance	8388.28 micro-S

Sample 2.21:



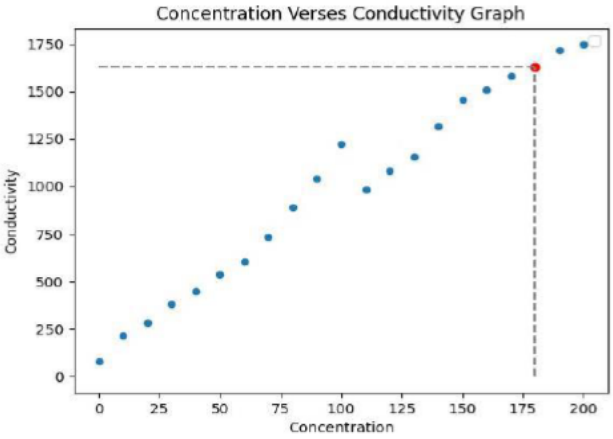
Parameter	Estimated Values
CMC	180 mg
Specific Conductivity	1574.33 micro-S/cm
CMC Moles	13.42
Molar Conductance	8746.28 micro-S

Sample 3.21:



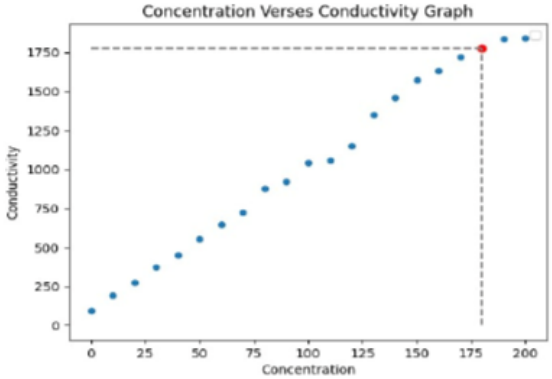
Parameter	Estimated Values
CMC	170 mg
Specific Conductivity	1753.18 micro-S/cm
CMC Moles	13.04
Molar Conductance	10312.28 micro-S

Sample 4.21:



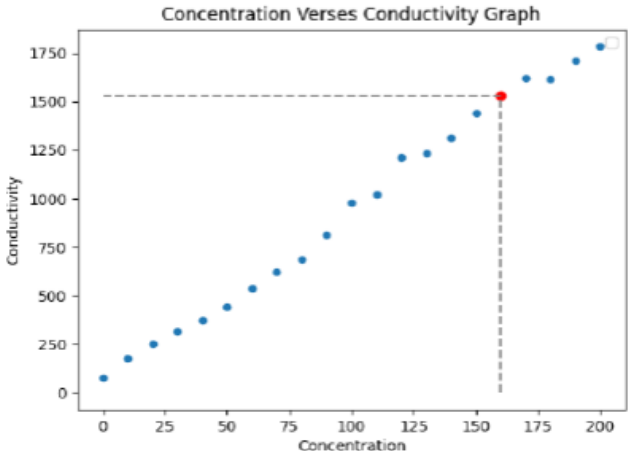
Parameter	Estimated Values
CMC	180 mg
Specific Conductivity	1628.88 micro-S/cm
CMC Moles	13.42
Molar Conductance	9049.33 micro-S

Sample 5.21:



Parameter	Estimated Values
CMC	180 mg
Specific Conductivity	1777.17 micro-S/cm
CMC Moles	13.42
Molar Conductance	9873.17 micro-S

Sample 6.21:



Parameter	Estimated Values
CMC	160 mg
Specific Conductivity	1531.96 micro-S/cm
CMC Moles	12.65
Molar Conductance	9574.75 micro-S

Discussion:

The CMC values obtained for the six detergent samples, highlighting the variability in surfactant behaviour across different formulations. Sample 6.21 exhibited the lowest CMC, suggesting a higher propensity for micelle formation at lower concentrations. Conversely, samples with higher CMC values may require higher concentrations for optimal performance.

Conductivity and molar conductance measurements provided further insights. Sample 3.21 displayed the highest values, indicating stronger ionization and potentially superior performance in applications where ion-mediated processes are crucial.

Interpretation:

By analysing these results, we observed a trade-off between micelle formation efficiency (CMC) and ionization (conductivity and molar conductance). Sample 6.21 demonstrated efficient micelle formation, while Sample 3.21 exhibited stronger ionization. This highlights the importance of considering both factors when selecting surfactants for specific applications.

5. FUTURE SCOPE

This research highlights the potential of the chemiprobe system for evaluating surfactant performance. Future efforts will focus on:

- Expanding the Sample Set: Analyzing a wider range of surfactants to enhance the system's applicability.
- Sensor Optimization: Exploring advanced sensors to improve accuracy and sensitivity.
- Machine Learning Integration: Integrating machine learning algorithms for automated data analysis and performance prediction.
- Real-World Applications: Investigating the system's utility in detergent formulation optimization, environmental monitoring, and industrial surfactant applications.

By addressing these areas, we aim to establish the chemiprobe system as a valuable tool for comprehensive surfactant performance evaluation and advancement in the field.

6. CONCLUSION

This research project successfully deployed a novel IoT-enabled chemiprobe system, offering a significant advancement in surfactant performance evaluation. Analyzing diverse samples, we gained valuable insights into surfactant behaviour, with implications for applications like detergent optimization. The project not only contributes to knowledge in both surfactant science and IoT technology, but the user-friendly web interface fosters informed decision-making across various fields. This innovative system lays the groundwork for further advancements in surfactant research and its broader impact.

REFERENCES

1. “Comparing Surfactants: Efficiency and Effectiveness” by Dennis Miller, Alexander Schrem and Matthias Löffler from 17th Surfactants in Solution Conference Berlin, Aug. 17-22, 2008.

2. “Determination of the critical micelle concentration of SDS” from Electronic Supplementary Material (ESI) for Dalton Transactions, The Royal Society of Chemistry 2012.

3. Research Article: “Conductivity, Surface Tension, and Comparative Antibacterial Efficacy Study of Different Brands of Soaps of Nepal” by Narendra Kumar Chaudhary , Arush Bhattarai, Biswash Guragain , and Ajaya Bhattarai from Hindawi Journal of Chemistry Volume 2020, Article ID 6989312, (https://doi.org/10.1155/2020/6989312).

❑ ❑ ❑ ❑ ❑ ❑

AET 044

A Multi-Sensor Approach : Optimizing Tennis Ball Collection

Shruti Vasudev Sawant
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
sawantshruti852@gmail.com

Aaradhya Milind Rane
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
mrunalvichare04@gmail.com

Mrunal Dipjay Vichare
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
mrunalvichare04@gmail.com

Swapnil Umesh Ambre
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
swapnilambre7902@gmail.com

Prof. Dr. Pratik Mhatre
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
pratik.mhatre1@vit.edu.in

***Abstract*— This paper focuses on the growing utilization of robotic systems across various domains, including sports and everyday life. Particularly, there is a surge in interest in tennis ball collection robots due to their potential to optimize training sessions by conserving time and energy. In this study, an autonomous tennis ball collection robot was engineered, employing LiDAR for environment mapping in two dimensions and a single camera for tennis ball detection. The paper approach was devised for the robot's path planning and navigation. The implemented robot demonstrated a commendable performance, successfully detecting 91% of the tennis balls.**

***Keywords*—LIDAR, Stereo Camera, OpenCV**

I INTRODUCTION

In the dynamic field of autonomous robotics, where innovation is key to enhancing interaction with the environment, robust visual perception stands as a critical element. This paper navigates through the intricate software architecture of an autonomous tennis ball collector, illuminating the integral role played by its image processing module. At the heart of this module lies the detect_ball function, a sophisticated component that leverages advanced techniques from the OpenCV library to discern and track tennis balls with precision. Through meticulous analysis, we unveil the inner workings of this function, elucidating its adeptness in interpreting visual data to autonomously guide the robot through its surroundings.[1] Seamlessly integrated with a camera system, the detect_ball function orchestrates he capture of frames, adeptly identifies tennis balls through intricate color space transformations and contour detection, and iteratively updates variables to facilitate informed decision-making. By unraveling the complexities of this function, our study not only underscores the technical sophistication of the robot's software but also emphasizes the indispensable role of image processing in the realm of autonomous robotics.[2] Furthermore, our findings contribute significantly to the ongoing discourse surrounding the convergence of robotics and computer vision, shedding invaluable light on the intricate software framework that empowers the autonomous tennis ball collector's visual perception capabilities, thereby paving the way for further advancements in autonomous robotic systems.[3]

In professional tennis, matches often last over three hours, with short breaks between sets. To manage the physical strain, players use ball-launching machines during practice. These machines can hold 100 balls and shoot around 10 balls every 29 seconds, allowing continuous operation for roughly 4 minutes and 50 seconds. However, when this limit is reached, players must stop to collect the scattered balls manually, a ask disliked for its time-consuming nature. Tests showed that gathering 85 balls takes approximately 2 minutes and 38 seconds.[4] Consequently, players spend about 35% of their practice time collecting balls, leading to longer breaks and increased fatigue. This inefficiency disrupts training and hampers players' ability to maximize practice time. Therefore, there's a crucial need for innovative solutions to streamline ball collection, reduce downtime, and enhance training efficiency in professional tennis.[7]

II. BLOCK DIAGRAM

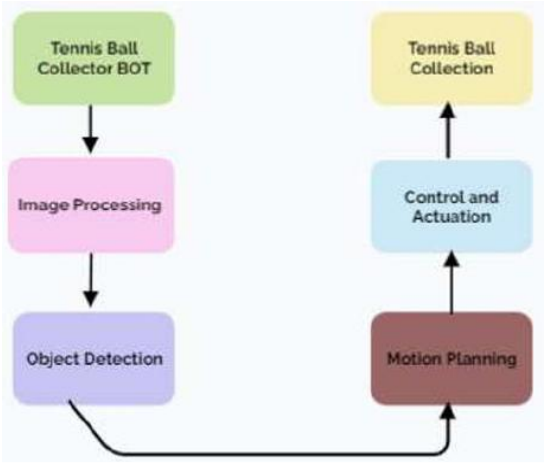


Figure 1 Basic functional block Diagram of the Proposed system.

The system shown in figure 1 can be broken down into five main components:

i. **Image Processing:** This stage is likely involved in the object detection phase, where images or video feeds are processed to identify the tennis ball.

ii. **Object Detection:** This is the initial stage where the bot detects the tennis ball. It likely involves the use of sensors or cameras to identify the ball's location.

iii. **Motion Planning:** Once the ball is detected, the bot plans its motion to reach the ball. This could involve obstacle avoidance and path planning algorithms.

iv. **Control and Actuation:** After the motion is planned, the bot needs to execute this plan. This involves controlling the motors or actuators to move the bot towards the ball. This is the overall system that integrates all the above components to enable the bot to identify and collect the tennis balls.

III. SOFTWARE IMPLEMENTATION

A computer vision system has been developed for a Raspberry Pi-controlled robot to detect and track a ball. The system involves a series of steps: The Raspberry Pi camera captures video frames, which are then processed using OpenCV functions. The processing begins with pre-processing steps such as orientation adjustment, noise reduction, and color space conversion. Subsequently, a mask is generated to isolate green objects, typically the ball, followed by morphological operations to refine the mask. Contours are then detected to locate potential objects, with the largest contour presumed to be the ball. Further analysis is conducted to determine the ball's centre coordinates and size based on the selected contour, with these measurements stored in global variables for the robot's control system. The system operates continuously, displaying the processed frame for visualization, and includes a termination condition for a smooth shutdown of the camera and image processing loop. This setup allows the robot to respond in real-time to the ball's position and dimensions.

IV. HARDWARE AND SOFTWARE INTEGRATION

The function updates pre-defined global variables to share the information about the detected ball with the main program.

These variables typically include:

- x: Ball's horizontal center coordinate within the image.
- y: Ball's vertical center coordinate within the image.
- radius: Radius of the minimum enclosing circle around the ball.
- center: Array containing the (x, y) coordinates of the ball's centroid.

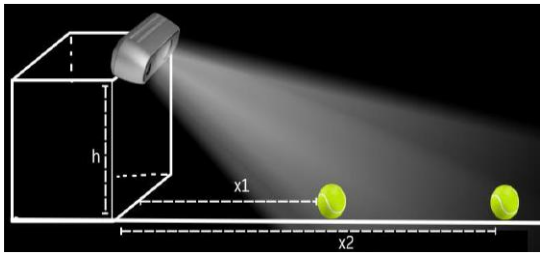


Figure 2: Orientation of camera on the Robot.

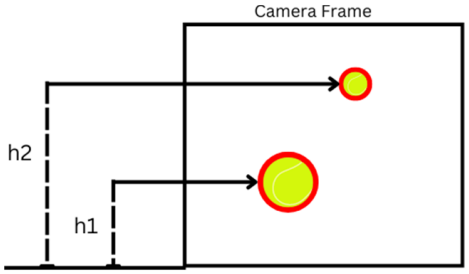


Figure 3: Vision of the Camera

From the camera frame of the bot shown in figure 3 and its orientation in figure 2, the ball is detected by the camera, indicating the ball's coordinate in x, y, and dimensions at the centre of the ball. Consequently, it indicates a basic path planning for the ball collecting by marking "x1" for the ball that is closest to you and "x2" for the ball that is farthest way.

V. Flowchart For Real Time Image Processing

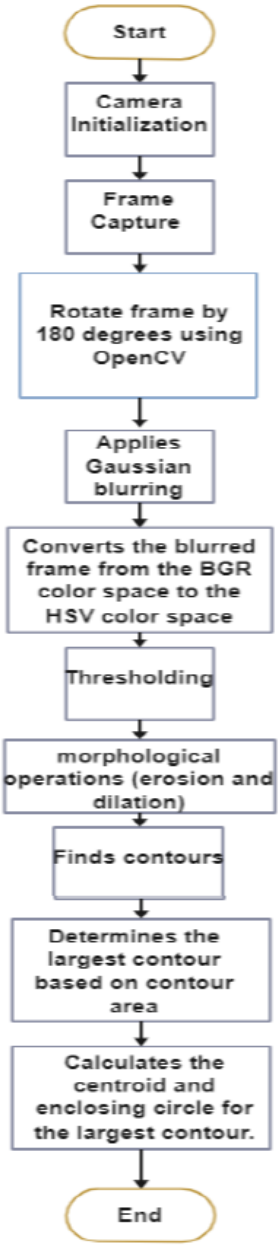


Figure 4: Flow Chart of the Image processing Function.

VI. IMPLEMENTATION APPROACH

1. **Camera Setup:** The detect ball function begins by setting up the Raspberry Pi camera. This involves configuring the desired image resolution and frame rate to capture video data efficiently.
2. **Image Buffer Creation:** A PiRGBArray object is created to serve as a temporary storage container for the raw image data captured by the camera. This buffer facilitates efficient image processing within the function.
3. **Frame Processing Loop:** The function enters a continuous loop, repeatedly acquiring frames from the camera and applying image processing techniques to detect the ball.

4. **Frame Capture:** Each iteration of the loop starts by capturing a single frame from the Raspberry Pi camera. This frame represents the current visual scene the camera observes.
5. **Orientation Calibration:** Depending on the camera's physical mounting, a 180-degree rotation might be necessary using OpenCV's cv.rotate function. This step ensures the image is oriented correctly for subsequent processing.
6. **Noise Reduction:** To mitigate the impact of image noise on object detection accuracy, a Gaussian blur is applied to the captured frame using the cv.GaussianBlur function. This technique smoothens the image, reducing the influence of random pixel variations.
7. **Color Space Conversion:** The function converts the image from the BGR color space (commonly used by OpenCV) to the HSV color space using cv.cvtColor. The HSV color space often provides a more intuitive separation of color information, simplifying object detection based on color properties.
8. **Color Thresholding:** A binary mask is constructed to isolate green-colored objects (presumably the ball) within the image. This mask is generated by applying thresholding with the cv.inRange function on the HSV image. The function utilizes predefined color boundaries (greenLower and greenUpper) to determine which pixels represent the target green color range.
9. **Morphological Refinement:** The generated mask might contain noise or small holes due to image imperfections. To address this, morphological operations like erosion (cv.erode) and dilation (cv.dilate) are applied. Erosion removes isolated pixels and small objects from the foreground (green areas), while dilation helps fill in small gaps or holes within the green regions. This process enhances the mask's accuracy in representing the target object.
10. **Contour Detection:** The function employs cv. Find Contours to identify meaningful shapes (contours) within the processed mask. These contours represent connected regions of pixels with similar characteristics (likely corresponding to objects in the scene).
11. **Nearest Ball Selection:** Assuming the largest contour corresponds to the ball, the function prioritizes this contour based on its area. This heuristic prioritizes larger objects as more likely candidates for the ball.
12. **Ball Feature Extraction:** For the chosen (largest) contour, the function calculates the centroid (center) and the minimum enclosing circle. The centroid represents the geometric center of the contour, while the enclosing circle defines the smallest circle completely encompassing the object. These features provide valuable information about the ball's position and size.
13. **Motor Movements:** Raspberry will get the co-ordinates of the balls and accordingly commands are sent to the motor driver for the.

VII. VALIDATION AND TESTING METHODOLOGY

A. Using Machine Learning

The Tennis Robot Computer Vision Project by Rob flow uses a publicly available dataset called the Tennis Robot Object Detection Dataset. This dataset contains 200+ images of tennis balls, people, and tennis courts. As shown in figure 5 the images are labelled with bounding boxes around the objects, which is used to train the computer vision model to detect and track these objects in real time. The dataset is relatively small, but it is sufficient for training a basic object detection model. The images are also well-balanced, with a good mix of different types of images, such as close-ups of tennis balls, people playing tennis, and wide shots of tennis courts. The major disadvantage of this technique is that the processor should be of high end for the onboard processing and recognition.

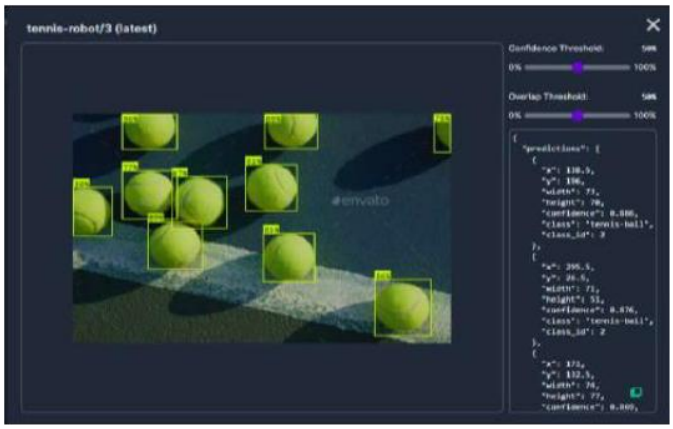


Figure 5: Obtained Results using Robo Flow.

B. Using LIDAR

LIDAR works very effectively for finding the exact position of the ball and its 3D mapping in its 360-degree plane. LIDAR requires a line of sight, and it only detects the objects in that plane, which gives a drawback of objects not been detected outside its line of sight. A combination of a camera and lidar with a larger plane of view will solve this problem. Shown in the Figure 6 these are the results obtained using LIDAR in MATLAB Simulink and it shows that the LIDAR requires line of Sight.[8]

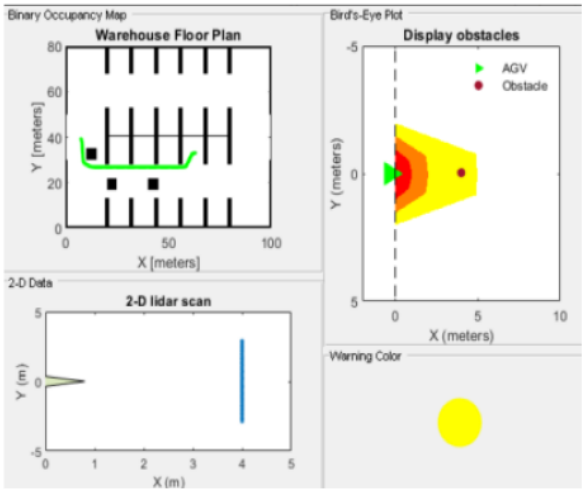


Figure 6: Obtained Results using MATLAB Simulink

C. Using LIDAR camera

The Intel Real Sense Camera is a camera which has a feature of RGB as well as Stereo Camera and is used to create accurate 3D scans of objects. The in-depth information can be used to create more detailed and accurate prediction of the location of the object. It has a range of 4m, and it

accurately detects the ball within its range as shown in figure 7, The Results Obtained are the 3D stereo images of the balls which are detected in figure 8, but this camera are expensive and will add to the additional cost of final product.

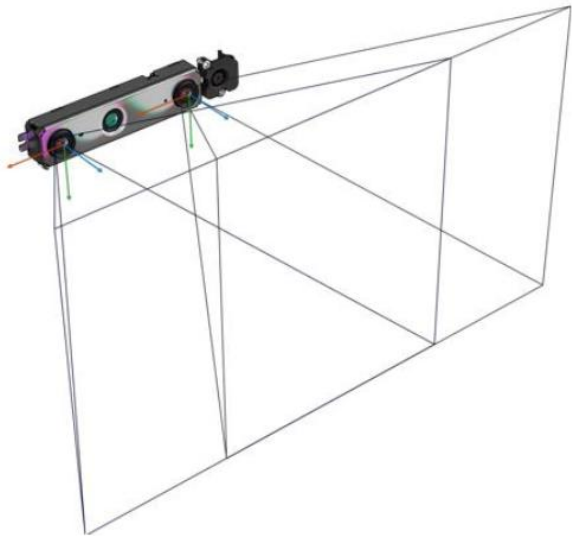


Figure 7: Range of the Camera

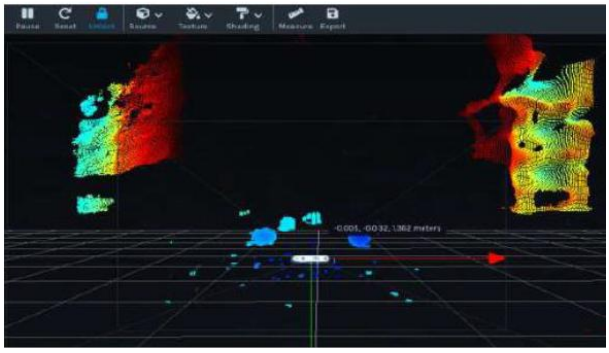


Figure 8: Obtained results using Intel Real sense Camera.

D. Using Raspberry pi Camera

The process involves real-time detection of tennis balls in video feeds using computer vision. Frames are captured and pre-processed for noise reduction and color segmentation. A binary mask isolates potential tennis ball regions, refined through morphological operations. Contours are detected, and the largest one is assumed to be a tennis ball, with its centroid and size estimated the results obtained are shown in figure 9. This method is very cost efficient by using processors such as Raspberry pi.

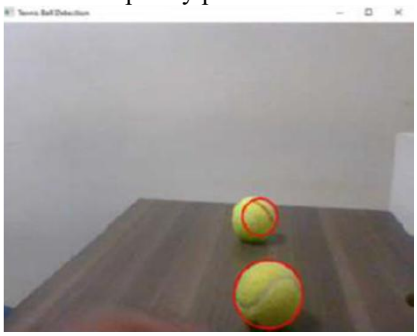


Figure 9: Results obtained using Raspberry pi and Raspberry pi Camera.

VIII. CONCLUSION AND FUTURE ADVANCEMENTS

This research has demonstrated the successful implementation of a robust computer vision system for detecting and tracking a ball in real-time using a Raspberry Pi. Using this algorithm an autonomous tennis ball collector can be implemented. The possibility of integrating machine learning algorithms to improve the robot's object detection and path planning capabilities and development of a mobile app that allows users to control the robot remotely and monitor its status.

REFERENCES

[1] T. Sato, K. Irie, S. Takaishi and H. Kawamoto, "Design and System Integration of Table Tennis Ball-Collecting Robot," 2024 IEEE/SICE International Symposium on System Integration (SII), Ha Long, Vietnam, 2024

[2] ME 481 Design Report Autonomous Tennis Ball Collector http://www.eng.uwaterloo.ca/~madham/files/ME481_Autonomous% 20Ball%20Collector.pdf

[3] Chen, Hung-Kuang & Dai, Jyun-Min. (2017). Interactive Smart Tennis Ball Collecting Vehicle. Recent Patents on Computer Science. 10. 1-1. 10.2174/2213275910666170502150602.

[4] Tennis-Ball Collecting Robot By Majdi Sultan Sohaib Abu Omar Mahmoud Sharabati https://scholar.ppu.edu/bitstream/handle/123456789/6320/Grad%20ProjectALL%20%281%29.pdf?sequence=2&isAllowed=y

[5] DEVELOPMENT OF AN AUTONOMOUS TENNIS BALL RETRIEVER ROBOT DOI: 10.22531/muglajsci.660496

[6] Tennis Ball Picking Robot Based on Bluetooth Control author = {Xiao, Min and Wen, Qingfeng and Ji, Li and Wang, Wei}, title = {Tennis Ball Picking Robot Based on Bluetooth Control}, year = {2021}, isbn = {9781450390453}, publisher = {Association for Computing Machinery}, address = {New York, NY, USA}, url = {https://doi.org/10.1145/3483845.3483848}, doi ={10.1145/3483845.3483848},

[7] V. U. Çabuk, A. Kubilay Şavkan, R. Kahraman, F. Karaduman, O. Kırıl and V. Sezer, "Design and Control of a Tennis Ball Collector Robot," 2018 6th International Conference on Control Engineering & Information Technology (CEIT), Istanbul, Turkey, 2018,

[8] Low Cost LIDAR Sensor System for Autonomous Vehicles by Akshaj Raut and Prof. Pratik Mhatre EasyChair Preprint №10993

[9] H. -K. Chen and J. -M. Dai, "An Intelligent Tennis Ball Collecting Vehicle Using Smart Phone Touch-Based Interface," 2016 International Symposium on Computer, Consumer and Control (IS3C), Xi'an, China, 2016,

[10] W. -S. Yoo, Z. Jones, H. Atsbaha and D. Wingfield, "Painless Tennis Ball Tracking System," 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo, Japan, 2018

[11] D. Chen and C. Zhang, "The design of tennis training platform based on the service robot," 2022 28th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), Nanjing, China, 2022.

❑ ❑ ❑ ❑ ❑ ❑

AET 045

SkyScan - Setting up a Satellite Ground Station

Chirag Sanap
Dept. of EXTC, Vidyalankar institute of technology
Mumbai, Maharashtra
chirag.sanap@vit.edu.in

Vaishnavi Joshi
Dept.of EXTC, Vidyalankar institute of technology
Mumbai, Maharashtra
vaishnavi.joshi@vit.edu.in

Rupesh Madake
dept. of EXTC, Vidyalankar institute of technology
Mumbai, Maharastra
rupesh.madake@vit.edu.in

***Abstract*— The suitable technology that needs to be integrated with weather images is especially satellite imaging. However, getting real-time weather data from satellites can be challenging. However, it is possible to get free weather data all the time from the NOAA 15, NOAA 18, and NOAA 19 weather satellites via Automatic Picture Transmission at various frequencies. This project integrated various data systems such as the QFH antenna,**
the eggbeater antenna, and RTL-SDR to obtain the available satellite frequencies. The antenna was tuned to 137.5 MHz in this project to obtain the data. This was then recorded using SDR, Audacity, and WXtoImg software, decoded, edited, and viewed back. The weather images were, on the whole, clearly identifiable.

***Index Terms*— Satellite Imaging, Software Defined Radio (SDR), Weather Data, NOAA Satellites, Antenna Design.**

I. INTRODUCTION

The number of satellites around the Earth is significantly growing, which is why their application domain is growing too. This allows for the use of satellite communications in new areas, while the educational component of the technology is becoming increasingly relevant. The general access to the local weather satellite image data in real-time remains very limited or complex to decode the hidden data presented in the image.

Since the APT signal is relatively public, the NOAA weather satellite could be an alternate choice to get weather satellite image data. The NOAA 18 and NOAA 19 satellite weather image data will be very useful in GIS. Nevertheless, since most of the existing satellites are LEO, it becomes mandatory to develop a reliable ground station network for receiving most of the satellite-transmitted data. Consequently, an increased number of evenly distributed ground stations should be developed, and a ground station network should be developed. In parallel with advancements in satellite technology, Software Defined Radio (SDR) has emerged as a significant innovation in the field of radio communication. SDR represents a reconfigurable. Radio, wherein the same hardware can be used for different functions at different times. Unlike Hardware Defined Radios, which are fixed-mode radios consisting of active filters, oscillators, mixers, and amplifiers, SDR's physical layer is significantly defined in software. The flexibility inherent in SDR allows it to overcome the limitations of hardware-defined radios. A flexible SDR can tune to any frequency band and receive any modulation across a wide frequency spectrum through programmable hardware. This adaptability not only enables diverse applications but also fosters innovation in radio communication technology. Moreover, SDR serves as an ideal platform for intelligent Cognitive Radio (CR) systems, which require high programmability and controllability for radio transmission and reception. In essence, SDR represents a shift from fixed-mode Hardware Radios towards flexible, low-cost solutions, offering a plethora of possibilities in radio communication technology.

II. HARDWARE IMPLEMENTATION

A. Quadrifilar Helix Antenna (QFH)

The materials required for building the QFH antenna include two sizes of PVC piping to be used for the body and supports, approximately 6 meters of RG 58 coaxial cable, an SMA Connector, and insulating tape with soldering flux. Tools for construction include a measuring tape, electric drill, marker, hand saw, filing tool, crimping tool, and a soldering gun. The construction process commences by procuring the materials and through calculation from the QFH calculator, determining the desired size of the antenna, given the frequency range of desire, for instance, 137.5 MHz for NOAA satellites. This calculation will give the total length of the PVC pipes needed for the base structure. Next, the PVC pipes are cut according to the dimensions obtained from the QFH calculator. This involves creating a drilling template of four circles that are aligned at 90° for drilling holes in the main pipe. Careful drilling and filing ensure that each hole is placed accurately in the pipe. Next, the supports are assembled according to the dimensions obtained from the QFH calculator. This entails cutting the supports using a saw to the required length and putting them into the main pipe to form the base structure of the QFH antenna. The coaxial cables are cut according to lengths obtained from the QFH calculator, ensuring that



Fig. 1. for larger loop



Fig. 2. for smaller loop

there is sufficient buffer length at the ends to avoid errors at the assembly stage. Soldering the coaxial cables is a sensitive
exercise whereby the four layers of coaxial cable are involved. Proper stripping of the wires to remove the braid and core and the soldering of the wires in certain set configurations, such as the connection of A to the braid of B and D, are vital to receiving the APT signals effectively. Lastly, it involves attaching the SMA Connector to the end of one coaxial cable using a crimping tool. This connector is used for the connection of the antenna to the receiver or tracking equipment. Finally, the construction process is completed, and the QFH antenna is ready to be used in the tracking of NOAA satellites.

B. The Eggbeater Antenna

The Eggbeater Antenna is designed with precision for amateur radio enthusiasts interested in satellite communication on the VHF and UHF bands. Its construction starts by procuring materials of use: copper or aluminium tubing/rods, RG 58 coaxial cable (50 ohms), PVC mast, and base components with screws and bolts. This antenna consists of two loops, normally made from the specified tubing or rods. The loops are ingeniously interconnected with RG 58 coaxial cable as a quarter-wave phasing line, with a precisely tuned impedance of 100 ohms. For operating on VHF, the circumferential length of each loop will be 215.28 cm, while for UHF, it's reduced to 75.90 cm. The corresponding lengths of the phasing line are 34.15 cm for VHF and 11.38 cm for UHF. These lengths are accurate for obtaining circular polarization and omnidirectional reception, two essential features for effectively tracking satellites moving through Low Earth orbit. In a building, the loops are positioned at a right angle to each other, forming a full wave, and carefully coupled with each other. To avoid interference and preserve performance, a non-conductive isolator, such as a nylon washer, is used at the junction of the loops, to ensure the loops do not touch each other directly. This will make the reception and transmission of the signal optimum, especially when trying to reach quickly moving satellite targets.

A 4:1 BALUN (balanced to the unbalanced transformer) is also included in the antenna assembly to ensure that the unbalanced output from the RG 58 coaxial cable (50 ohms) is matched to the balanced loading of the antenna, which normally has a terminating impedance of 200 ohms. For calculating the dimensions of the antenna of desired frequency we use the antenna calculator : https://www.ws6x.com/antcalc.htm

III. SOFTWARE IMPLEMENTATION

A. SDR-sharp

SDR-sharp is a software for software-defined radio processing and reception used by many users from amateur enthusiasts to professionals. Its user-friendly interface supports a range of SDR hardware in use, with tuning across a wide frequency range of VLF to microwave frequencies. It features realtime spectrum visualization, frequency scanning, and signal demodulation to cover a wide range of radio signals from AM to digital formats. The expandability of the software due to plugins, which adds extensibility to it, makes it a versatile tool for exploration in the radio spectrum and experiments in communication, both leisure and professional, in the SDR community. We successfully set up SDR-sharp to receive audio signals from NOAA weather satellites at 137 MHz through precise frequency settings of, "NFM" mode selection, and the appropriate filter bandwidth of 34 kHz to reach all corners of the signal. By adjusting the RF Gain, as well as proper audio output choosing WAV format with appropriate bitrate, and managing the files, the recordings were high quality with easy access. Systematic antenna alignment, monitoring of the signal, and timely initiation of the recording ensured stable reception of the signal and data capture. Post-recording enhancements and decoding of data further supported the audio quality to meet the accuracy and clarity requirements of NOAA satellite signal reception at 137 MHz.

B. WXtoImg

WXtoImg is a specific software application for amateur radio enthusiasts and weather enthusiasts that decodes audio signals from weather satellites like those run by NOAA—National Oceanic and Atmospheric Administration. It is specifically made to decode audio signals from the satellites, which represent the audio recordings made by various equipment such as software-defined radio (SDR) receivers.

Subsequently, the recorded audio is turned into visible images of different satellite types: visible, infrared, and thermal imagery. These are the meteorological images, which bring quite useful information on the weather conditions, including the amount of clouds, temperature patterns, and storm systems, which can be used to help forecast the weather, do research, and education. Among the features that the software contains are image processing, calibration, and customization, which improve the quality and accuracy of the decoded images. The software is used around the world by amateur radio operators, meteorologists, educators, and weather enthusiasts.

C. Audacity

Audacity is a popular free and open-source audio editing software. It's widely used for recording and editing audio tracks on various platforms like Windows, macOS, and Linux. Audacity offers a range of features for editing audio files, including cutting, copying, and pasting audio segments, as well as applying various effects and filters. It's a handy tool for tasks like podcasting, music production, and sound editing for videos.

D. Virtual Audio Cable

Virtual Audio Cable is a software application for creating virtual audio devices that allows you to connect different

audio applications. It is like a virtual patch panel, enabling you to route audio streams from one application to another within your computer.

IV. SOFTWARE DEFINED RADIO (SDR)

Radio technology has been dramatically changed with software-defined radio, whereby radio components such as

modulators, demodulators, and tuners are migrated from analog hardware to software that is enabled by today's computing and an analog-to-digital converter. Consequently, the final price of such a little device such as the RTL-SDR, a USB dongle of approximately Doller 30, has a computer-based radio scanner without the need for an Internet connection. RTL-SDR can indeed receive live radio signals on a wide frequency range—from 500 kHz to 1.75 GHz, depending on the model—and is freeware, supported by community-based software. It, however, achieves a maximum stable sample rate of 2.56 MS/s, which is some tens of kilohertz shy of the nominal 3.2 MS/s for the fear of sample drops at a higher rate. Additionally, RTL-SDR dongles are generally designed with an input impedance of close to 75 Ohms, though this may vary widely from frequency to frequency.

V. WORKFLOW

The first step of our satellite tracking process is to track the movement of select satellites and their orbital paths. These satellites have a constant path and pass over a selected location twice per day. We get useful information from websites such as N2Yo, which gives us details on satellite passes above our region, including elevation and the total duration of pass. We then select a clear, unobstructed location in which to receive the signal, a location free of trees or buildings. At this location, we attach a set of specialized antennas tuned to 137MHz frequency reception because this is the frequency above which the satellite transmits. The most popular antennas used include the Quadrifilar Helix (QFH) or the Eggbeater antennas. These antennas are then attached to an SDR dongle using SMA connectors. The SDR dongle plays a vital role since it converts the analog signals that come from the satellite into a digital signal. This digital signal is then sent to a computer with the software needed, which includes SDR-sharp designed to work with SDR dongles. SDR-sharp allows us to finetune the frequency range that is needed for our use, which

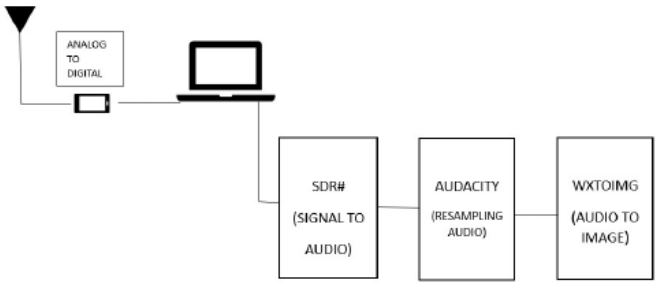
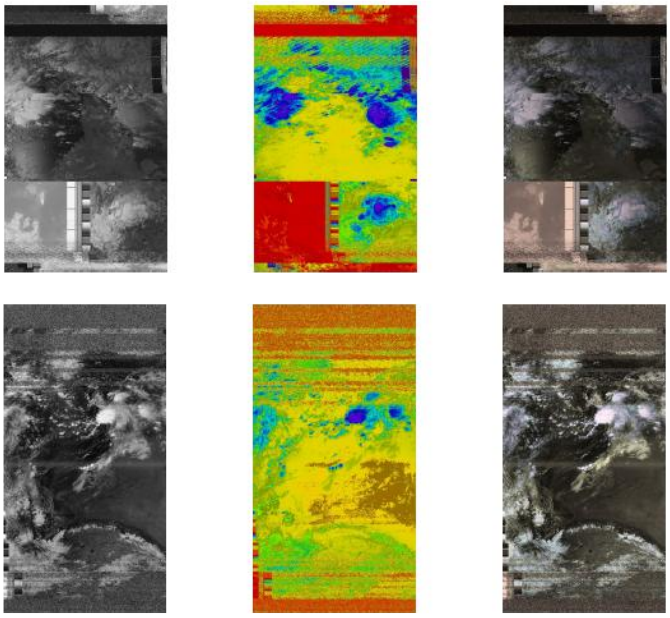


Fig. 3. Workflow

makes this digital signal into audio files. After finetuning, the software allows us to finetune the bandwidth needed so that the unwanted noise can be filtered so that the actual signal can be recorded. After recording, the audio files are then processed, which include resampling to a standardized 11025Hz frequency using software like Audacity. After resampling, the audio signal is decoded and is transformed into an image. This image is the overall view of the Earth's surface that shows water bodies, landmasses, and clear outlines of clouds. From this image, we can get an idea of the weather patterns. For instance, the cloud cover of the analysis gives an idea of these patterns. For Temperatures, we use thermal filters to obtain thermal images at different points on the map. Thus, in short, our systematic approach amalgamates the cutting edge of technology with serious data analysis in offering valuable insight from satellite imagery, mainly for the prediction and monitoring of temperature.

VI. RESULT AND DISCUSSION

Throughout our project, we tracked and decoded signals from three leading weather satellites: NOAA 15, NOAA 18, and NOAA 19. We did so using one very careful setup that included two very different kinds of circularly polarized antennas: the Quadrifilar Helix (QFH) antenna and the Eggbeater antenna. Because of its innate advantages, including relative immunity to such atmospheric perturbations as signal loss caused by rain and interference, circular polarization was our preferred mode of satellite communication. While Eggbeater antennas theoretically have higher gain, we realized that the QFH antenna was dramatically superior to the Eggbeater antenna in practice: our signals had a higher received signal quality and suffered less from the noise floor. This may be because the structure of the QFH antenna, which does not contain the open elements that are part of the Eggbeater antenna, removes a point of weakness for the signal strength. This gives the signal strength a better means for maintaining integrity. We determined that our decoding function divided every pass into two separate channels: Channel A and Channel B. These two channels are sent from NOAA satellites and have two distinct purposes. Channel A provides visible frequency range imagery during daytime but transmits data only when the sun is up. Channel B, in contrast, provides infrared imagery regardless of the time of day. We determined that the processed



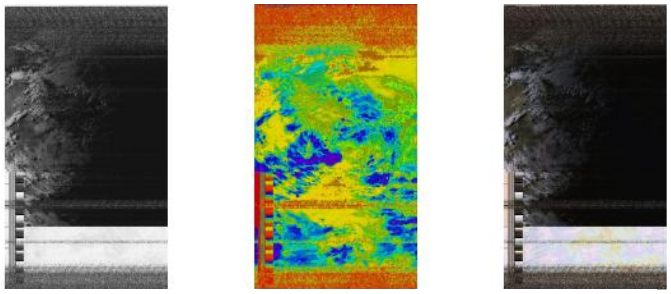


Fig. 4. Raw, Thermal and False colour images of NOAA 15,18,19

nature of the decoded images had an adequate influence on the elevation of the satellite tracked. As the elevation of the tracked satellite increases, the receipt quality of the signal gets better and results in improved imagery. We included thermal filters in our image processing pipeline to gain deeper insights into temperature distributions occurring at some points in the decoded images. Applying thermal filters yielded the capacity to estimate temperatures almost precisely, contributing to our understanding of weather patterns and phenomena shown in the satellite images. Altogether, our tracking and decoding of satellite signals by critical setup design and image processing techniques revealed very useful information for weather monitoring and prediction, which highlights how careful design of setup and image processing techniques are imperative to all forms of satellite communication.

VII. CONCLUSION

Strong satellite signals were received using a proper antenna tuned for 137 MHz for weather satellites. The difficulties that came across for this project were RF interferences. Although there were RF interferences for the reception of the signal, it was mitigated by receiving the signal at a place where there is a low usage of different RF equipment. After the recording of the satellite signal, the signal was further processed using different software for decoding. The decoding procedure for

APT signals and LRPT signals were different so different software had to be used. Once the decoding process for both signals was complete, stunning images, containing clouds and other information were Displayed. All these images from the weather satellites can be used to forecast the weather. The

software that allowed further processing of the image was used which gave multiple versions of the image.

ACKNOWLEDGMENT

We gratefully acknowledge our teachers for their invaluable guidance and support throughout this project. We also extend our appreciation to our team members for their collaborative efforts. Special thanks to our college for providing the necessary resources. Lastly, we thank our families and friends for their encouragement and understanding during this project.

REFERENCES

[1] Min Jae Lee, Decoding Signals From Weather Satellites Using SDR, May 2018

[2] Ajith Kumar, Charan Yadav Np, Chandan M N,Pavan Kalyan Reddy Y, SDR Based Ground Station for Image Reception from Weather Satellites, May 2020

[3] Ubong Ukommi, Emmanuel UbomAnd Ikpaya Ikpaya, Ground Station Design for Satellite and Space Technology Development, July 2021

[4] A.M. Cailean, Cezar Lesanu, Mihai Dimian, Adrian done, Design and implementation of a satellite communication ground station, July 2017

[5] Priyanka S. Kamble, Bhalchandra B. Godbole, A review paper on Software Defined Radio, June 2016

[6] Romi Wiryadinata, Muhammad Khoirussolih, Neneng Rohanah, Imamul Muttakin, Teguh Firmansyah, Image Data Acquisition for NOAA 18 and NOAA 19 Weather Satellites Using QFH Antenna and RTL-SDR, ICIEE 2018

[7] George Slade, The Basics of Quadrifilar Helix Antennas, March 2015

❑ ❑ ❑ ❑ ❑ ❑

AET 046

Design and Implementation of Remote Labs

Ved Tikekar
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology Mumbai, India
ved.tikekar@vit.edu.in

Pooja Shunmugaraja
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology Mumbai, India
pooja.shunmugaraja@vit.edu.in

Saurabh Mehta
Professor & CAO
Vidyalankar Institute of Technology Mumbai, India
saurabh.mehta@vit.edu.in

Pratham Chavan
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology Mumbai, India
pratham.chavan20.1@vit.edu.in

***Abstract*— Remote laboratories can transcend geographical barriers and offer students immersive practical learning experiences. This paper highlights the pivotal role of remote laboratories in education, particularly amidst the challenges posed by the COVID-19 pandemic. Our research focuses on the development of three remote labs: a computational power lab for remote code execution, a temperature remote sensing and control lab and a remote robotics control lab. By ensuring 24/7 availability, these labs offer students flexibility in engaging with experiments and learning activities, regardless of their time zone. Furthermore, remote laboratories can be utilized to control and monitor equipment in remote locations and industries, offering unprecedented access and insights into real-world applications. This dual functionality not only ensures educational continuity during crises but also extends its value in various industrial settings.**

***Keywords*— Remote laboratories, Hands-on experimentation, Accessibility, Flexibility, Time zones, Experiments, Simulations, Raspberry Pi, Sensors, Practical application, Versatility.**

I. INTRODUCTION

In the realm of modern education, remote laboratories have emerged as transformative entities, reshaping conventional models of learning and experimentation. These virtual environments, also referred to as online or virtual laboratories, serve as digital platforms facilitating hands-on experimentation and practical learning experiences for students via the internet [1]. Unlike their traditional counterparts confined by physical constraints, remote labs transcend geographical boundaries, granting students unprecedented access to a diverse array of experiments and simulations from virtually any location with internet connectivity. There are three specific situations in which offering experimental work to students remotely can facilitate the provision of such work more easily [2]:

- When the students are studying away from the institution.
- When the equipment needed for the intended experimental work is deemed excessively costly.
- When it becomes challenging to accommodate a large number of students given the available lab space.

Remote laboratories herald a paradigm shift in educational pedagogy, offering a versatile and accessible platform for practical learning and experimentation. By harnessing technology's power, remote labs have the potential to revolutionize STEM education, empowering students to explore, innovate, and excel in an increasingly interconnected and digitized world. As educational institutions embrace digital transformation, the integration of remote laboratories will undoubtedly play a pivotal role in shaping the future of learning and preparing students for success in the 21st century.

Aspect	Remote Lab	Virtual Lab	Physical Lab
Advantages	- Accessible from anywhere	- Scalability	- Direct hands-on experience
	- Flexibility in scheduling	- Cost-effective	- Personalized guidance
	- Integration with online learning	- Simulations for hazardous experiments	- Opportunity for in-depth exploration
Disadvantages	- Lack of direct supervision	- Dependent on software availability	- Limited accessibility
	- Limited to certain types of experiments	- Limited physical interaction	- Maintenance overhead
	- Dependency on internet connectivity	- Potential for cheating	- Equipment wear and tear
Challenges	- Limited hands-on experience and accessibility	- Dependency on technology	- High cost of setup and maintenance
	- Lack of in-person supervision	- Network connectivity issues	- Limited scalability and space

Table 1. Comparison of Remote Lab, Virtual Lab, Physical Lab.

The problem of allocating the resources of one or more laboratories to a greater number of students practically defines the architecture. The simplest solution is a basic client-server architecture with one or more servers and numerous clients. A more serious issue that emerges from a closer look is how to guarantee that the resources will be distributed efficiently while preventing server overload and answering every query. This issue can be resolved by introducing a time slot booking system that ensures that there is no server overload by multiple requests.

II. PROPOSED SOLUTION

In this paper, we will illustrate the suggested methodology using a system identification experiment design practice as an example. We will describe this practice's hardware, tools, and methodology. While the primary focus of the paper is on the significance of remote labs in education, with particular emphasis on Raspberry Pi as a tool for implementation, we will also provide insights into its practical application.

Additionally, we have developed two remote labs as part of this study: 1. A computational power lab designed to facilitate remote code execution and 2. A remote temperature sensing and control lab aimed at monitoring and controlling environmental conditions using Raspberry Pi and sensors and, 3. Remote Robotics lab which enables users to control robotic movements remotely based on user input such as angle movements. These remote labs serve as practical demonstrations of the effectiveness and versatility of remote lab setups in educational contexts. The proposed solution is divided into three sections:

- Software: This section includes a description of the tech stack, pseudocode/workflow diagram, and the whole operation of the scheduling system for remote labs.
- Hardware: For reader's reference, we have provided a list of hardware components together with information on their purpose and functionality in remote labs, a schematic, and actual images of hardware settings.
- Results: This section presents screenshots of the computational lab's, remote temperature sensing and control lab's, and remote robotics lab's web interface results. Additionally, we included screenshots of the Thingspeak interface for comprehensive visualization of data and system operation.

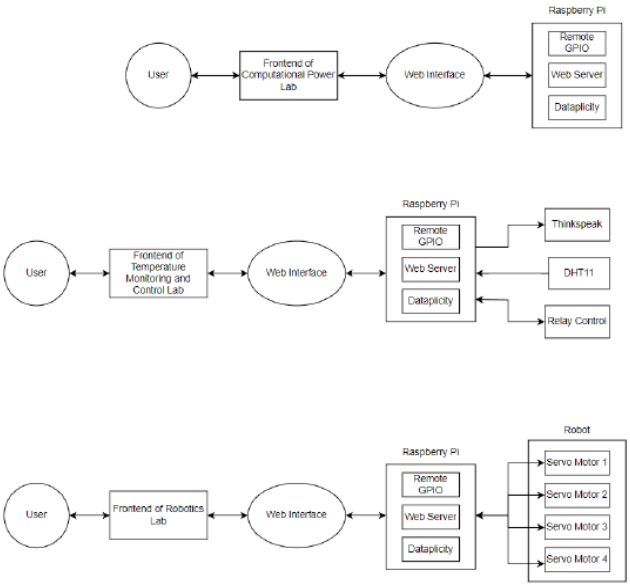


Fig. 1. Block Diagram of remote lab setup.

The Fig.1 describes the block diagram of our remote lab's setup where in there are various components such as Temperature sensors, servo motors, relay module which are mainly connected to raspberry pi as a central processing unit.

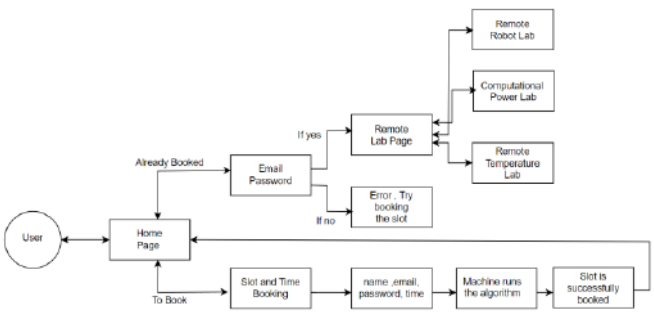


Fig. 2. System Architecture of of Booking system.

A. Software

In our setup, we have devised a sophisticated scheduling system using PHP, HTML, CSS and MySQL databases to streamline user access to our remote laboratory. This scheduling system enables users to reserve specific time slots for accessing the remote lab, enhancing organization and efficiency in lab usage. Integrated seamlessly into our software, the scheduling system presents users with a user-friendly calendar interface, allowing them to select their preferred time slots. Upon attempting to access the lab, the system validates the user's credentials against the current time slot's reservation. If the credentials match, the user is granted access and redirected to the remote lab running on a web server hosted on our Raspberry Pi otherwise it is redirected to the slot booking page as shown in Fig.2 .This stringent authentication process ensures that only one user can access the lab at any given time, mitigating the risk of conflicts or overuse. Furthermore, it guarantees undisturbed access for each user during their reserved time slot, thus maintaining fairness and equitable resource utilization.

Pseudocode for scheduling system :

1. Connect to a MySQL database.
2. If a date is selected:
   i) Fetch all bookings for that date.
3. If a booking is submitted:
   i) Check if the timeslot is already booked.
   ii) If not, insert the new booking into the database.
4. Define the timeslots for bookings.
5. Generate the timeslots based on the defined start time, end time, duration, and cleanup time.
6. If a 'submit' POST parameter is provided:
   i) Get the 'email' POST parameter.
   ii) Prepare a SQL statement to select all bookings for the provided email.
   iii) Execute the SQL statement.
   iv) If the execution is successful, get the result.
   v) If there are rows in the result:
      a) Check if the email matches the email in the database.
      b) If the emails match, check if the current date and time fall within the booked date and timeslot.
      c) If the current date and time fall within the booked date and timeslot, redirect to a specific URL.
      d) If not, set a message indicating that the current time is not within the booked timeslot or the dates do not match.
7. If there are no rows in the result, set a message indicating that the email does not match.

This pseudocode provides a high-level overview of your PHP scripts. They handle database connections, fetch bookings for a selected date, handle booking submissions, generate timeslots for bookings, and check if the current date and time fall within a booked date and timeslot for a provided email. If a timeslot is fully booked, it shows an "Already booked" message; otherwise, it inserts a new booking into the database. If the current date and time fall within a booked date and timeslot, it redirects to a specific URL; otherwise, it shows an appropriate error message.

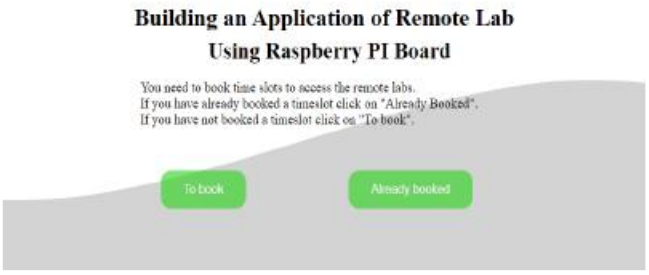


Fig. 3. Home Page of remote lab.



Fig. 4. Calendar page (current month, when clicked on "To book" button).



Fig. 5. Next month.



Fig. 6. Previous month.



Fig. 7. Choosing available date.

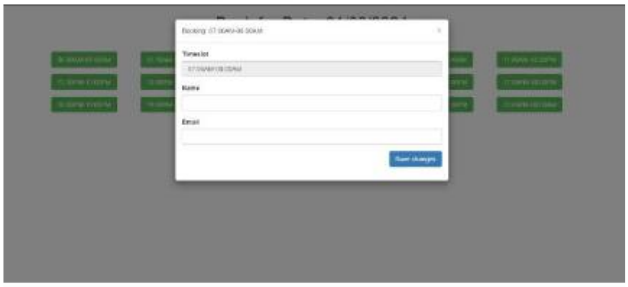


Fig. 8. Fill the form with appropriate details.



Fig. 9. Booked successfully.

B. Hardware

An extensive summary of the hardware parts used in the remote laboratory configuration is given in this section. The purpose of each component, including its connections, integration, and configuration, is explained in excellent detail to provide readers with a complete grasp of the Hardware configuration.

One of the most essential components is the Raspberry Pi 3 Model B+, which acts as an interface for peripheral connectivity as well as data processing which is common for all the labs as shown in Fig. 1. The temperature sensor (DHT11), which provides real-time data required for environmental control, is an integral component of the testing environment's ambient air temperature monitoring system. The 3.3V single-channel relay module also serves as an interface between the Raspberry Pi and external hardware, which permits the remote control of a wide range of gadgets, such as heaters, motors, and lights .With these components and the addition of remote GPIO capabilities and Adafruit library, we possess the flexibility to monitor and adjust environmental conditions from any location, offering enhanced control and convenience.

With its robust computing capabilities and GPIO support, the system is able to facilitate a wide array of programming tasks and hardware integration projects. Using the computational lab we can execute code on the Raspberry Pi terminal remotely from anywhere in the world by only accessing the remote lab. In our remote robotics lab, we have meticulously selected hardware components that work seamlessly together.

The Raspberry Pi acts like the central controller and is the computer part of our arrangement. It communicates with peripherals like servo motors in our case, which move the robot. To control these parts, we have utilized libraries such as RPi. GPIO for managing communication between raspberry Pi and peripherals and gpio zero to simplify motor control.By making sure every part is properly calibrated and configured, we build a system that is prepared for testing in education.

A three-step approach was implemented to create seamless remote management and access capabilities of the Raspberry Pi within the lab setting. This involved connecting the Raspberry Pi to a remote device management platform and configuring the device's software:

1. Raspberry Pi configuration To enable remote interaction with Raspberry Pi, we first activated key functionalities on hardware such as RealVNC for viewing and controlling the raspberry pi desktop from distance and Remote GPIO which enables the control of GPIO pins from remote location.

2. To streamline the management of remote device access, we opted for Dataplicity, a purpose-built platform designed to simplify this process. Beginning with the creation of an account, we proceeded to register our Raspberry Pi device onto the Dataplicity platform. This registration involved executing specific commands within the Raspberry Pi terminal, as outlined

comprehensively in the platform's documentation. By adhering to these instructions, we seamlessly integrated our device into Dataplicity's ecosystem, laying the groundwork for efficient remote access management.

3. Following device registration, we configured remote access to the Raspberry Pi's desktop, meticulously adjusting settings per Dataplicity's guidelines. Key to this was setting the accessibility port to 80, facilitating smooth desktop interaction and enabling seamless engagement with the remote laboratory on the device's web server. This adherence to Dataplicity's instructions established a robust remote access framework, empowering users to maximize the Raspberry Pi's potential.

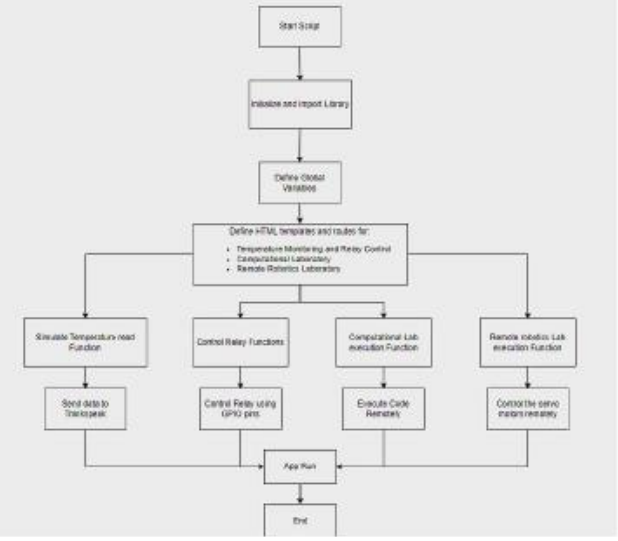


Fig. 10. Codeflow of remote lab running on Raspberry Pi.

Fig 10 shows the code flow of the Remote Lab indicating how the commands given by the user are executed on the server side. All the functionalities of the Laboratory are confined to predefined functions allowing users to interact with the laboratory system through a structured and organized framework. By encapsulating the functionalities within predefined functions, the Remote Lab ensures that users can easily understand and utilize its capabilities. This modular approach facilitates the efficient execution of user commands on the server side, enhancing the overall user experience and system reliability.

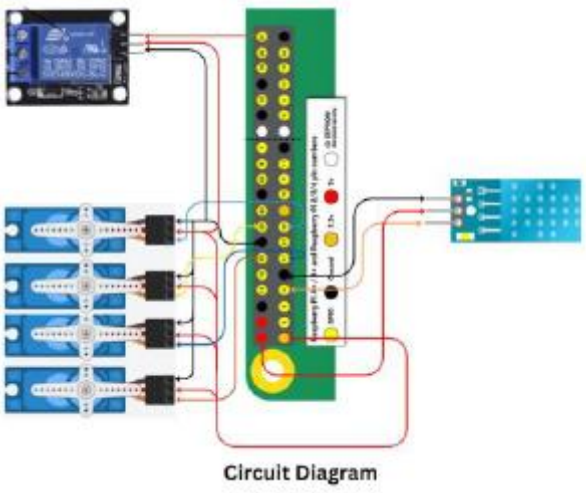


Fig. 11. Circuit Diagram.

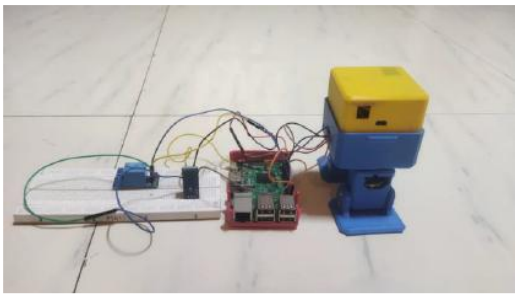


Fig. 12. Actual Hardware Connection.

C. Results

Lab Home Page :



Fig. 13. Home Page of remote lab.



Fig. 14. Computational code execution.

Lab 2 :

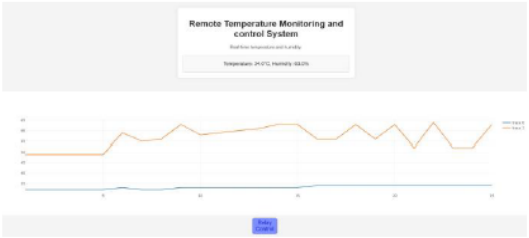


Fig. 15. Remote Temperature sensing and control.

Lab 3:



Fig. 16. Remote robotics lab interface.



Fig. 17. Thinkspeak interface.

As a result, we have developed three remote labs utilizing the Raspberry Pi Board. The first lab serves as a computational power lab which facilitates remote code execution as shown in Fig. 14 .Additionally, we established a remote Temperature Sensing and Control Lab in Fig. 15, where the relay module is activated if the temperature exceeds 28 degrees Celsius, allowing for automated temperature regulation. The integration of a Remote Robotics Lab enables users to remotely control robotic movements as per user's input as shown in Fig 16.

III. APPLICATIONS

Remote laboratories worldwide are proving to be effective in supporting various subjects like engineering, robotics, and physics, demonstrating the efficiency of an environment that mimics reality. This approach eliminates many problems associated with real laboratory operations. While this observation does not aim to eliminate local lab practices, it suggests the importance of new learning tools to increase course offerings with lower installation, operation, and maintenance costs. Astronomy, robotics, and materials science studies are all supported by remote labs. They make it possible to observe cosmic events, test control systems, and study the properties of material. Additionally, Engineers are trained by industries to operate industrial automation technologies, which improve productivity. Remote labs monitor environmental parameters in real time, aiding research and resource management. They offer flexible, engaging, and accessible hands-on learning experiences, enhancing science education and research.

Remote laboratories, such as the Virtual Instrument Systems in Reality (VISIR), Aquarium, Shake Table 2DOF, Turbulent to Laminar Flow, Truss Testing, and Virtual Machines, have been utilized in various projects. The first is a real laboratory connected to the internet, allowing students to build electronic experiments and verify signals remotely. The second is an underwater life verification in an aquarium, allowing users to feed fish, turn on lights, and watch movements remotely. [6]

In [7], after a brief overview of the latest developments used in the creation of remote laboratories and a breakdown of current, curious examples of remote laboratories in number of fields associated with industrial electronics education, it also discusses some current trends and challenges. A practical robotic situation can be interacted with remotely by the students. The Networked Control System Laboratory (NCSLab) is a remote laboratory that seeks to provide a range of experiment services to consumers distributed throughout the globe and incorporate many worldwide equipment for experiments [8].

IV. CONCLUSIONS

Remote labs are a promising solution for education and scientific exploration, offering accessibility, flexibility, and cost-effectiveness. They transcend geographical barriers, promoting inclusivity and equity in STEM education. Remote labs are used in various fields, from classrooms to university research laboratories, facilitating hands-on experimentation across disciplines like physics, chemistry, biology, and engineering. They enable collaborative learning, promote inquiry-based exploration, and provide a platform for real-world applications. Our contribution through the development and implementation of remote lab experiments underscores their importance in the education sector.

In summary, remote labs represent a vision of inclusive, experiential learning that transcends physical boundaries and empowers learners to explore the frontiers of science and technology. By embracing accessibility, collaboration, and innovation, educators can harness the transformative power of remote labs to inspire curiosity, drive discovery, and shape the future of education for generations to come. Recognizing that AI and machine learning algorithms could revolutionize data analysis and decision-making, the future of remote labs appears optimistic. Researchers can obtain deeper insights from the massive amounts of data collected by sensor networks by utilizing sophisticated computational techniques. Increasing the number of networks of sensors allows more thorough monitoring and more accurate measurement of a wider variety of environmental parameters. The integration of cloud computing assets improves accessibility, scalability, and durability by giving users on-demand access to processing and storage capacity, enabling easy data analysis, and remote collaboration.

REFERENCES

[1] M. Bower et al., "A conceptual framework for remote laboratories," Computers and Education, vol. 114, pp. 201-213, 2017.

[2] M. Cooper and J. M. M. Ferreira, "Remote Laboratories Extending Access to Science and Engineering Curricular," in IEEE Transactions on Learning Technologies, vol. 2, no. 4, pp. 342-353, Oct.-Dec. 2009.

[3] A. Almeida et al., "Remote laboratories in engineering education: A systematic review of empirical literature," Computers and Education, vol. 157, p. 103992, 2020.

[4] L. Zeng et al., "Design and implementation of a remote laboratory for power electronics experiments," IEEE Transactions on Industrial Electronics, vol. 66, no. 9, pp. 7337-7347, 2019.

[5] N. Zary et al., "Education in telemedicine and telehealth: Remote online case-based learning," Journal of Telemedicine and Telecare, vol. 22, no. 3, pp. 155-161, 2016.

[6] Luis C. M. Schlichting, Golberi de S. Ferreira, Daniel D. de Bona, Flavio de Faveri, Jose A. Anderson, "Remote Laboratory:´ Application and usability", Conference Paper · June 2016 DOI: 10.1109/TAEE.2016.7528355

[7] L. Gomes and S. Bogosyan, "Current trends in remote laboratories," IEEE Trans. Ind. Electron., vol. 56, no. 12, pp. 4744–4756, Dec. 2009.

[8] Y. Qiao, G.-P. Liu, G. Zheng, and W. Hu, "NCSLab: A web-based global-scale control laboratory with rich interactive features," IEEE Trans. Ind. Electron., vol. 57, no. 10, pp. 3253–3265, Oct. 2010

❑ ❑ ❑ ❑ ❑ ❑

AET 047

Design & Development of Multimodal Approach to Virtual Labs for Advanced Communication

Shravani Dorlikar
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
Mumbai, India. shravani.dorlikar@vit.edu.in

Udbhav Kamath
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
Mumbai, India. udbhav.kamath@vit.edu.in

Pratik Bhalake
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
Mumbai, India. pratik.bhalake@vit.edu.in

Adilya Phatkare
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
Mumbai, India. aditva.patkare@vit..edu.in

Prof. Pratik Mhatre
Department of Electronics and Telecommunication Engineering
Vidyalankar Institute of Technology
Mumbai, India. pratik.mhatrel@vit.edu.in

Prof.Sushopti Gawade
Department of Information Technology Engineering
Vidyalankar Institute of Technology
Mumbai, India.
sushopt.i.ga.wade@vitedu.in

***Abstract*—The paper presents a multimodal approach to virtual labs, enhancing advanced communication skills and understanding of digital communication and antennas. The virtual laboratory platform "VLABS" facilitates immersive learning experiences in analog and digital modulation techniques, along with antenna communication, offering flexible accessibility and personalized learning journeys. Emphasizing interactive engagement, the virtual labs bridge the gap between theoretical knowledge and practical application, fostering proficiency in communication domains. Real-time surveys with TE and BE engineering students revealed a 70-80% comprehension and engagement rate, underscoring the transformative potential of virtual labs in shaping dynamic educational landscapes.**

In summary, this study highlights how holistic virtual labs may revolutionize engineering education. Through the integration of theory, simulation, and interactive learning, these virtual labs enable students to get a more profound comprehension of intricate ideas while promoting cooperation and involvement in a virtual setting.

***Keywords*– interactive learning , accessibility, proficiency**

I. INTRODUCTION

The world of education has experienced a dramatic upheaval toward remote and virtual learning modalities under the extraordinary changes brought about by the COVID-19 epidemic.[l] Because of the restrictions or inaccessibility of traditional classroom settings, educational institutions and instructors around the world have resorted to creative alternatives to maintain learning continuity. With the ability to give students authentic laboratory experiences from the comfort and safety of their homes, virtual labs and remote learning platforms have become essential instruments in this shift.[3]

The pandemic's significant effects on education have brought attention to the pressing need for accessible and scalable alternatives to conventional laboratory settings.

Virtual laboratories are in high demand as educational institutions struggle with the problems of social distancing measures and remote instruction[3][4].

A fundamental shift in educational processes is seen in the stunning 300% increase in online learning enrollment rates above pre-pandemic levels, as evidenced by statistics from the COVID-19 era.[1][2]

A key project in addressing this increasing demand has been Virtual Labs (VLABS), which is being led by the MHRD Govt. of India. As a mainstay of engineering education, virtual labs (VLABS) offer students unmatched access to top-notch laboratory experiences, with over 700,000 independent uses and over 800 experiments accessible on the platform. Virtual laboratories have the potential to completely transform education, as seen by the continuous investment in virtual laboratory infrastructure by governments and educational institutions around the world. [5]

In this environment, the creation and execution of virtual laboratories for advanced communication, such as Signals and Systems, Antenna Technology, and Advanced Digital Communication Systems, becomes increasingly important. These labs provide students with a route to continuity, engagement, and academic success in the face of the pandemic's educational interruptions. They are a source of hope. These virtual labs promise to close the gap between theoretical knowledge and practical application by precisely customizing, adapting, and integrating assessment tools. This will ensure that students graduate with the skills and competencies required to prosper in an increasingly digital world[2] [3] [4].

It is impossible to overestimate the role that virtual laboratories will play in influencing education in the future as we negotiate the complexity of the post-pandemic educational landscape[3]. Virtual labs have the potential to reshape education and usher in a new era of creativity, accessibility, and excellence by utilizing technology to overcome geographical barriers, accommodate varied learning styles, and promote collaborative learning environments. [5] [2].

II. CHALLENGES

Challenges such as simulation accuracy, hardware limitations, UI design, experimental variability, and feedback and assessment procedures come up during the development of VLabs for antenna technology. While balancing technology restrictions to create realistic simulations, ensuring accurate simulations requires sophisticated modeling methodologies and validation

against real-world data. Usability requires designing interfaces that are easy to use and facilitate investigation of antenna systems. Robust simulation techniques are necessary to reproduce experimental heterogeneity in virtual laboratories, and the subjective character of antenna design makes it difficult to develop efficient evaluation and feedback mechanisms. In order to give students a real-world learning experience in virtual labs on antenna technology, these obstacles must be overcome using creative solutions.

III. SOLUTION APPROACH

An organized method designed to address particular elements essential to productive learning environments was used in the development of virtual laboratories for antenna technology. The project started with a thorough needs assessment and requirements analysis to guarantee that everyone understood the objectives and scope. Next, curriculum coordination was done to match material to learning outcomes. In order to provide interactive exercises and simulations that promote student engagement and comprehension, development relied on new libraries created with HTML, CSS, and JavaScript. In order to improve learning experiences, interactive components like simulations and quizzes were incorporated into user interface design, with an emphasis on intuitive interfaces. Iterative development allows for continual improvement based on input from users and stakeholders, and quality assurance procedures were strictly followed to ensure functionality and dependability. Even Bloom's Taxonomy helped in designing the organized approach and Through this methodical process, the virtual labs aimed to provide an engaging and dynamic learning platform conducive to student success in antenna technology.

Software and hardware: Hardware requirements include Compatibility with modern devices, minimal CPU and RAM for smooth performance, and standard input/display with internet connectivity. For development, software needs include VS Code for coding, Figma for design, Chrome/Firefox/Safari for testing, Git for version control, Node.js/npm for dependency management, and optionally, Photoshop/Illustrator for graphics.

VI. INSPIRATION & DESIGN IDEOLOGY

The COVID-19 epidemic presented significant hurdles for students, disrupting typical educational settings and limiting access to physical labs. These challenges served as inspiration for the establishment of virtual laboratories for digital communication and antenna communication. Despite these challenges, virtual lab simulations became apparent as a workable way to close the knowledge gap between theory and practice by giving students hands-on learning experiences. These simulations, which used cutting-edge technologies like HTML5, CSS3, and JavaScript, were created to be dynamic, immersive, and accessible. Their goal was to mimic real-world laboratory experiences.

The design concept placed a strong emphasis on efficacy, accessibility, and interaction. User-friendly interfaces and eye-catching images increased user engagement. While scalability and platform-agnostic design provided accessibility across many devices and environments, flexibility and adaptability were focused to support varying learning needs.

V. PROPOSED SOLUTION

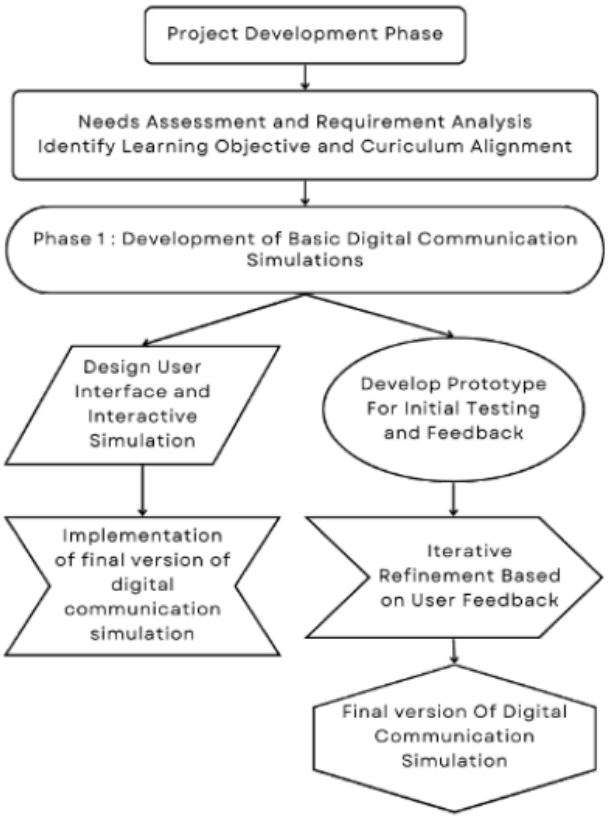


Fig 1: Phase-I development

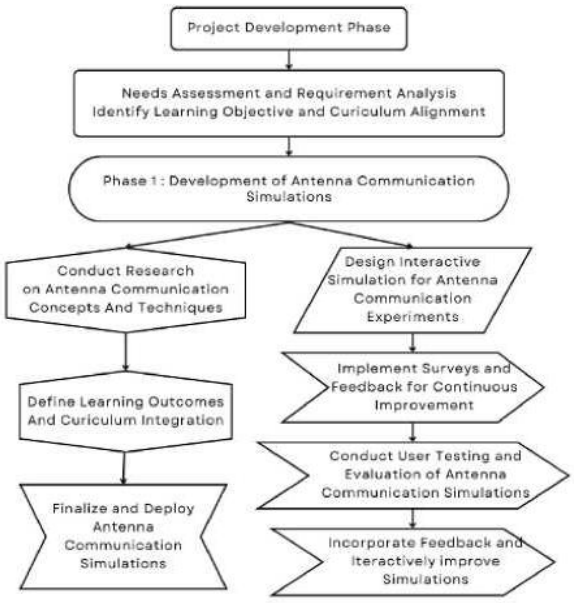


Fig 2: Phase-II development

The proposed solution involves creating virtual labs for digital and antenna communication, aiming to provide students with interactive learning experiences. Development phases ref fig 1 & fig 2 include building basic digital communication simulations like AM, FM, ASK, and FSK, followed by refining antenna communication simulations based on iterative feedback also a step by step process in each of the development phases included research, aligning curriculum to build relevant simulation, Conduct survey to enhance user experience and incorporate feedback in better development of VLABS.

Phase 1 : Development of Basic Digital Communication Simulations

Our first phase of virtual lab development ref fig 1 & fig 2 focuses on creating fundamental digital communication simulations, covering key modulation techniques such as amplitude modulation (AM), frequency modulation (FM), and digital modulation techniques such as amplitude shift keying (ASK) and frequency shift keying (FSK). This requires meticulously aligning curriculum goals, assessing needs, and iterative prototyping to provide intuitive user interfaces and effective knowledge transmission. Following completion, these simulations provide students with a firm foundation in digital communication fundamentals.

Phase 2: Development of Antenna Communication Simulations

The second Phase focuses on the creation of antenna communication simulations. Extensive research is being performed to better understand the concepts and procedures of antenna communication, ensuring that simulations accurately reflect real-world conditions. Learning outcomes are determined, and the curriculum is integrated to ensure that it is consistent with educational goals. Interactive simulations are designed to let students experiment with antenna designs, propagation characteristics, and signal transmission.

VI. METHODOLOGY

The section will explain the steps of methodology used for designing the antenna & digital communication. The discussion will cover key considerations such as the cognitive levels addressed by the experiments, Blooms Taxonomy for better output & the structuring of lessons. All of the above factors will be presented in a detailed, technical manner in the virtual experiment & also the flow of experiment is as follows

I. Interactive Introduction
II. Video Theory
III. Simulation & Performance
IV. Post Assessment Quiz
V. Chat Bot (Lab Bot).
VI. IMPLEMENTATION

VII. IMPLEMENTATION

TABLE I: COMPARATIVE ANALYSIS

Characteristics	Previous Digital Communication Experiment	Current Antenna Experiment
Realism	Simplified model of an AM System	Realistic model of antenna system
Complexity	Basic Concepts of AM	Antenna Design, Propagation, and Performance
Interactivity	Static Simulation	Interactive Simulation
Bloom's Taxology	Knowledge and Comprehension	Knowledge, Comprehension, application and analysis
Structural Pedagogy	Simple Instruction	Guided Inquiry

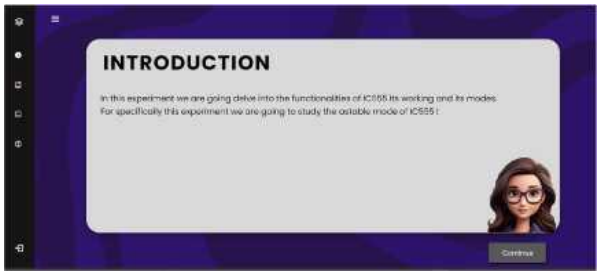


Fig 3: Interactive Introduction



Fig 4: Video Theory

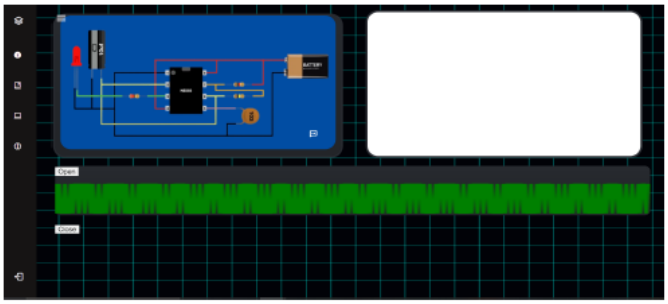


Fig 5: Simulation & Performance



Fig 6: Post Assessment Quiz

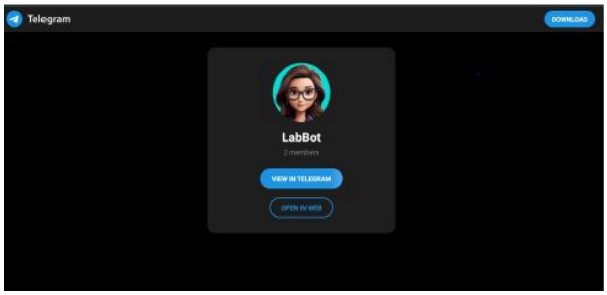


Fig 7: Chatbot (Labbot)

Ref. Table I. Two independent phases have been implemented to significantly improve the use of virtual labs for digital communication and antenna testing. The amplitude modulation (AM) systems models used in the early digital communication experiments were mainly

simplified versions that provide a fundamental grasp of the ideas. Realistic models that replicate the complexities of antenna design, propagation, and performance have been implemented, though, in the present phase of antenna experimentation, which has seen a trend towards realism.

In terms of intricacy, the present antenna experiments get into the intricacies of antenna design and performance evaluation, whereas the preceding digital communication experiments focused on basic ideas of AM. Students are now able to interact with more complex subjects because to this development, which develops their analytical and understanding abilities.

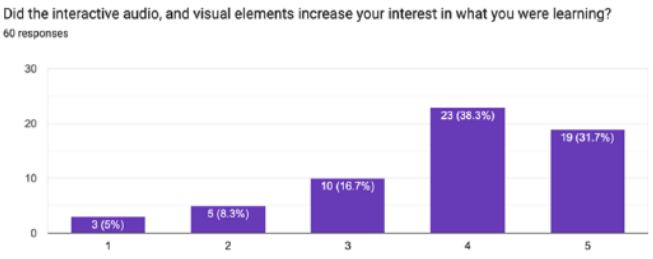
The shift from static to interactive simulations has resulted in a noticeable improvement in interaction. This improvement makes it possible for students to actively engage in the experimental process by changing parameters and seeing results in real time, which promotes a more engaging learning environment.

Furthermore, the current phase's use of Bloom's Taxonomy has broadened the scope of education. Unlike the previous trials, which focused mostly on knowledge and comprehension levels, the current phase encourages students to apply their theoretical knowledge to real-world situations through analysis and application.

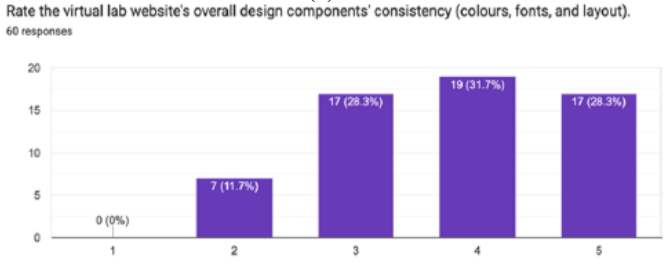
The virtual experiment's phases are all thoughtfully planned to accomplish particular goals. By acquainting students with the goals and methods of the experiment, the interactive introduction sets the scene. Comprehensive explanations of fundamental concepts are offered by the video theory fig 3, 4 which enhances the usefulness of simulations. ref. fig 5 Students can actively participate in the experiment during the simulation and performance phase by changing variables and seeing the results.

ref. fig 6, 7 While the chatbot acts as a helpful guide, offering support and clarification when needed, the post-assessment quiz assesses students' comprehension and recall of concepts.

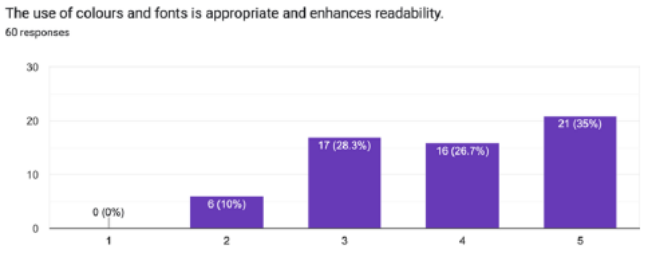
In conclusion, the learning process has been enhanced by the iterative development of the virtual lab implementation, which has given realistic simulations, fostered analytical abilities, and offered complete support during the experimentation phase.



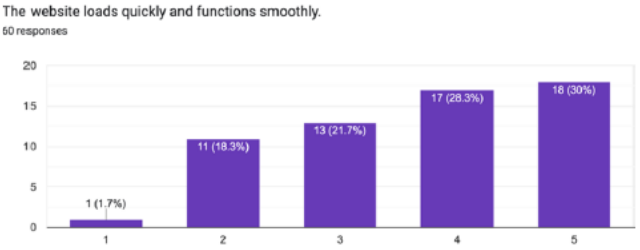
(a)



(b)

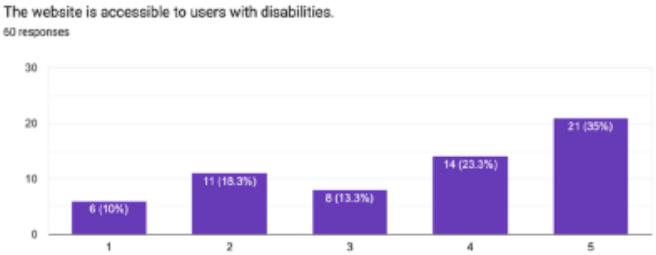


(c)

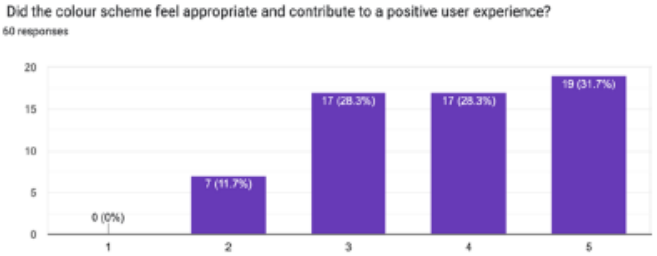


(d)

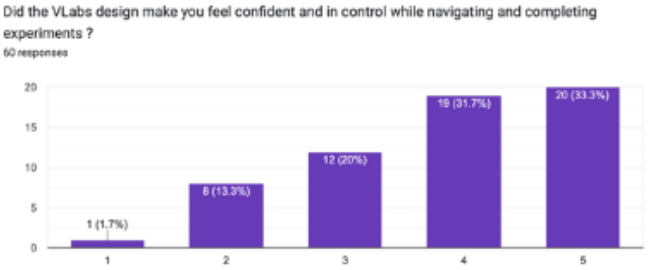
Fig 8 (a, b, c, d)



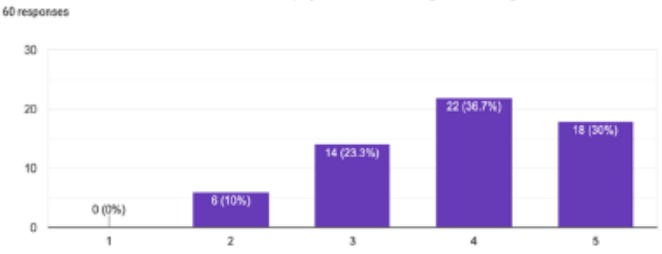
(e)



(f)

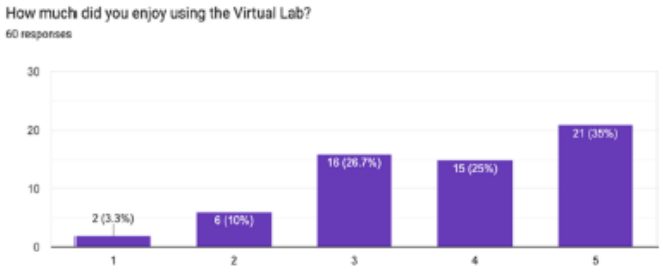


(g)

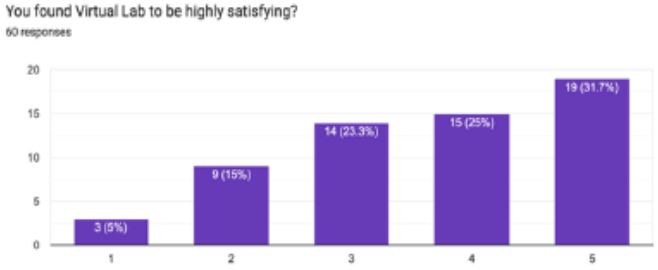


(h)

Fig 9 (e, f, g, h)



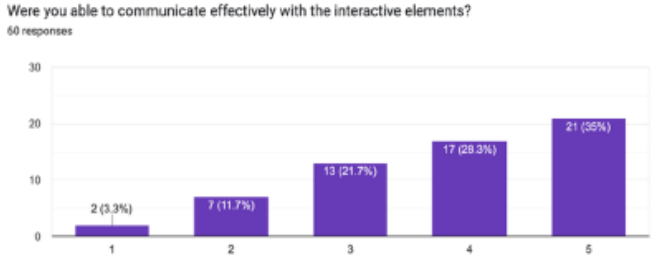
(i)



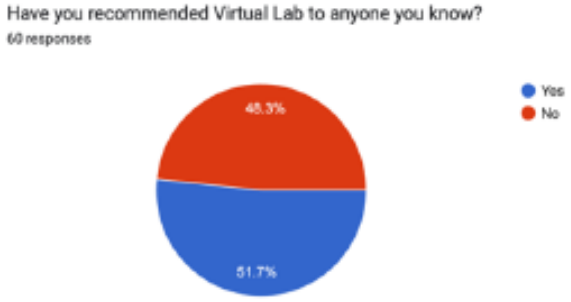
(j)



(k)



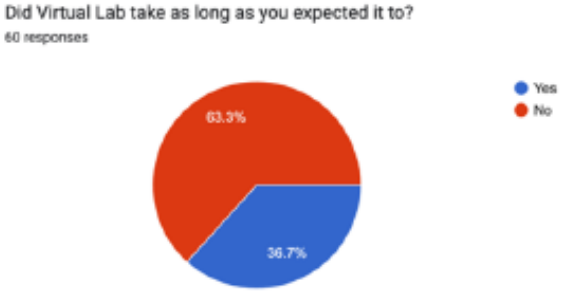
(l) Fig 10 (i, j, k, l)



(m)



(n)



(o)

Ref. Fig 8, 9, 10, 11 This section presents an in-depth analysis of user experience and design interaction in the context of virtual labs, leveraging real-time survey data. The statistical insights derived from these surveys are crucial in understanding user perceptions, identifying areas for improvement, and guiding the enhanced development of virtual labs. Real-time surveys were conducted to assess user perceptions of virtual labs' usability, interface design, content relevance, and learning effectiveness. Statistical analysis revealed:

1. Usability concerns were identified initially, with users citing navigation complexity and interface intuitiveness issues. Statistical analysis indicated an average usability rating of 3 out of 5, highlighting the need for navigation and interface design improvements.

2. Feedback on interface design varied, with some praising its simplicity while others suggested enhancements for better organization. Statistical analysis of interface design ratings provided insights into areas requiring refinement, such as layout structure and visual hierarchy.

3. Users expressed the need for more relevant and engaging content aligned with curriculum objectives. Statistical analysis of content relevance ratings highlighted areas requiring expansion or revision to enhance learning experiences.

4. While users valued the hands-on nature of virtual labs, some faced challenges in applying theoretical concepts practically. Statistical analysis of learning effectiveness ratings identified areas for optimization to better support learning outcomes and foster understanding.

XI. ANALYSIS OF THE VLABS TOOL VIA REAL-TIME SURVEY AND ASSESSMENT

Ref. Fig 8, 9, 10, 11 The real-time survey data offered valuable insights into user experiences and perceptions of virtual labs. Through rigorous statistical analysis, significant findings emerged, leading to actionable recommendations for enhancing virtual lab development:

1. Iterative Design Refinements: Utilizing user feedback and statistical insights, iterative design refinements were initiated to boost usability, streamline interface navigation, and enhance content relevance.
2. Enhanced Content Development: Insights from user surveys guided the creation of more pertinent and captivating content tailored to meet curriculum objectives and fulfill user needs effectively.
3. Interactive Feature Integration: Statistical analysis steered the incorporation of interactive features, including simulations, quizzes, and virtual experiments, to heighten user engagement and promote active learning.
4. Accessibility Enhancements: Measures were taken to ensure accessibility for all users by integrating features like keyboard navigation and screen reader compatibility, addressing accessibility issues identified through statistical analysis.

X. CONCLUSION AND FUTURE SCOPE

Virtual Laboratories for Signals and Systems, Antenna Technology, and Advanced Digital Communication Systems.

In an ever-changing educational landscape, our project launches a ground-breaking initiative: Virtual Laboratories for Signals and Systems, Antenna Technology, and Advanced Digital Communication Systems. These virtual laboratories are intended to bridge the gap between theoretical knowledge and practical application, resulting in an immersive and enjoyable learning experience. Key parts of the project process include needs assessment, technology selection, customisation, automated assessment, and pilot deployment. The ultimate goal is to provide students with an accessible and adaptive learning environment that exceeds the limitations of traditional laboratories. Our research paper contributes to education by enhancing learning outcomes, improving accessibility, and promoting cost-efficiency. Through this report, we aim to underscore the potential of virtual laboratories to transform education, providing students with the tools they need to excel in these vital subject areas.

REFERENCES

Here are the references for the research paper "Multimodal Approach to Virtual Labs for Advanced Communication":

[1] S. Gawade, U. Kulkarni, H. Palivela, S. Chopade, S. Chopade.“Student Learning Time Analysis during COVID-19 with Online and Offline mode. ”

[2] S. Tang, R. Tse, C. Cheong, C. Lam, P. Lei “Overcoming the Sudden Conversion to Online Education During the COVID-19 Pandemic: A Case Study in Computing Education.” IEEE December 2020.

[3] M. Kumar, J. Emory, V. Choppella. “Usability Analysis of Virtual Labs.” IEEE. July 2018

[4] S. Frerich, D. Kruse, M. Petermann, A. Kilzer. “Virtual Labs and Remote Labs: Practical experience for everyone.” IEEE. April 2014.

[5] O. Panasiuk, L. Akimova, O. Kuznietsova,I. Panasiuk. “Virtual Laboratories for Engineering Education.” IEEE. Sept. 2021.

[6] B. Aktan, C. A. Bohus, L. A. Crowl, and M. H. Shor, ‘‘Distance learning applied to control engineering laboratories,’’ IEEE Trans. on Education, vol. 39, Aug. 1996.

[7] O. Dziabenko, J. Garc´ıa-Zub´ıa (Eds.), IT Innovative Practices in Secondary Schools: Remote Experiments, University of Deusto, 2013.

[8] S. R. Singer, . L. Hilton, H. A. Schweingruber, America’s lab report: Investigations in high school science, Tech. rep., National Research Council, National Academy Press, Washington, DC (2006)

[9] J. Ma, J. V. Nickerson, Hands-on, simulated, and remote laboratories: A comparative literature review, ACM Computing Surveys 38 (3), Article 7.

[10] D. A. Kolb, Experiential learning: Experience as the source of learning and development, Prentice Hall, 1983.

[11] M. A. Prada, J. J. Fuertes, S. Alonso, S. Garca, M. Domnguez, Challenges and solutions in remote laboratories. application to a remote laboratory of an electro-pneumatic classification cell, Computers & Education 85 (2015)

[12] J. Saenz, J. Chacon, L. De La Torre, A. Visioli, S. Dormido, Open and low-cost virtual and remote labs on control engineering, IEEE Access 3 (2015)

[13] C. C. Ko, B. Chen, J. Chen, Y. Zhuang, K. Chen Tan, Development of a web-based laboratory for control experiments on a coupled tank apparatus, IEEE Transactions on Education 44 (1) (2001)

[14] D. Gillet, A. V. N. Ngoc, Y. Rekik, Collaborative web-based experimentation in flexible engineering education, IEEE Transactions on Education 48 (4) (2005)

[15] M. Casini, D. Prattichizzo, The automatic control telelab: a user-friendly interface for distance learning, IEEE Transactions on Education 46 (2003)

[16] ormido, R. Pastor, F. Morilla, A java/matlab-based environment for remote control system laboratories: illustrated with an inverted pendulum, IEEE Transactions on Education 47 (3) (2004)

[17] B. Balamuralithara, P. C. Woods, Virtual laboratories in engineering education: The simulation lab and remote lab, Computer Applications in Engineering Education 17 (1) (2009)

[18] M. Koretsky, D. Amatore, C. Barnes, S. Kimura, Enhancement of student learning in experimental design using a virtual laboratory, IEEE Transactions on Education 51 (1) (2008)

[19] J. E. Corter, J. V. Nickerson, S. K. Esche, C. Chassapis, S. Im, J. Ma, Constructing reality: A study of remote, hands-on, and simulated laboratories, ACM Transactions on Computer-Human Interaction 14 (2).

[20] A. Valera J. L. Diez, M. Valles, P. Albertos, Virtual and remote control laboratory development, IEEE Control Systems 25 (1) (2005).

[21] M. Ogot, G. Elliott, N. Glumac, An assessment of in-person and remotely operated laboratories, Journal of Engineering Education 92 (1) (2003).

[22] C. A. Jara, F. A. Candelas, F. Torres, S. Dormido, F. Esquembre, O. Reinoso, Real-time collaboration of virtual laboratories through the internet, Computers & Education 52 (1) (2009).

[23] O. Bellmunt, D. Miracle, S. Arellano, A. Sumper, A. Andreu, A distance plc programming course employing a remote laboratory based on a flexible manufacturing cell, IEEE Transactions on Education 49 (2) (2006).

[24] C. Colwell, E. Scanlon, M. Cooper, Using remote laboratories to extend access to science and engineering, Computers & Education 38 (13) (2002).

[25] Maffeis, J. Pereira, B. Russo, E. Cardozo, M. Bergerman, M. Magalhaes, Real: a virtual laboratory for mobile robot experiments, IEEE Transactions on Education 46 (1) (2003).

[26] A. Ferrero, V. Piuri, A simulation tool for virtual laboratory experiments in a www environment, IEEE Transactions on Instrumentation and Measurement 48 (3) (1999).

[27] C. Schmid, A remote laboratory using virtual reality on the web, Simulation 73 (1) (1999).

[28] C. R¨ohrig, A. Jochheim, The virtual lab for controlling real experiments via internet, in: International Symposium on Computer Aided Control System Design, Kohala Coast-Island of Hawai, USA, (1999).

[29] B. Barros, T. Read, M. F. Verdejo, Virtual collaborative experimentation: An approach combining remote and local labs, IEEE Transactions on Education 51 (2008).

[30] F. Schaf, C. Pereira, Integrating mixed-reality remote experiments into virtual learning environments using interchangeable components, IEEE Transactions on Industrial Electronics 56 (12) (2009).

[31] M. Garcia-Famoso, R. Moya, Towards the integration of remote laboratories into learning management systems, in: IEEE Conf. on Emerging Technologies and Factory Automation, Vol. 2, Catania, Italy, (2005).

[32] Z. C. Zacharia, G. Olympiou, Physical versus virtual manipulative experimentation in physics learning, Learning and Instruction 21 (3) (2011).

[33] E. Fabregas, G. Farias, S. Dormido-Canto, S. Dormido, F. Esquembre, Developing a remote laboratory for engineering education, Computers & Education 57 (2) (2011).

[34] M. Koretsky, C. Kelly, E. Gummer, Student perceptions of learning in the laboratory: Comparison of industrially situated virtual laboratories to capstone physical laboratories, Journal of Engineering Education 100 (3) (2011).

❑ ❑ ❑ ❑ ❑ ❑

Guidelines for Manuscript Submission

- Paper in IEEE two column format can be submitted for possible publication in the "Applied Engineering and Technologies" Journal of VIT. The template for the IEEE two column format is available at : http://www.vit.edu.in in Instructions to authors link

- The paper should be submitted in MS word format to editor.aet@vit.edu.in

- Authors of the accepted papers have to sign the copyright form which is available at: http://www.vit.edu.in Instructions to authors link

- For more information and queries refer http://www.vit.edu.in or email at: editor.aet@vit.edu.in

CONTENTS

Paper Id	Title	Page No.
AET016	LevelUp- A Smart Way Towards Safety	82
AET017	Computational Stellar Model	92
AET018	Mediplant: Medicinal Plant Identifying Platform	97
AET019	Yantrahasth: A Mechatronics Sorting Machine	102
AET020	Secure Collaborative Learning: Demonstrating Federated Next Word Prediction	106
AET021	'Abode Alchemy'- An Augmented Reality based Flutter Application	111
AET022	AiAuthor: AI-Powered Technical Review Paper Generator	116
AET023	YogaWise: Enhancing Yoga with Intelligent Real Time Tracking using TensorFlow MoveNet	122
AET024	TalkSpace - Advanced Emotion Recognition and Personality Analysis Platform	129
AET025	Mental Health Assistant using Sentiment Analysis Through NLP	135
AET026	GeoCause – Geolocation Event Finder	139
AET027	Sentiment Analysis for Hindi Movie Reviews	144
AET028	AI based Fashion Stylist Recommendation System	148
AET029	VitaCompass : Nutrition Tracker	154
AET030	Automatic Pill Dispenser	158
AET031	Patient Monitoring System for Comatose Patient	163
AET032	A Review of Erythrocyte Shape Detection under Normal and Pathological Conditions	167
AET033	Implementation of a Portable Integrated Biosensor System for Real-time Vital Sign Monitoring using IoT	171
AET034	Patient Information System (PIS)	175
AET035	Skin Cancer Detection using TensorFlow in Python	178
AET036	Automated Inventory Analyzer	183
AET037	Deep Learning-based Plant Identification App: Leveraging Efficient Net Model and Modern Web Technologies	187
AET038	EcoFlow: Enhanced Water Efficient Farming	192
AET039	Verification and Tracking of Humans in Campus	195
AET040	Automatic Medicine Dispenser using QR code	201
AET041	Identification of Chronic Wound and Prediction of Wound Size and Depth using Deep Learning Approach.	204
AET042	Smart Cradle for Baby	208
AET043	IoT based Chemiprobe for Surfactant Performance Modelling	212
AET044	A Muilti-Sensor Approach : Optimizing Tennis Ball Collection	217
AET045	SkyScan - Setting up a Satellite Ground Station	221
AET046	Design and Implementation of Remote Labs	225
AET047	Design and Development of Multimodal Approach to Virtual Labs for Advanced Communication	230