

Enhancing Mobility for the Visually Impaired using Artificial Intelligence and Machine Learning in IKSHANA

Dr. Raghu C.N¹, Dr. G. Raghavendra², Malashree G³, Gopinath K⁴, Dr. Pavithra G⁵ and Dr. T.C.Manjunath⁶

¹Director, School of Electrical & Electronics Engineering, REVA University, Bengaluru

²Associate Professor, Dept. of Electrical & Electronics Engg, Sapthagiri College of Engineering, Bengaluru

³Assistant Professor, Dept. of Electrical & Electronics Engg, Sapthagiri College of Engineering, Bengaluru & Research Scholar, School of Electrical & Electronics Engineering, REVA University, Bengaluru

⁴Assistant Professor, Dept. of Electrical & Electronics Engg, Sapthagiri College of Engineering, Bengaluru

⁵Associate Professor, Dept. of Electronics & Communication Engineering, Dayananda Sagar College of Engineering, Bangalore, Karnataka

⁶ Professor & Head, Dept. of Electronics & Communication Engineering, Dayananda Sagar College of Engineering, Bangalore, Karnataka

Corresponding Author Email: dimpu213@gmail.com

Abstract— This paper explores the utilization of Artificial Intelligence and Machine Learning algorithms to empower visually impaired individuals, enabling satisfactory movement in their daily lives. The research focuses on enhancing the lives of the visually impaired by addressing challenges they face in unexpected situations. In today's highly technological society, the need for independent living is crucial, especially for visually impaired individuals who often experience social exclusion detection of objects, facial based recognitions, emotion based recognitions, currency based recognitions, web type automations, screen readings & emergency callings, we enhance the quality of life for the visually impaired. Implemented through ML, DL, speech to text & vice-versa, Selenium based web type automation frameworks, these technologies empower individuals to effectively adapt and respond to their environment...

Index Terms— Artificial Intelligence, Machine Learning, Object, Detection, Emotion, Intelligence, Currency, Recognition, Face, DL.

I. INTRODUCTION

The field of AI involves the study & the developing of the computer based systems to have the ability for perceiving, reason, and take actions. This encompasses learning, decision-making, and problem-solving capabilities. AI technology has significantly improved the lives of visually impaired individuals, allowing them to lead lives comparable to those without visual impairments. Detection of objects, which could be treated as a interdisciplinary branch of computer visions, plays a crucial role by identifying semantic objects, like humans, building, automobile in photographs & in videos, providing a comprehensive understanding of the surroundings [1]. Facial recognition systems, when combined with object detection, not only detect individuals in frames but also identify them, offering enhanced awareness for the visually impaired [2]. Emotion recognition technology analyzes facial expressions from images and videos, providing insights into others' emotional states. This helps

visually impaired individuals understand the emotions of people around them and respond appropriately [3]. Web automation facilitates internet access for visually impaired users through speech commands, ensuring a streamlined and efficient browsing experience with minimal keystrokes [4]. The integration of a currency-recognizing module allows automatic identification of currency values, aiding in day-to-day transactions without human intervention. Screen readers, software that converts on-screen text into an audible format, further assist visually impaired users in processing information displayed on computer screens [5]. These technologies work collaboratively, utilizing input from a Bluetooth camera embedded in a hat worn by visually impaired individuals. This comprehensive system empowers them to navigate their surroundings, recognize individuals, interpret emotions, manage currency transactions, and access digital content seamlessly, ultimately fostering a life that closely resembles that of individuals without visual impairments [6] [4]

II. OBJECTIVES OF RESEARCH

The research's objective outlined in the 2nd section are categorized in to six stream, addressing crucial areas such as:

- Developing object detection embedded with facial recognition software to enhance the adaptability of the visually type impaired [6].
- Creating an Emotion detection model capable of identifying emotions of individuals in the visually impaired's frame [7].
- Developing currency recognition software to ensure secure transactions for the visually impaired [8].
- Automating web-surfing through voice type command thro' the visually impaired for a more effective and efficient browsing experience [9].
- Creating software to alert registered phone numbers during panic situations which are experienced by the visual type impaired persons [10].

III. LITERATURE REVIEW

Here, review of literature is presented. The developed technologies focuses on obstacle detections to alerting individuals, since it lacks information about object types and the environment. To address this limitation, we've integrated detection of objects using the embedded type of facial based recognitions. This enhancement provide details about the objects in the environment & which will identifies individuals, enabling visually impaired individuals to interact more effectively with their surroundings [11]. While the designed model detects the presence of an individual, it doesn't provide identity information. To overcome this, facial recognition has been embedded into object detection, enabling the identification of individuals present in the frame if their images are in our database [12]. Extending previous work on automating web browsing for Wikipedia and Gmail, our goal is to enhance the shopping experience for visually impaired individuals on platforms like Amazon. This involves voice command-driven operations without human intervention [13]. In their Android application, object detection and emotion intelligence software were implemented. Building on this work, we aim to introduce currency detection features to provide assurance about transactions for visually impaired individuals [14]. Expanding the capabilities further, we are working on incorporating an Emergency calling function through voice commands. This feature would be invaluable in helping visually impaired individuals contact necessary individuals during panic situations [15].

IV. OBJECT DETECTIONS USING MOBINET – SSD ARCHITECTURE

In our research, we utilized the COCO dataset, encompassing data related to 81 distinct objects, as the input for training the network. For RTI of detection of objects, we selected the Mobinet type of Single Shot's Detector's architecture. Then, process involves a series of steps, illustrated diagrammatically in Fig. 1, representing the architecture of Mobinet as a single shot detector [16].

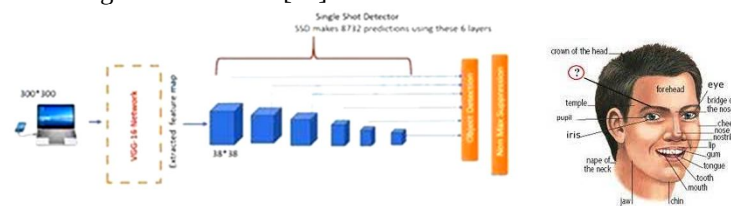


Fig. 1 : Block diagrammatic module of the proposed model & Face finding or identification using the image of the human face

V. BLOCK DIAGRAMMATIC MODULE OF THE PROPOSED MODEL

In our research, the chosen dataset is the COCO dataset, and our approach involves three key steps [17]:

1. **Training the Network (Step 1):** During the training phase, a training images, comprising of ground truth's box for every objects, serves as input to the VGG-16 network. This network includes convolution layers with 3x3 filters (stride 1, same padding) and a 2x2 Maxpooling layer (stride 2).
2. **The VGG-16 network produces extracted feature maps for each object. Generating Bounding Boxes (Step 2):** The extracted feature maps from the VGG-16 network are then inputted into six progressively decreasing convolutional layers. These layers are designed to produce 8735 bounded box for every detected objects, aligned using the ground truth's boxes of the trained images.
3. **Bounding Box Refinement (Step 3):** The output from the decreasing convolutional layers, representing 8732 bounding boxes with various aspect ratios, is refined into a solo bounded boxes for every objects. This refinement is achieved through a technique known as non-maximum suppressions, that is dependent on the intersections of the unions objective conceptions that are taken into consideration. The goal is to select the best-fitting bounding box for the provided training image among the candidates generated in the previous step [20].

VII. FACE FINDING OR IDENTIFICATION USING THE IMAGE OF THE HUMAN FACE

A. Emotional based Intelligence Use in the process of Design

The training processes for emotion recognition - This utilizes the FER2013 dataset, and the key steps in the emotion based recognition processes using the landmark is illustrated in Fig. No. 3. The 3 major steps considered in this process is as follows [25]:

B. Face Detection

This initial step focuses on detecting faces within the given data or images.

C. Feature Extraction

Once faces are detected, the system extracts relevant features from these facial landmarks. Landmarks could include points such as the eyes, nose, and mouth.

D. Feature Classification

The extracted features are then subjected to a classification process, where the system categorizes them based on predefined patterns or criteria. This step is crucial for determining the emotional expression portrayed in the facial features. These steps collectively contribute to the effective recognition of emotions using landmarks, enhancing the understanding and interpretation of emotional cues in the given dataset.

VIII. FACE FINDING OR IDENTIFICATION USING THE IMAGE OF THE HUMAN FACE

Step #1- Face Detections

In the initial steps, facial detection employs a model trained for facial recognition, utilizing the HOG approaches are used to identify faces.

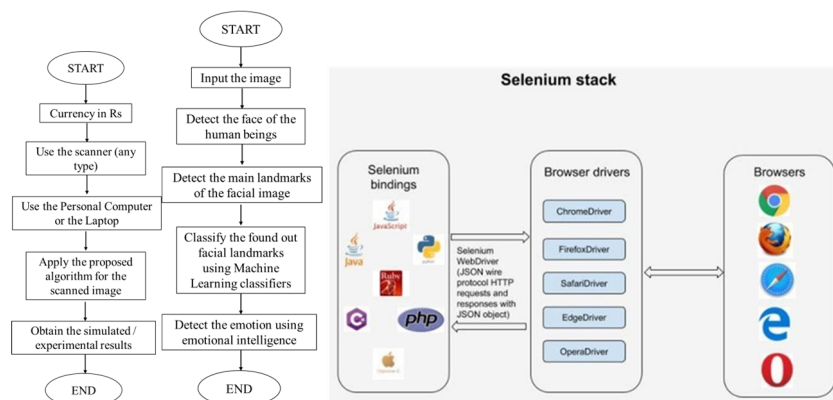


Fig. 4 : Processes involved in currency detection recognition steps & Stages in the emotion recognition procedure utilizing facial landmarks & Web automation using selenium stacks

Step #2 -: Feature based Extractions

In this stage, the process entails isolating landmarks from the identified faces. Utilizing Dlib's pre-trained Facial Landmark Detection Model, which identifies 68 2-Dimensional points on the human face, facial landmarks defining key features like eyes, brows, nose, mouth, and jawline are determined. Fig. 4 gives the stages in the emotion recognition procedure utilizing facial landmarks.

Step #3 - Feature based Classifications

Feature classification generates a list of the soft-max score for seven emotional categories. Then, user selects the primary emotion by choosing the emotions using the higher scores, found out by the calculations of the Euclidean distances from the entire search result.

IX. STAGES IN THE EMOTION RECOGNITION PROCEDURE UTILIZING FACIAL LANDMARKS

- *Currency Recognitions* - Then, currency recognition process involves a series of sequential steps designed to automatically identify the financial values of the money & facilitate conversion to other type of currencies w/o human interventions. Then, the steps are outlined below:
- *Image Input*: The process begins with providing an image of the currency to the currency recognition system.
- *Pre-processing*: The input image undergoes pre-processing to enhance its quality and optimize it for subsequent analysis.
- *Feature Extraction*: Relevant features, such as specific patterns, symbols, or characteristics unique to each currency, are extracted from the pre-processed image.
- *Pattern Recognition*: The system employs pattern recognition techniques to match the extracted features with known patterns associated with various currencies.
- *Monetary Value Determination*: Based on the recognized patterns, the system determines the monetary value of the currency in the provided image.
- *Currency Conversion (Optional)*: If required, the system can perform automatic currency conversion, providing the equivalent value in other currencies.
- Fig. 5 visually represents the steps involved in the currency recognition process, illustrating the flow of information and decision-making within the system..

X. STEPS IN IMAGE PROCESSING CONCEPTS

The imaging process for currency detection involves several steps, as illustrated in Fig. 7:

Image Pre-processing: This step aims to enhance important image features and suppress undesired distortions. It includes image adjusting and smoothening to remove noise.

Edge Detections & the ROI Determinations : Detecting of edges are employed to find region boundaries in the image analysis process. This step results in a binary image with black and white colors, separating the background and foreground and defining the Region of Interest (ROI).

Extracting the useful info from the R o I : The values of the HSV considering all the pels in the ROI are extracted, and the average of their H, S, V features is calculated. These values are then compared with the database, utilizing Euclidean distances. The currency value using the least Euclidean distances are displayed as the result.

Camera module utilized to capture input from visually impaired surroundings

Our aim is to employ a wireless webcam/camera for capturing inputs from the surroundings, serving as inputs for diverse applications like object detections, emotional recognitions, currency recognitions & text to speech conversion. Then, the outcomes are then delivered as voice outputs thro' Bluetooths earphone connected to the hats. The Fig. No. 8 illustrates the CCD cameras module that are embedded in the hats, facilitating wireless input's acquisition.

XI. RESULTS & DISCUSSIONS

This study demonstrates the successful implementation of an object detection model with integrated facial recognition using AI and ML methodologies. Figure 9 illustrates the model's accurate detection of various denominations of Indian currencies...

A. Detecting the Indian Money

This research showcases the successful integration of an object detection model with a facial recognition system. Using the coco dataset and the Mobinet-SSD architecture, our model accurately identifies 81 different objects

within the visually impaired individual's frame. The embedded facial recognition further distinguishes and identifies individuals if their data is stored in the database. The presented outcomes highlight the module's effectiveness in accurately recognizing various objects, including a persons, cell phones, and zebras. The trained model has been tested with various inputs, consistently delivering correct and precise results..

B. Exemplary result from the object detection model

In the following simulation step, we successfully implemented the emotional intelligence model. Utilizing the FER-2013 dataset, our model identifies seven emotions, including anger. The graph above illustrates the model accuracy and loss throughout the training process. This implementation naturally detects emotions on all faces in the webcam feed, achieving a test accuracy of 65% after 60 epochs using a simple 4-layer CNN..

C. Plot of the model accuracies & the model losses

In the third module, a successful implementation of a currency recognition model has been achieved to assist the visually based impaired in the daily transaction. Utilizing the random fores't algo, Scikit-learn, & OpenCVs, the model can accurately identifications of the denomination of Indian currency notes (10, 20, 50, 100, 200, 500, 2000). Moving on to the fourth module, a Web-automation framework has been implemented for Gmail & the shopper's stop website for the Amazon. This framework, developed utilizing the Selenium automation's tool, accepts voice commands as input. In the case of email automation, it sends emails to the designated recipient, and for shopping websites like Amazon, it locates a product based on user preferences and places an order. In the final module, an emergency based caller softwares has been developed utilizing the Twilio models in Python. This software is designed to call the authroized phone numbers when ever the visually type of impaired person perceives a panic situation, providing Immediate assistance when needed.

XII. CONCLUSIVE REMARKS

AI-ML-DL has significantly improved the lives of visually challenged individuals, addressing various challenges faced by disabled, autistic, elderly, blind, and other people in need of assistance. As technology continues to advance, there is a growing potential for machines to understand human emotions and problem-solving, offering promising prospects for the future. In our work, we have harnessed these technological advancements to empower visually impaired individuals, enabling them to lead more normal and independent lives. The integration of detection of objects using the face recognitions & the emotion type of intelligent softwares which provides an effective means for visually impaired individuals to comprehend their surroundings and respond appropriately. Currency recognition facilitates seamless transactions, while textual recognitions aids in reading & to analyze the texts. Then, the implementation of a virtual based assistants offers a user-friendly method for visually impaired individuals to access various websites, eliminating the need for complex keyboard shortcuts or extensive knowledge of screen reader usage. Moreover, enhancing processing performance and reducing data's inputs latency starting with the module of the camera to the generalized module of the camera model can enable higher frame resolution and frame rates, leading to more precise environmental perception.

REFERENCES

- [1] Shashank S., Kushal K., Abhay Surya Shankar, Madan Kumar G., Dr. Pavithra G., Dr. Sindhu Sree M., Padmavathy M., Dr. T.C. Manjunath, "RFID based attendance system with SMS Alert", Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 774 – 782, Jul. - Sept. 2023.
- [2] Abhishek, Sujith M.S., Jeevan D., P. Kamalesh, Dr. Sindhu Sree M., Dr. Pavithra G., Dr. T.C. Manjunath, "Design & Development of a Table Assisted Robotic Arm", Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 818 – 822, Jul. - Sept. 2023.
- [3] Apeksha U., Chithrashree G.S., Divya N.M., Shalmali S. Mankikar, Dr. Sindhu Sree M., Dr. Pavithra G., Dr. T.C. Manjunath, "Wireless LoRa Communication Between Two Arduino Uno for Military Application in Soldier Tracking", Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN : 1001- 4055, Vol. 44, No. 3, pp. 768 – 773, Jul. - Sept. 2023.
- [4] Akarsh Kesharwani, Ayush P. Chaudhary, Bhanu Pratap Singh, Ved Kumar, Padmavathi M., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, "A Study on Hand Motion Controlled Robotic Arm", Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 812 – 817, Jul. - Sept. 2023.
- [5] Leena Jeyakumar, Prerana Aithal, Vismitha R., Pradhan Aithal, Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, "Development of Smart Bridge – Automatic Height Increase When Floodings Take Place", Scopus Journal Q3, Schimago

Ranking SJR 2022 0.32, HIndex 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 763 – 767, Jul. - Sept. 2023.

- [6] Sushanthi Raj, Manohar R., Bhuvan G.S., Deepthi. S.R., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, “Design and Development of Obstruction Detection and Removal Bot”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 807 – 811, Jul. - Sept. 2023.
- [7] Anagha, Jhanavi M., Khushi M.S., Nithin Kumar S., Dr. Sindhu Sree M., Dr. Pavithra G., Dr. T.C. Manjunath, “Paralysed Patient Healthcare Monitoring Device”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 758 – 762, Jul. - Sept. 2023.
- [8] Ashmeet Singh, Harshit Goenka, Prakhar Sahu, Venkatesh L., Pamavathi M., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, “Development of an Automatic Fire Extinguisher”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 802 – 806, Jul. - Sept. 2023.
- [9] Vaishnavi Patil, Dr. Pavithra G., Dr. T.C. Manjunath, “Design, Development of a Diversified Implementation of a Supervisory Control And Data Acquisition based VLSI System (SCADA) framework Utilizing Microcontroller based Programmable Logic Controllers”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 879 – 890, Jul. - Sept. 2023.
- [10] Kavya P., Sanjana S., Harika, Teju R., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, “Design & Development of Drones Using Radio Frequency Controllers”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 707 – 710, Jul. - Sept. 2023.
- [11] S.G. Swathi, Aliya Bhandari, Srushti M. B. Kavya A., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, “Voice Control Robot – Design & Development for Various Domestic Applications”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 791 – 801, Jul. - Sept. 2023.
- [12] Vaishnavi Patil, Dr. Pavithra G., Dr. T.C. Manjunath, “Design of Smart Wheelchair For Disabled (Handicapped) Persons Using Real Time Embedded Systems & Internet Of Things Approach”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 871 – 878, Jul. - Sept. 2023.
- [13] Biswendu Biswas, Madan V.L., Rakesh B.S., Prathik Chandrapal, Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, “Development of remotely operated military purpose aerial vehicles”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 638 – 641, Jul. - Sept. 2023.
- [14] Akanksha Dash, Amrutha G., Krutika S. Ganpur, Sneha Chatter, Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, “Obstacle Avoiding Robotic Car Using Arduino with Bluetooth and Voice Control”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 783 – 790, Jul. - Sept. 2023.
- [15] Vaishnavi Patil, Dr. Pavithra G., Dr. T.C. Manjunath, “Simulation & design of a VLSI embedded system using Verilog Coding with Modelsim approach in FPGA scenarios for AI applications in automotive sectors”, Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology, ISSN: 1001-4055, Vol. 44, No. 3, pp. 862 – 870, Jul. - Sept. 2023.