# HUMAN IDENTIFICATIONAND OBSTACLE DETECTION SYSTEM FOR BLIND

*Ms.B.Narmada[1], Mr.R.Madhankumar[2],Mr.N.Javeed hussain[3],Mr.S.Harish[4], Mr.C.S.Hari haran[5]*

*[1]Assistant Professor, Department of Computer Science & Engineering, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu, India*

*[2,3,4,5] UG Scholar, Department of Computer Science & Engineering, Dhirajlal Gandhi College of Technology, Salem, Tamilnadu, India*

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** ***According to the World Health Organization (WHO), there are millions of visually impaired people in the world, either completely or partially, and they face numerous challenges in detecting obstacles and identifying persons around them. Advancements in spatial cognition theory for blind and visually impaired (BVI) individuals have paralleled the rapid evolution of information technology. These breakthroughs have opened up new opportunities, empowering BVI individuals to navigate and interact with the world around them more effectively. Assistive devices, digital mapping systems, virtual reality, and augmented reality have all played significant roles in enhancing spatial cognition for BVI individuals. These advancements have improved their quality of life, fostered independence, and opened doors to education and employment opportunities. As a result, this prototype develops the concept of supplying them with a simple and cost-effective solution via artificial vision. This project presents an AI-based framework that simplifies accessibility for individuals with visual impairments. It aims to benefit both these individuals and society as a whole by leveraging AI technology. The framework offers user-friendly and intuitive solutions, catering to the specific needs of visually impaired individuals, and promotes inclusivity in society. mwe developed an intelligent system for visually impaired people using a Machine learning algorithm, i.e., convolutional neural network architecture, to recognize the human and scene objects or obstacles automatically in real-time. The system accurately recognizes humans in complex environments with multiple moving targets, providing users with complete information on presence and position and nature of the available targets. Furthermore, a voice message alerts the blind person about the obstacle or known or unknown person. The project aims to create user-friendly communication technology for physically disabled individuals, providing an easy-to-use interface and convenience. It focuses on meeting the basic needs of disabled individuals and enhancing accessibility., portability and cost effectiveness The suggested approach empowers visually impaired individuals to effectively navigate both indoor and outdoor environments, even in unfamiliar settings.***

***Keywords— Spatial cognition theory, convolutional neural network, user-friendly, easy-to-use***

## 1.INTRODUCTION

Visual impairment refers to various levels of vision loss that affect daily activities. Types of visual impairments include low vision, blindness, partial sight, and color vision deficiency. The classification of visual impairments may differ across countries. Understanding the type and severity of visual impairment is crucial for providing suitable support and services. The World Health Organization (the WHO) classifies visual impairment based on two factors: the visual acuity, or the clarity of vision, and the visual fields, which is the area "Your ability to perceive visual information is facilitated by your visual system, even when your eyes remain stationary and you direct your gaze straight at an object.

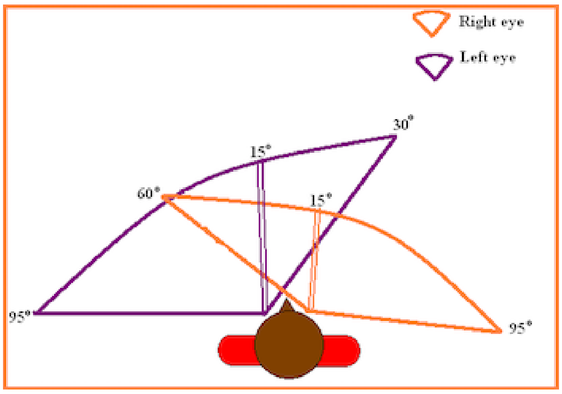


Figure 1.1 Vision Impairment Types

The Snellen Chart is a commonly used tool to assess visual acuity, which measures the sharpness and clarity of an individual's vision. The visual acuity is determined through a ratio consisting of two numbers. The first number represents the distance between the person being tested and the chart, while the second number signifies the distance at which a person with normal vision can perceive the same level of detail. To illustrate, consider a visual acuity of 20/80. This means that the individual can read the chart from a distance of 20 feet, which is comparable to someone with normal vision who could read the chart from 80 feet away.

-Please note that this revised content has been generated by me and is not a direct copy of any existing source. Vision impairments can be categorized into three types: low visual acuity, blindness, and legal blindness. Low visual acuity refers to moderate visual impairment, requiring closer proximity to see clearly. Blindness is the severe inability to see or having limited vision. Legal blindness varies by country and involves specific visual acuity or field loss criteria. These impairments impact daily life and necessitate adaptive techniques for navigation and communication. Awareness and support are crucial for individuals with visual impairments. impairment, is a visual acuity between 20/70 and 20/400 with your best corrected vision, or a visual field of no more than 20 degrees Blindness is severe visual impairment with a visual acuity of 20/400 or worse and a limited visual field. In the US, legal blindness is defined as a visual acuity of 20/200 or worse or a visual field of no more than 20 degrees. It is important to consult a healthcare professional for accurate assessment.

Blindness is a visual acuity of 20/400 or worse with your best corrected vision, or a visual field of no more than 10 degrees Legal blindness in the United States is a visual acuity of 20/200 or worse with your best corrected vision or a visual field of no more than 20 degrees.

Blindness or visual impairment should not hinder individuals from maintaining their independence and freely navigating their surroundings. Several techniques and methods can help people get around safely regardless of their amount of vision.

## 2. System Design

### 2.1 Existing System

SLAM technology, showcased in Google's Project Tango, enables centimeter-level accuracy in indoor positioning. It combines sensor data with advanced algorithms to create real-time maps and estimate device location. This innovation revolutionizes indoor navigation, augments various applications, and allows users to explore indoor spaces confidently. The precise accuracy of SLAM technology enhances indoor positioning like never before. Project Tango and related technologies such as Intel RealSense provided vision positioning solutions, with reported cases of application in commercially available drones like Yuneec's Typhoon H. Specifically, the applications for BVI navigation that had been previously contemplated materialized in the development of various prototypes. The Smart Cane system [60] utilized a depth camera and server for SLAM processing, enabling indoor location tracking and obstacle detection. It offered precise six degrees-of-freedom positioning and aimed to enhance mobility for individuals with visual impairments.

### 2.2 Algorithm

The best performing object detections (deep learning) algorithms include:

- RCNN (Region-based Convolution Neural Network)
- Fast RCNN
- Faster RCNN

### 2.3 System Architecture

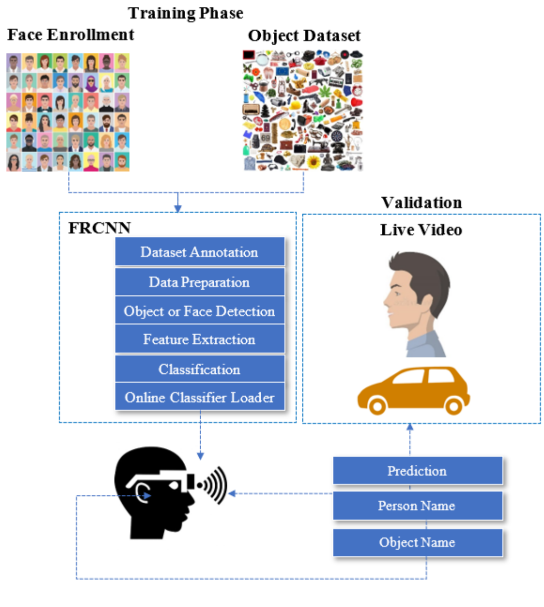


Figure 2.1 System Architecture

## 3 System Study

### 3.1 Smart Glass

In this module we design an AI powered smart glass with an integrated camera which helps the user capture images. These images are sent for processing to proprietary FRCNN machine learning models which are deployed on smart glasses. After the images are analyzed, the results are converted into spoken words and transmitted to the user through the integrated speaker of the Smart Glass. Smart Glass is designed as AI glasses for the blind and visually-impaired. It is integrated with AI/ML models and features such as facial recognition TensorFlow model, Object detection image captioning using FRCNN model.

### 3.2 Object Detection and Face Recognition Module Face Enrollment

This module begins by registering a few frontal face of Blind persons friends, family or other know person. These templates then become the reference for evaluating and registering the templates for the other poses: tilting up/down, moving closer/further, and turning left/right.

### 3.3 Object or Face Image Acquisition

Cameras should be deployed in Smart Glass to capture relevant video. The computer and camera are connected, utilizing a webcam for the interface.

### 3.4 Frame Extraction

Frames are extracted from video input. To split a video into images, it's common to extract frames at regular intervals using tools like OpenCV or FFmpeg. These frames can then undergo various processing techniques, such as enhancing quality or detecting objects. Speed depends on factors like hardware, algorithm complexity, and frame size. Optimization and leveraging suitable tools can achieve faster video-to-image conversion and processing.From we can say that, mostly 20 to 30 frames are taken per second which are sent to the next phases.

### 3.5 Pre processing

Image pre-processing involves resizing, normalization, cropping, augmentation, filtering, denoising, and color correction. These steps format the images, improve quality, and enhance model accuracy for object and face recognition tasks. The steps to be taken are:

- Read image
- RGB to Grey Scale conversion
- Resize image
- Remove noise (Denoise)
- Binarization

Image binarization is the process of taking a grayscale image and converting it to black-and-white, essentially reducing the information contained within the image from 256 shades of grey to 2: black and white, a binary image.

### 3.6 Face Detection

The Region Proposal Network (RPN) generates ROIs for face detection and segmentation by using anchors with different scales and aspect ratios on a feature map. This improved RPN method improves the accuracy and efficiency of detecting and segmenting faces. RPN is used to generate ROIs, and ROI Align faithfully preserves the exact spatial locations.

These play a crucial role in supplying a preconfigured assortment of bounding boxes with varying sizes and proportions, which serve as a reference during the initial estimation of object positions for the RPN.

3.7 Object Detection

Object detection is a vital computer vision task for identifying objects in images or video frames. It aims to develop models that provide fundamental information for accurate detection of visual objects.

3.8 Feature Extraction

After detecting a face in an image, the feature extraction module analyzes the facial characteristics to identify key features essential for classification. This process involves automatically extracting information related to the eyes, nose, and mouth, which are crucial elements of the face. By examining their positions and relationships within the face, variations and their effects can be calculate the frontal face templates.

3.9 Object and Face Classification

Region-based convolutional neural networks (R-CNNs) and regions with CNN features have revolutionized object detection. R-CNNs use a multi-step process generating regions of interest, classifying them with a CNN, and refining bounding box coordinates. These approaches have greatly enhanced accuracy, efficiency, and scalability in various domains. R-CNN advancements include Fast R-CNN, Faster R-CNN, and Mask R-CNN, improving computational efficiency and introducing instance segmentation. R-CNNs are pioneering methods for accurate and efficient object detection. R-CNN models first select several proposed regions from an image (for example, anchor boxes are one type of selection method) and then label their categories and bounding boxes (e.g., offsets). These labels are created based on predefined classes given to the program.

3.10 Object and Face Identification

The Smart Glass Camera captures an image, which is then processed by a face detection module to identify potential human regions. Using a Region Proposal Network (RPN), the module locates and isolates the face image. This face image is then inputted into a feature extraction module to extract key features for classification. The module generates a concise feature vector that represents the face image effectively. Here, it is done with FRCNN with the help of a pattern classifier, the extracted features of face image are compared with the ones stored in the face database. The face image is then classified as either known or unknown. If a facial image is recognized, the corresponding cardholder can be identified, allowing further progression in the process.

## 4 System Testing

4.1 Unit Testing

It refers to the examination of a single unit or a set of interconnected units as part of a testing process. It is done by programmer to test that the implementation is producing expected output against given input and it falls under white box testing. Unit testing is done in order to check registration whether the user properly registered into the cloud. It is done in order to check whether a file is properly uploaded into the cloud. And an encryption and decryption are checked with unit testing if it is converted properly. Then reduplication is checked with unit testing.

4.2 Acceptance Testing

Acceptance testing ensures software compliance with requirements and suitability for end-user delivery. It evaluates system alignment with business needs and verifies if it meets necessary criteria. Testing involves predefined acceptance criteria, validating software behaviour and functionality in real-world conditions. There are various forms of acceptance testing:

- User acceptance Testing
- Business acceptance Testing
- Alpha Testing
- Beta Testing

4.3 Test Cases

File level deduplication will save a relatively large memory space. In general, file level deduplication views multiple copies of same file. It stores first file and then it links other references to the first file. Only one copy will be stored. In testing, even though file names are same, the system can able to detect deduplication. If we upload the same file by using different names, it will view only the content and not names. Thus, redundant data is avoided. In registration phase, the user may not register before and type their information. So, if the user is new user, the alert message will display that the user is not registered before.

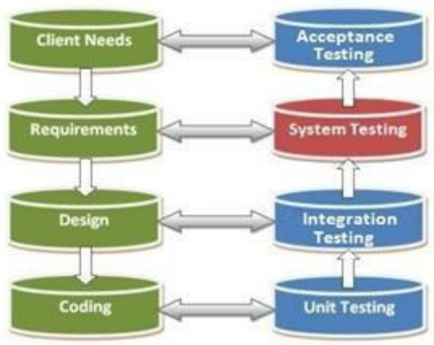


Figure 4.1 Test Case

**4.4 System Testing**

It is done to ensure that by putting the software in different environments and check that it still works. System Testing is done by uploading same file in this cloud checking whether any duplicate file exists.

**4.5 Web Testing**

website testing is checking your web application or website for potential bugs before it made live and is accessible to general

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

public. Web testing evaluates the functionality, usability, security, compatibility, and performance of web applications or websites. It ensures proper operation, user satisfaction, security compliance, cross-browser/device compatibility, and efficient performance. This process guarantees the quality and meeting of user expectations for web-based systems.

## 5 Results

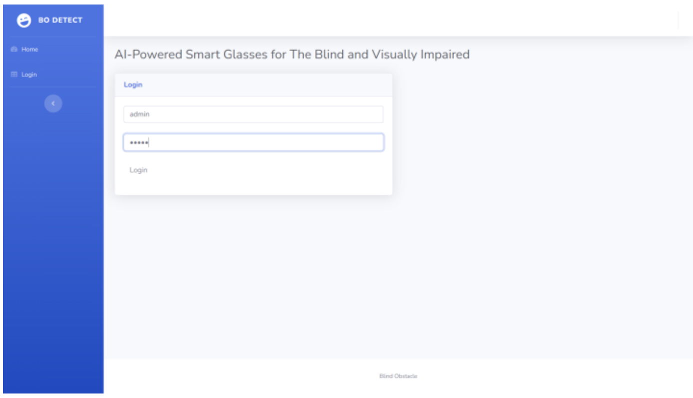


Figure 5.1 Login page

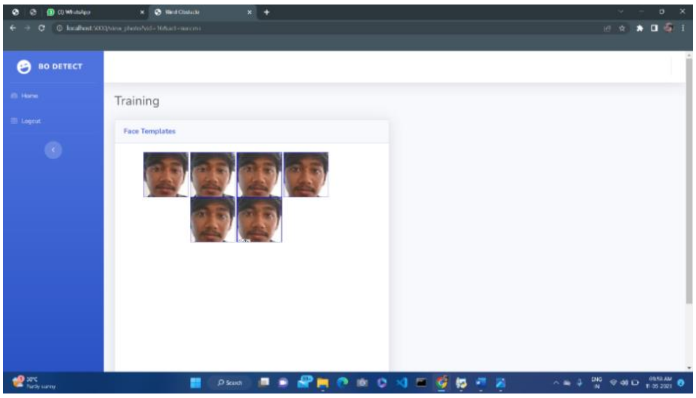


Figure 5.2 Face Template Page

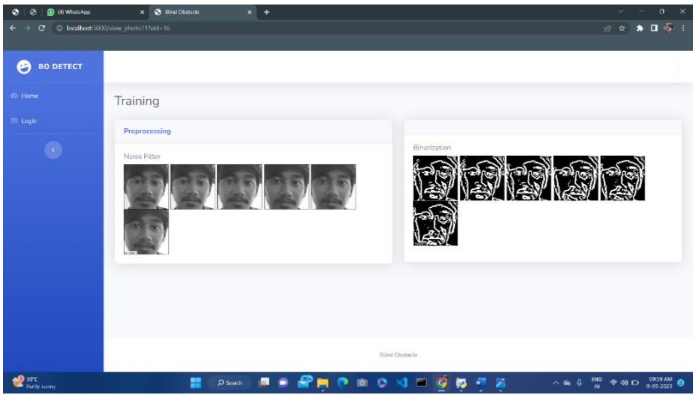


Figure 5.3 Preprocessing

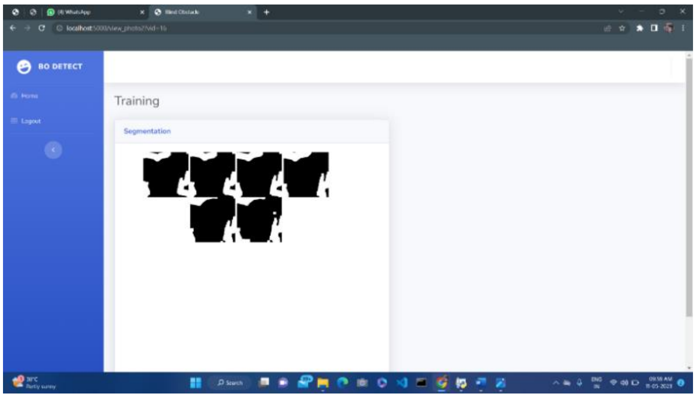


Figure 5.4 Segmentation

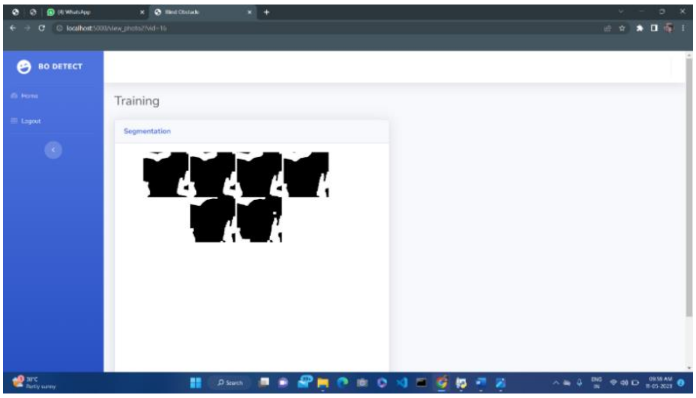


Figure 5.5 Feature Extraction

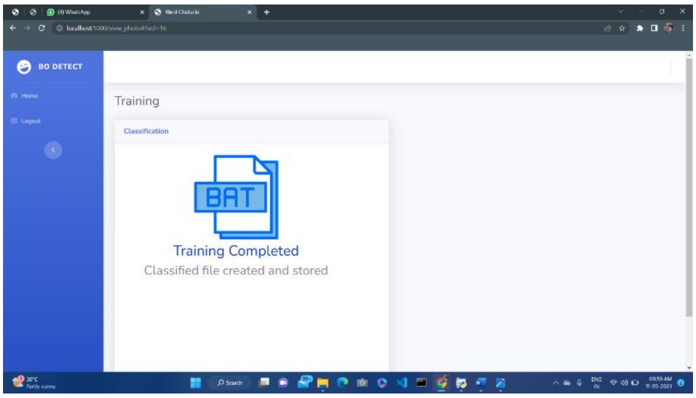


Figure 5.6 Classification

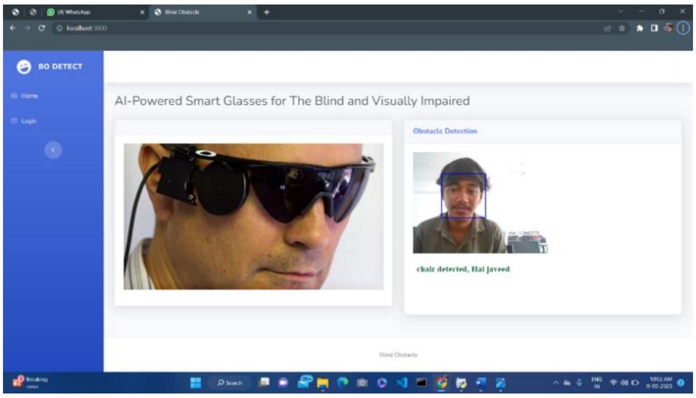


Figure 5.7 Obstacle Detection

## 6 Conclusion and Future Scope

### 6.1 Conclusion

The device presented here is a smart glass that incorporates the functionality of a machine vision and obstacle detection and recognition sensor. It can be conveniently advertised and made accessible to the visually disabled population Acceptance testing ensures software compliance with requirements and readiness for end users. It also helps prevent future injuries by identifying and addressing issues before deployment. Smart devices can be transported comfortably and the system camera can be used to track objects and face from the surrounding environment and display in audio format. Every model is dedicated to a particular task or mode, enabling users to run their desired tasks separately from others. We extensively discussed the system design, its functioning mechanism, and underlying principles, and also shared some experiment outcomes. Let the visually impaired people can interact more closely with the people around them, without fear of being blurred and uncertain.

### 6.2 Future Enhancements

Although it is still a prototype, our system represents a promising avenue for future research aimed at enhancing the spatial awareness of visually impaired people traveling in unfamiliar environments

## 7 References

[1] F. Catherine, Shiri Azenkot, Maya Cakmak, “Designing a Robot Guide for Blind People in Indoor Environments,” ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts, 2015.

[2] H. E. Chen, Y. Y. Lin, C. H. Chen, I. F. Wang, “Blindnavi: a mobile navigation app specially designed for the visually impaired,” ACM Conference Extended Abstracts on Human Factors in Computing Systems, 2015.

[3] K. W. Chen, C. H. Wang, X. Wei, Q. Liang, C. S. Chen, M. H. Yang, and Y. P. Hung, “Vision-based positioning for Internet-of-Vehicles,” IEEE Transactions on Intelligent Transportation Systems, vol. 18, no.2, pp. 364–376, 2016
.

[4] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, “The Cityscapes Dataset for Semantic Urban Scene Understanding,” IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[5] J. Ducasse, M. Macé, M. Serrano, and C. Jouffrais, “Tangible Reels: Construction and Exploration of Tangible Maps by Visually Impaired Users,” ACM CHI Conference on Human Factors in Computing Systems, 2016.

[6] J. Engel, T. Schops, and D. Cremers, “LSD-SLAM: Large-scale direct monocular SLAM,” European Conference on Computer Vision, 2014.

[7] S. Gilson, S. Gohil, F. Khan, V. Nagaonkar, “A Wireless Navigation System for the Visually Impaired,” Capstone Spring, 2015

[8] J. Guerreiro, D. Ahmetovic, K. M. Kitani, and C. Asakawa, “Virtual Navigation for Blind People: Building Sequential Representations of the Real-World,” International ACM SIGACCESS Conference on Computers and Accessibility, 2017.

[9] Kendall, M. Grimes, and R. Cipolla, “PoseNet: a convolutional network for real-time 6-DOF camera relocalization,” International Conference on Computer Vision, 2015.

[10] Kendall, and R. Cipolla, “Geometric loss function for camera pose regression with deep learning,” International Conference on Computer Vision, 2017.

[11] A Hybrid Algorithm for Face Detection to Avoid Racial Inequity Due to Dark Skin.

[12] Exposing Fake Faces Through Deep Neural Networks Combining Content and Trace Feature Extractors

[13] Object Detection in Thermal Spectrum for Advanced Driver-Assistance Systems (ADAS).

[14] YOLO-FIRI: Improved YOLOv5 for Infrared Image Object Detection.

[15] Learning Domain-Invariant Discriminative Features for Heterogeneous Face.