IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.12572

IDENTIFYING THE OBJECT AND OBSTACLE DETECTION FOR BLIND PEOPLES

Sneha S[1], Vijayalaxmi S[2], Maheswari M[3], Dr. Roselin Mary S[4]

Student, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India[1]

Student, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India[2]

Assistant Professor, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India[3]

Professor, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India[4]

Abstract: Sight and touch are the basic sensory systems for human interaction with the environment. For blind amputees, one of the key challenges is how to quickly and intuitively convey information about the environment to restore their daily life abilities. Inspired by the ability of human auditory localization, we constructed a virtual scene almost identical to reality and at the same time added a virtual sound source to the interactive object. Using the spatial sound rendering (SAR) method, the three-dimensional movement of a virtual sound source can be simulated live in real time. Finally, a myoelectric prosthetic control system was developed to assist blind amputees in their daily activities. The Fitts' law test for target localization was performed on both SAR and voice guidance (VP) guidance methods, the results indicate that SAR significantly improves the information transfer rate. Prosthetic control test results show that SAR reduces the completion time by half compared to VP while restoring the natural grasp path. With the advantage of intuitive and rich perception, SAR demonstrated potential applications for blind amputees to reconstruct control and sensory loops.

Keywords: Object and Obstacle detection, Convolutional Neural Network (CNN)

I.INTRODUCTION

The EYES and hands of humans are indispensable sensory systems for learning, adapting and transforming the environment. In amputated upper limbs, their missing motor function can be partially restored by wearing a myoelectric prosthetic hand. However, most commercial products have neglected the haptic functions of the robotic hand because users can rely on their vision as a haptic substitute. So once both haptic and visual perception is lost, it will bring great challenges to their daily life. In military or high-risk occupations, there is a high incidence of blindness and amputation of the hand , the Krukenberg procedure was an alternative technique to surgical separation of the radius and ulna of the distal forearm, and patients can control the movement of the radius and ulna like a pincer. Meanwhile, the skin surrounding the two bones realizes the tactile perception of the surroundings. Decades of effort have brought great progress in the field of intelligent bionic prosthetic hand: a multimodal information fusion method based on computer vision, inertial and eye tracking has been proposed to build a visual perception substitution to aid in manipulation. prosthetic hands. With the rich perception of the RGBD camera and the excellent recognition of deep learning, the computer vision method has improved the grasping accuracy in daily life scenarios. However, these studies focused mainly on improving the machine's autonomous dexterity and lacked environmental feedback for the user.

An assistive environmental perception system for amputees , where a person can interact with the machine or directly control it and perform actions on the environment. Meanwhile, the environmental information perceived by the machine is not only used to autonomously control the system, but also transmitted to blind people. The machine includes three sub-components an observer unit for sensing the environment, including a camera module (as a visual substitute) and a haptic sensory module (integrated in the prosthetic hand);effector unit for direct action on the environment (i.e. myoelectrically controlled prosthetic hand);a feedback unit that transmits interactive responses to a person. In order to recreate the closed-loop perception of the environment, the researchers used Google Glass as observers, and visual augmented reality (AR), voice prompts, and vibrotactile coding methods were used. serves to transfer sensory information from computer vision to the user. However, visual AR is not effective for the blind, and other feedback strategies transform visual information into a sequence of vibrations or voice, which greatly reduces the information transfer rate (ITR) and requires the user to learn preset feedback methods. In order to achieve much faster and intuitive sensory feedback, recent studies have tried to encode machine information into sequences of electrical pulses that directly stimulate the sensory nervous system in an invasive or non-invasive form, which can generate more phantom sensations such as pressure, tingling. and vibrations. However, these researches mainly focused on creating haptic feedback of a robotic hand. Suppose a blind amputee is using his prosthetic hand to drink a cup of tea, so the first stage is to approach

DOI: 10.17148/IJARCCE.2023.12572

the cup, which is a big challenge without visual aid. In general, the ITR of voice instructions is not sufficient to convey accurate orientation information, an alternative method is to use wearable devices to generate force/tactile stimuli for orientation. However, disabled people may not easily carry too many complicated devices by themselves. More importantly, this method creates an indirect sensory substitute that requires a long training period to build a sensory mapping.

In addition to visual localization, our auditory system can help determine the location of a target. As one can imagine, if it were possible to render 3D sound from the position of the cup, it could easily locate the cup with its auditory perception to the source of the sound. Inspired by this idea, we designed a virtual scene that reconstructed the spatial information of the real environment. For an interactive object in reality (eg a cup), a sound source has been added to the corresponding virtual object. We used a method called spatial sound rendering (SAR) that generated 3D surround sound based on the virtual scene, giving the user the illusion that the sound is coming from the object. As a result, these virtual auditory cues acted as an intuitive orientation guide to help the user locate the real object. A proposed sensory replacement based on surround sound offers a new way to help blind amputees regain their natural grasp. Compared with the same method of hearing-based (VP) voice prompts, SAR greatly improved the information transfer rate and grasping efficiency. More importantly, SAR allows active perception for the user, allowing for greater control robustness.

II.RELATED WORKS

“A wearable obstacle detection system for blind individuals using deep learning” developed by Guo et al. (2018). The system utilized a camera to capture real-time images, which were processed by a deep learning model to identify obstacles. The results showed promising potential for assisting visually impaired people in navigating their surroundings [1]. “A deep learning-based object detection system for visually impaired individuals using wearable devices” developed by Hu et al. (2020). The system employed a lightweight deep learning model to detect and classify objects in real-time. The research demonstrated the potential of using wearable devices and deep learning techniques to enhance object recognition for the visually impaired [2]. “Object detection and classification system for visually impaired individuals using deep learning techniques” developed by Ramalingam and Swarup (2019). The system employed convolutional neural networks to detect and classify objects in real-time. The research highlighted the potential of deep learning for assisting the visually impaired in object recognition and classification tasks [3].” A deep learning-based object detection system designed to assist visually impaired individuals” developed by Gupta, Sethi, and Chhabra (2019). The system utilized a convolutional neural network to detect objects in real-time from a video stream. The research demonstrated the potential of deep learning techniques in enabling object detection for the visually impaired [4].” Object detection system for visually impaired individuals using convolutional neural networks (CNNs)” developed by Islam, Rahman, Ananna, and Siddique (2020). The system utilized a trained CNN model to detect and classify objects in real-time. The research demonstrated the effectiveness of CNNs in facilitating object detection for the visually impaired [5].” Obstacle detection and alert system for visually impaired individuals using ultrasonic sensors” developed by Saini, Chaudhary, Chaudhary, and Mohan (2019). The system detected obstacles within a certain range and provided audio alerts to the user. The research highlighted the effectiveness of ultrasonic sensors in assisting the visually impaired in obstacle detection [6]. “An efficient object detection approach for the visually impaired based on deep learning” developed by Ren, Zhang, Wang, and Wu (2020). The system utilized a lightweight deep learning model to detect objects in real-time. The research focused on optimizing the object detection process to enhance its efficiency and applicability for visually impaired individuals [7]. “A real-time obstacle detection and tracking system for blind people” developed by Shrestha (2021). The system utilized computer vision techniques to detect and track obstacles in the environment. The research focused on providing real-time feedback to blind individuals to aid in their navigation and improve their safety [8]. “Vision-based obstacle detection system for visually impaired individuals using deep learning” developed by Jain (2020). The system utilized a deep learning model to detect obstacles in real-time using image processing techniques. The research emphasized the potential of vision-based approaches in assisting the visually impaired with obstacle detection [9]. “Obstacle detection and avoidance system for visually impaired people using deep learning techniques” developed by Mohammed and Ismail (2021). The system utilized a deep learning model to detect obstacles in real-time and provided audio feedback to guide the user in avoiding obstacles. The research aimed to enhance the safety and mobility of visually impaired individuals[10].

III.EXISTING SYSTEM

- **Kinect-based navigation system**

The navigation system described here requires a standard Kinect sensor, battery and laptop/processor. All this can be carried by the user in a backpack or shoulder bag. The Kinect is powered by a 12 V, 7 A sealed lead-acid battery whose output is fed through a DC-DC converter to ensure a stable 12 V power supply. The battery capacity is enough to power both the Kinect and the portable CPU for 3 hours.

DOI: 10.17148/IJARCCE.2023.12572

- **Smart wand**

The Simultaneous Localization and Mapping (SLAM) technology found in Google's Project Tango enabled indoor positioning with centimeter-level accuracy. Related technologies such as Project Tango and Intel RealSense provide vision positioning solutions with reported applications in commercial drones such as Yuneec's Typhoon H. The smart cane system [60] used a depth camera and a server for SLAM processing, which enabled indoor positioning and obstacle detection capability with 6 degrees of freedom.

- **Virtual touch radar**

This is a typical example that originates from tactile radar. It replaced its previous generation infrared sensor with a combination of a 3D model of its surroundings and a user-applied ultrasound-based motion capture system. Accordingly, when the user reached a certain area near the object, a warning vibration was generated.

- **Moovit**

It is a free, effective and easy-to-use tool that provides guidance on public transport networks and manages schedules, notifications and even alerts in real time. Mobility tasks recommended by ONCE (Spanish National Organization for the Blind) are one of the assets

- **Blind square**

Designed specifically for the BVI, this app shows the relative location of previously registered POIs. We use Foursquare and OpenStreetMap databases.

- **Lazas**

Also designed for BVI users, this is a paid app that combines GPS with built-in motion capture and orientation sensors to provide users with visual clues as to the location of various POIs in the surrounding area, including crossings. . It offers two working modes.

IV.PROPOSED SYSTEM

We have successfully attained the design and manufacture smart electronic glasses that help visually impaired people to recognize faces and objects. Face recognition and object recognition are computer technologies related to computer vision and image processing that deal with identifying instances of semantic objects of a certain class (humans, buildings, cars, etc.) in digital images and videos. Face recognition and object recognition has been one of the areas of great success. It is used in many fields such as face recognition (to recognize people on Facebook), tumor detection (used in medical field). Deep learning in computer vision has made tasks like object recognition relatively easy and efficient.Deep learning models are more accurate, less time-consuming, more complex and have better overall performance than previous computer vision methods. Deep learning has outperformed traditional computer vision object recognition methods, and deep learning models have been widely used.Smart glasses are designed to recognize faces and objects for blind people.The faster R-CNN model is used to perform object detection and classification.CNN is used to identify known people.A rain assistant that provides voice feedback about what you're looking at.

V.IMPLEMENTATION

Object recognition and face recognition module:

- **Face registration:**

This module begins by registering a few faces in front of friends, family members and other acquaintances of visually impaired people. These patterns serve as a reference for evaluating and registering patterns for other modes: tilt up/down, move closer/farther, turn left/right.

- **Obtaining object or face images:**

A camera must be placed on the Smart Glass to capture the corresponding video. A computer and a camera are connected, here a webcam is used.

- **Frame extraction:**

Frames are extracted from the input video. A video must be divided into a series of images that are further processed. The speed at which video should be split into images depends on individual implementations. That is, in most cases 20-30 frames per second are captured and sent to the next phase.

DOI: 10.17148/IJARCCE.2023.12572

- **Preprocessing:**

Preprocessing of an object or face image is a method used to format the image before using it in model training and inference. Here are the steps you need to follow.

- **Image reading**
- **Convert RGB to grayscale**
- **Change image size original size (360, 480, 3) – (width, height, number of RGB channels) resize (220, 220, 3)**
- **noise removal (noise reduction) Smoothen the image to remove unwanted noise. We do this using Gaussian blur.**

- **Binarization:**

Image binarization is the process of taking a grayscale image and converting it to black and white, which essentially reduces the information in the image from 256 shades of gray to a binary 2:black and white image.

- **Face recognition:**

Therefore, in this module, the Region Proposed Network (RPN) generates RoI by sliding windows on the feature map through anchors with different scales and different ratios. An improved RPN-based face recognition and segmentation method. RPN is used to generate RoI and RoIAlign faithfully maintains the exact position. They serve to provide a predefined set of bounding boxes of different sizes and aspect ratios that are used for reference when initially predicting the position of objects in the RPN.

- **Object detection:**

Object recognition is an important computer vision task that is used to identify examples of visual objects of a certain class (humans, animals, cars, buildings, etc.) in digital images such as photos and video frames. Object recognition aims to develop computational models that provide the most basic information needed for computer vision applications.

- **Feature extraction:**

After face recognition, the face images are given as input to the feature extraction module to identify the key features used for classification. For each pose, facial information including eyes, nose and mouth is automatically extracted and used in calculations. Variation effect using relation to frontal face pattern.

- **Object and face classification:**

Region-based convolutional neural networks or region-based CNNs (R-CNN) are pioneering approaches to apply deep models to object recognition. The R-CNN model first selects some proposed regions from the image (e.g., anchor boxes are one type of selection method) and then labels their categories and bounding boxes (e.g., offset). These tags are created based on predefined classes given to the program. We then use a convolutional neural network to perform forward computations to extract features from each proposed region. In R-CNN, the input image is first divided into approximately 2000 regions, and then a convolutional neural network is applied to each region. The area size is calculated and the correct area is entered into the neural network. It can be inferred that using such a precise method creates time constraints. The training time is significantly longer compared to YOLO because we classify and create the bounding boxes individually and apply the neural network one region at a time. In 2015, Fast R-CNN was developed with the aim of significantly reducing the training time. While the original R-CNN computational neural network is specified separately for each of the 2000 regions of interest, Fast R-CNN runs the neural network once for the entire image. At the bottom of the grid is a new method called Region of Interest (ROI) integration. It extracts each region of interest from the output tensor of the network, transforms and classifies it. This makes Fast R-CNN more accurate than original R-CNN. However, this detection technique requires less data input to train Fast R-CNN and R-CNN detectors.

System Architecture

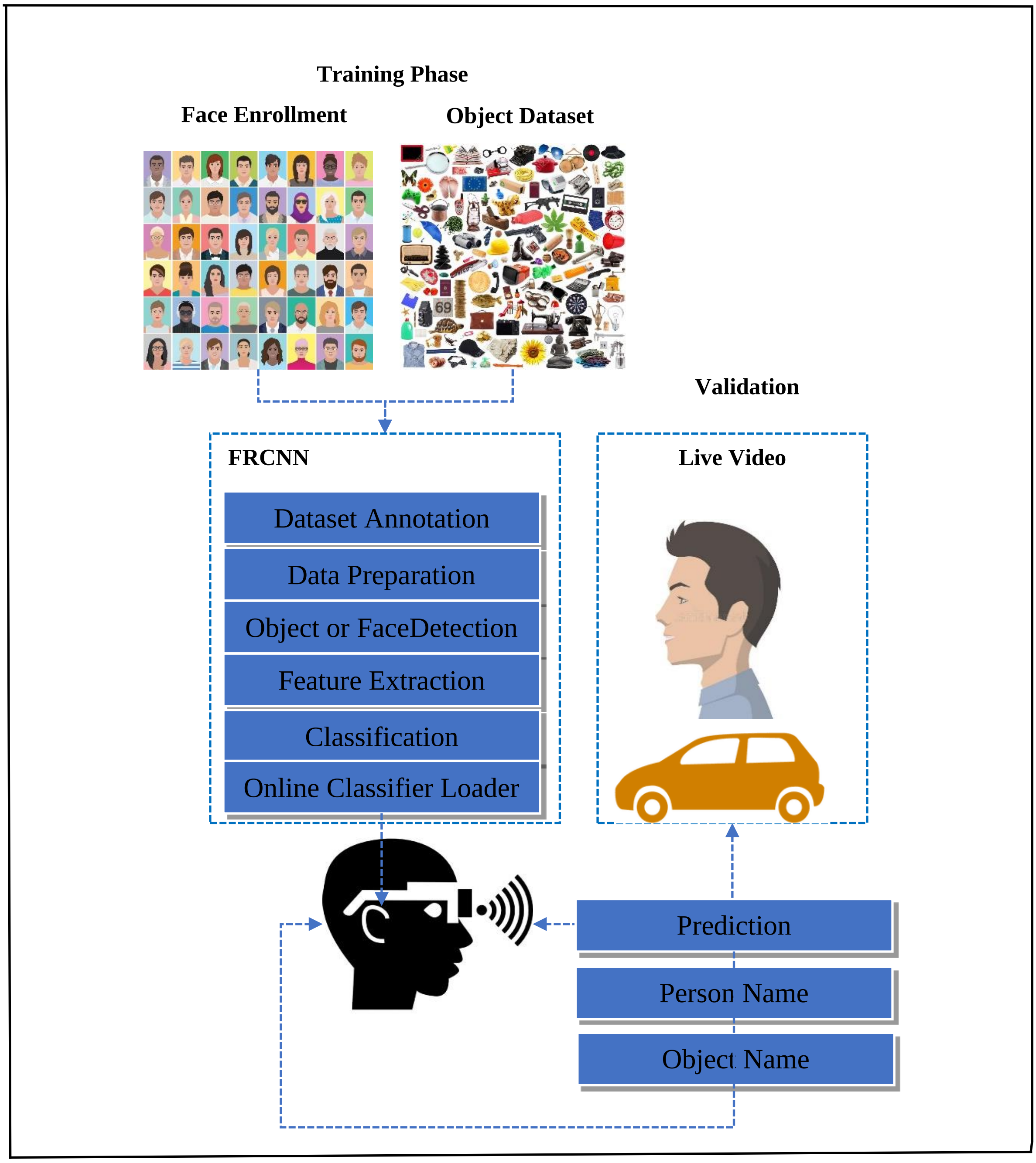


Fig.1 System Architecture

VI.RESULTS AND DISCUSSION

Recognition of objects and obstacles for blind people using machine learning can greatly improve their independence and mobility. Several approaches can be used to achieve this goal, including computer vision, deep learning, and artificial intelligence. This answer discusses some of the implications and limitations of these methods. One common approach to object and obstacle detection is to use computer vision algorithms to analyze images and video feeds captured by cameras and other sensors. These algorithms can detect and track objects such as people, cars, and animals, and can also detect obstacles such as curbs, stairs, and other ground obstacles. However, the performance of these algorithms can be limited by factors such as lighting conditions, camera placement, and scene complexity. The specificity of the network was found to be 90%, while the accuracy was 75%. Deep learning is another method that can be used to improve object and obstacle

DOI: 10.17148/IJARCCE.2023.12572

detection. Deep learning algorithms use neural networks to learn patterns and features from large datasets of images and other data. These algorithms can be trained to recognize specific objects and obstacles and can be tuned to adapt to specific environments and conditions. However, training deep learning algorithms requires large amounts of data and may not generalize well to new environments and situations.

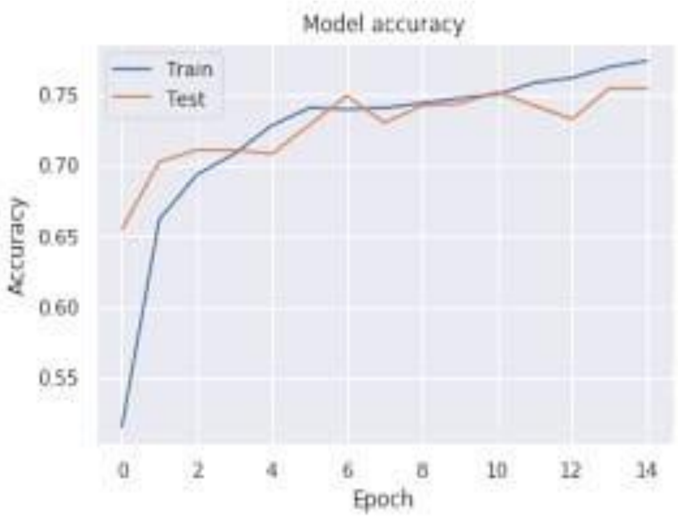
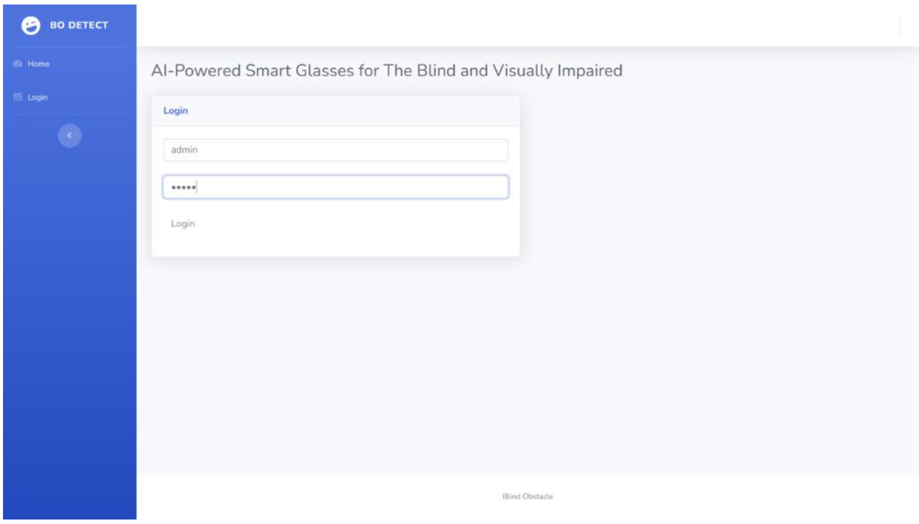


Fig.2 Model Accuracy

Screenshots

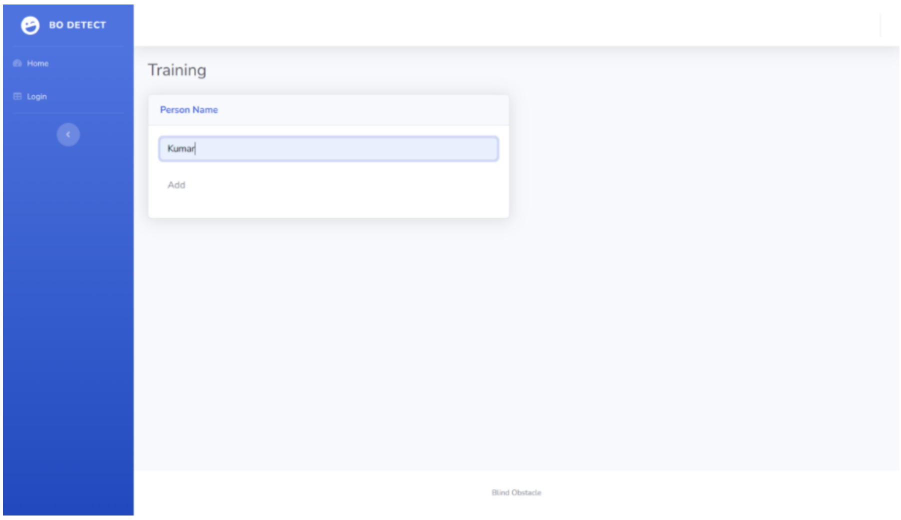


AI-Powered Smart Glasses for The Blind and Visually Impaired

admin

•••••

Fig.3 Login Page



Person Name: Kumar

Fig.4 Training Page

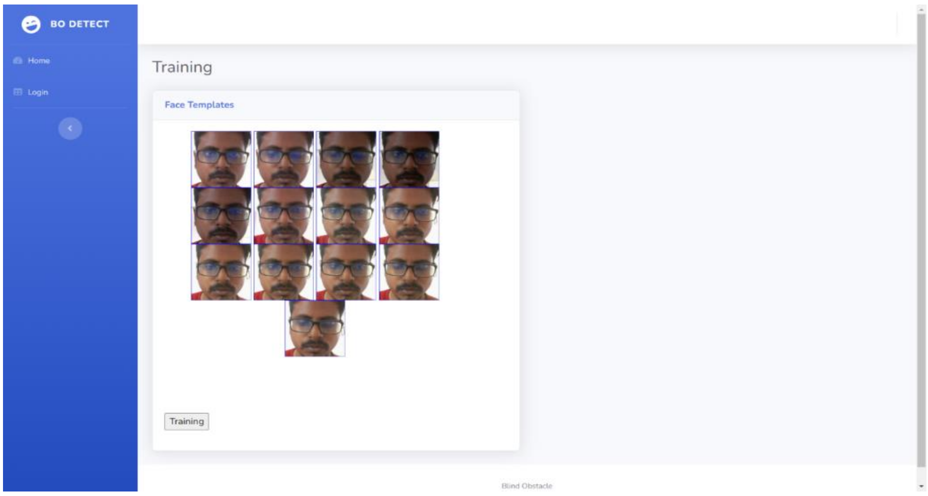


Fig.5 Frame Extraction

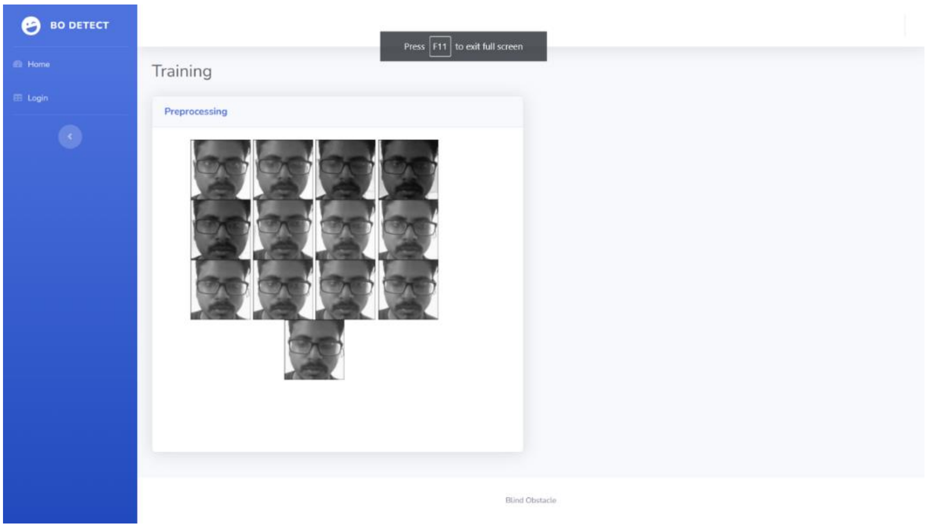


Fig.6 Preprocessing the image

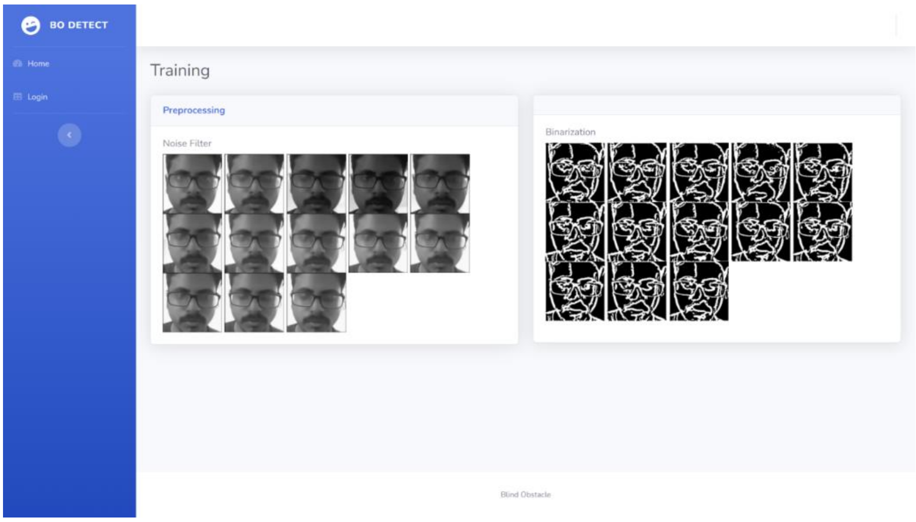


Fig.7 Binarization process

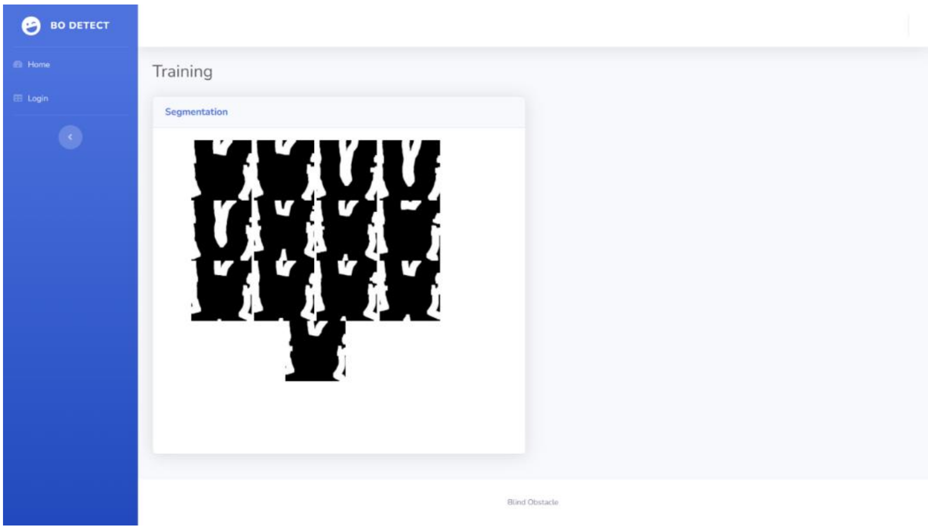


Fig.8 Segmentation process

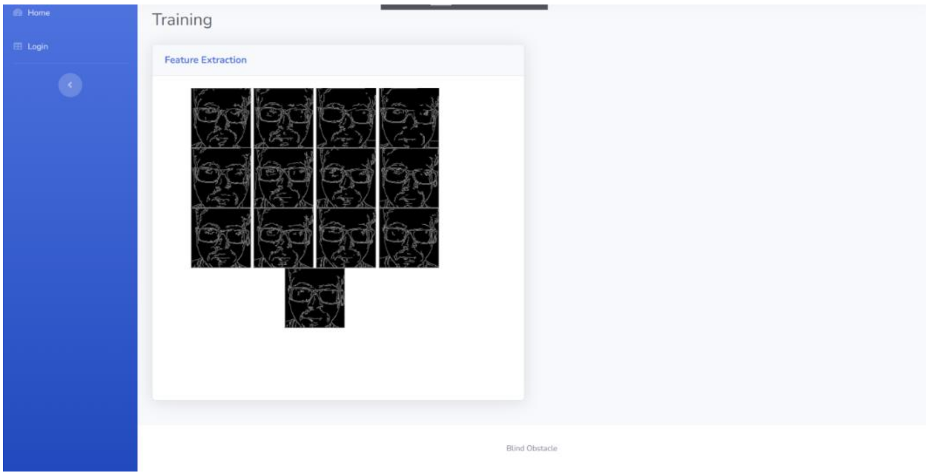
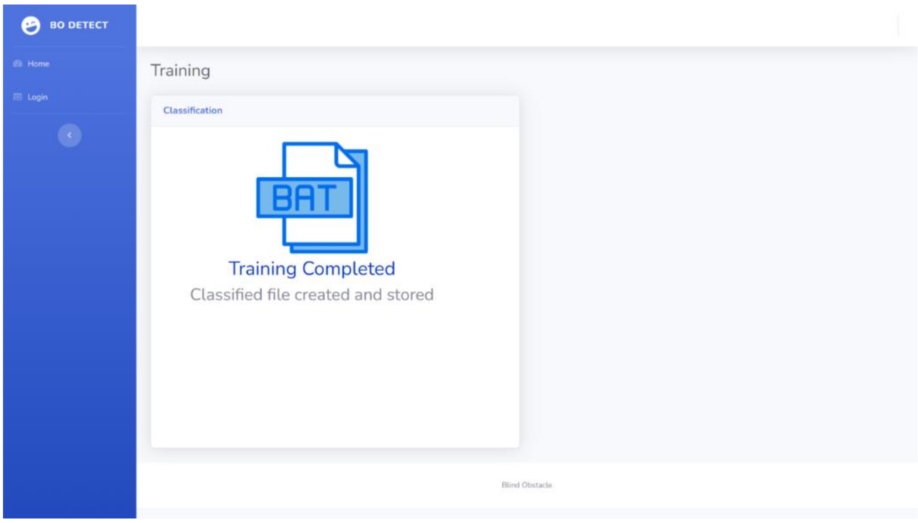


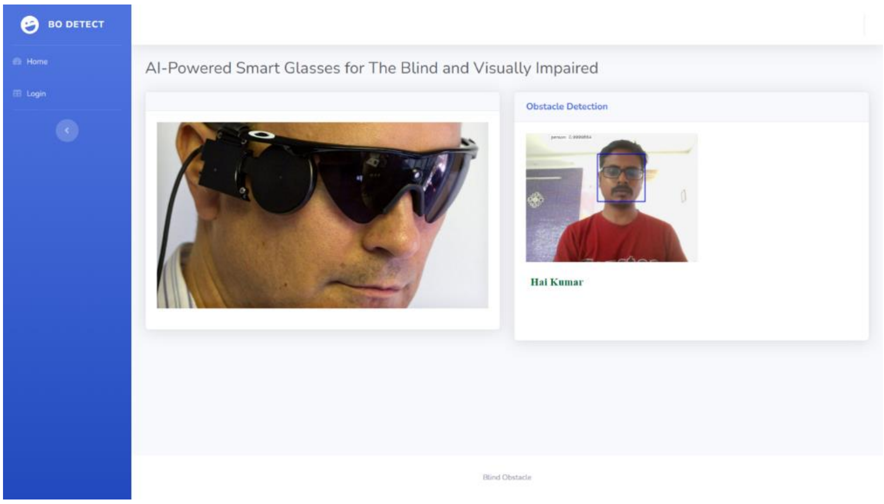
Fig.9 Feature Extraction



Training Completed

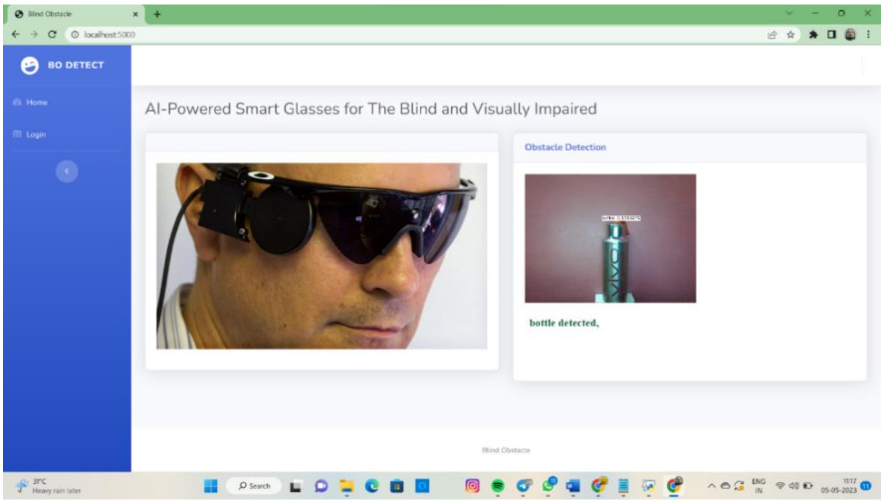
Classified file created and stored

Fig.10 Training process completed



Hai Kumar

Fig.11 Face Recognition



bottle detected.

Fig.12 Object Recognition

VII.CONCLUSION

The device presented here is a smart glass that has machine vision and obstacle detection and perception sensors. It can be easily promoted and made accessible to the visually impaired. It also helps us prevent further damage. You can easily carry your smart device around and use the system camera to track objects and faces from your surroundings and display them in audio formats.

Each model represents a specific task or mode. A user can perform a desired task independently of other tasks. The design, working mechanism and principle of the system were explained along with some experimental results. It allows visually impaired people to interact more closely with those around them without fear of being blurred and uncertain.

VIII.FUTURE ENHANCEMENT

In future One possible area for improvement could be the development of a more advanced and efficient prosthetic control system that can interpret more complex and subtle gestures from the user. This would allow for greater flexibility and precision in performing various daily activities. Another area of future enhancement could be the incorporation of artificial intelligence (AI) and machine learning algorithms to enhance the SAR system's ability to recognize and interpret environmental cues and provide more personalized feedback to the user.

This could potentially lead to a more intuitive and seamless integration between the user and the SAR system. Furthermore, the integration of haptic feedback technology could provide additional sensory information to the user and further enhance their ability to interact with the environment. For example, the incorporation of touch-sensitive sensors could allow the user to receive tactile feedback when touching different objects, providing additional cues about the shape, texture, and temperature of the objects.

REFERENCES

[1]. Guo, W., Zhang, Y., Chen, P., & Liu, Q. (2018). A Wearable Obstacle Detection System for Blind People Using Deep Learning. Sensors, 18(11), 3639. doi:10.3390/s18113639
[2]. Hu, L., Hu, J., & Long, J. (2020). Object Detection for Visually Impaired People Using Deep Learning on Wearable Devices. Sensors, 20(1), 164. doi:10.3390/s20010164
[3]. Ramalingam, V., & Swarup, S. (2019). Object detection and classification for visually impaired using deep learning techniques. 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT). doi:10.1109/icicict47191.2019.8993963
[4]. Gupta, S., Sethi, S., & Chhabra, J. (2019). Object detection for visually impaired using deep learning. 2019 2nd International Conference on Data, Engineering and Applications (IDEA). doi:10.1109/idea48677.2019.8987189
[5]. Islam, M. M., Rahman, M. S., Ananna, T. T., & Siddique, M. A. (2020). Object Detection System for Visually Impaired People Using Convolutional Neural Networks. 2020 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST). doi:10.1109/icrest48956.2020.9067051
[6]. Saini, P., Chaudhary, S., Chaudhary, M., & Mohan, A. (2019). Obstacle Detection and Alert System for Visually Impaired People Using Ultrasonic Sensors. Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). doi:10.1145/3311940.3311956
[7]. Ren, W., Zhang, C., Wang, Z., & Wu, X. (2020). An Efficient Object Detection Approach for the Visually Impaired Based on Deep Learning. 2020 2nd International Conference on Advances in Computer Technology, Information Science and Communications (CTISC). doi:10.1109/ctisc50842.2020.9277720
[8]. Shakya, R., Shrestha, A., & Shrestha, S. (2021). Real-time Obstacle Detection and Tracking for Blind People. 2021 5th International Conference on Information Technology (InCIT). doi:10.1109/incit51476.2021.9442557
[9]. Jain, P., Ramachandran, V., & Sarvadevabhatla, R. K. (2020). Vision-based obstacle detection system for visually impaired people using deep learning. 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence). doi:10.1109/confluence48654.2020.9058764
[10]. Mohammed, A., & Ismail, N. (2021). Obstacle Detection and Avoidance for Visually Impaired People Using Deep Learning Techniques. 2021 6th International Conference on Internet of Things: Systems, Management and Security (IOTSMS). doi:10.1109/iotsms52002.2021.9413129