

Real-time Smart Navigation System for Visually Impaired People

S.R.D.Sudaraka Keshara
Department of Software

Engineering
Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
it19109190@my.sliit.lk

Jathusanan Ellankovan
Department of Software
Engineering

Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
it19049878@my.sliit.lk

W.R.J.M.Weragoda
Department of Software
Engineering

Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
it19008424@sliit.lk

Madushan.W.A
Department of Software
Engineering

Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
it19042152@sliit.lk

Sanjeevi Chandrasiri
Department of Information
Technology

Sri Lanka Institute of Information
Technology
Malabe, Sri Lanka
sanji.c@sliit.lk

Abstract - Visual sense plays a primary role in guiding sighted people through an unfamiliar environment and assisting them to reach a destination safely. Visual impairment describes the actual damage that makes it difficult to accomplish visual tasks because it makes it difficult to see clearly. This paper proposes an approach to overcome the challenges faced by visually impaired people with the help of machine learning. This proposed system combines a smart cane and a wearable smart glass. Obstacle and pothole detection help to increase the safety and comfort of visually impaired users by detecting and displaying obstacles, and the Smart Walk-lane Navigation assists in navigating through the walk-lanes without letting them into the main roads and helps to prevent accidents. Road sign detection allows users to follow road signs and cross the roads safely, while the face and emotion detection allows users to recognize well-known people and their emotions.

Keywords - Visual Impairment, Machine learning, Smart Cane, Smart Glass.

I. INTRODUCTION

While most people in the world have been blessed with proper eyesight, a community suffers from a disability named visual impairment and therefore struggles to build the life they desire due to poor vision, which causes shortcomings such as lack of environmental identification and navigational struggles. According to the recent global estimations of the World Health Organization, at least 2.2 billion people are suffering from various visual impairment defects. Upon further inspection, it was found out that almost half of these cases could have been prevented if they were correctly diagnosed at the proper age and given treatment accordingly. Many of those struggling with this disability overcome their shortcomings to a certain extent through heightened other senses such as good hearing and navigation using the palms of the hands.

Many supporting systems are in place to assist the visually impaired community, such as walking sticks and Electronic Travel Aids (ETA). With the technological advancement of the current world and the ever-growing need for improved technical devices, the concept of image processing plays a

significant role in producing the next generation of ETAs with more advanced features to assist the visually impaired community. Though these devices are incredibly efficient for a visually impaired person, several factors have caused a handful of people to acquire this assistance, such as the high price factor in the technology and systems only having a few useful features. This situation creates a need for a system that not only appeals to a large crowd of people with disabilities but is also affordable at a reasonable rate.

The study's main aim is to create an affordable and user-friendly navigation system for all age groups while maintaining the accuracy and efficiency of highly developed systems. Smart Navigation application utilizes multiple devices such as a walking stick incorporated with sensors and smart glass to obtain accurate time information from the environment and relay them to the user in a coordinated pattern through audio instructions so that the user can navigate safely around various environments and be more encouraged to lead an active life.

II. LITERATURE REVIEW

A. J. Ramadhan discussed the "Wearable Smart System for Visually impaired people" [1]. A wearable system with a microcontroller and many sensors was suggested to help visually impaired people navigate the outdoors and alert them by vibrating their wrists. When a user requires assistance, the system warns those nearby, and To registered mobile phones of family members and caregivers, the alarm and the system location are delivered as a message.

Shan Xu talked on a method of structured road lane recognition for blind travel aid[2]. A Median Filter, which has been used to process images first, has been used to mark off the region of interest in the initial image. Then, the image was segmented using a threshold and Canny Edge Enhancement, and a modified Hough Transformation was performed to suit the road lane. This algorithm's extreme reliability and timeliness have been demonstrated.

Ramer talked about the Wearable Jogging Navigation System for the Visually Impaired on Less Structured Paths

with Adaptive, Color-Based Lane Detection[3]. Based on the 2D and 3D cameras for surrounding perception, feedback is given as vibrotactile feedback to the users to guide them towards the center of the running lane. For lane detection, 2D images of the camera, which is mounted on the chest, are used, and 3D sensors are used to prevent collisions. A 2D/3D camera, a laptop for data processing, and an elastic chest belt with built-in vibration motors make up the system. In addition, the camera is actively stabilized to enhance the image's quality due to the upper body's movements during jogging.

T. Nguyen, S. Phung, and A. Bouzerdoum proposed a hybrid deep learning-based method for pedestrian lane detection in an unstructured environment [4]. This method suggests using a hybrid deep learning-Gaussian process (DL-GP) network to separate lane and background regions from a scene image. This combines a small nonparametric hierarchical GP classifier with a powerful convolutional encoder-decoder net. As a result, overfitting issues can be reduced by using fewer parameters. A new data set of 5000 images is also introduced for training and testing the algorithm.

At the very least, there have been some initiatives to help the blind go about their daily lives. Most initiatives mix modern, cutting-edge technology with traditional commodities that blind people use to transport. As a result, blind people need to carry fewer things. There are also remote human help services available, whereby blind people can receive telecommunications-based assistance for various tasks. Large corporations more frequently employ the latter.

Another intriguing work by Foo et al.[5] was published to help grocery shoppers who are blind or visually impaired. The cost of supplying groceries to those who are blind or visually impaired can be very significant for supermarket owners. People can use this app to help them make independent food purchases. For persons who are blind or visually handicapped, food is a must. Therefore, this can make their life easier. The text-to-speech notifications, object identification, and audio localization employed in this program are all fundamental components of computer vision.

The Grozi project's capacity to aid someone in finding an aisle using sound localization and identifying merchandise and aisle sections using object recognition was one of its defining characteristics. After specific audio was played repeatedly, a multi-dimensional microphone system was used to pinpoint the exact location of each[6] aisle in a 3-dimensional landscape. This is a practical choice, but it can only be applied in peaceful settings. This is absurd because of how noisy grocery stores may be all day. On the other hand, a method based on images that can read QR codes at the beginning of aisles would be more efficient.

The authors advised that people with vision difficulties use a blind stick. They created the method and incorporated it into the smart pole to help with obstacle detection and pathfinding. While the ultrasonic sensor looks at obstructions, the buzzer[7] alerts the user. The process can only be used for a short distance, and the visually impaired person can only find it inside that area. The stick also encourages blind people's independence.

Ultrasonic sensors were utilized in developing the Ultrasonic Blind Walking Stick [8], where the main objective is to alleviate the shortcomings of conventional stick instruments. The ultrasonic sensor component is meant to help

the blind find objects in their field of vision. With pinpoint accuracy, this device can find uphill or downward areas. In addition, this blind stick can detect obstructions and challenges, which is helpful for impaired individuals traveling alone in more difficult muddy terrain.

Mishra and Koley [9] created a relatively affordable tool to locate and track down blind or visually impaired persons. The objective is to develop low-cost equipment that can locate a location for a person with vision impairment in a remote area without the assistance of a guide. A GPS tracking system is utilized to determine the device's location, and an ultrasonic sensor is used to look for obstructions. The system also has a voice-based communication mechanism for sending and receiving data.

The authors created an electronic travel aid for blind individuals [10]. The prior processing system was a single-board device. The author uses this equipment to take a photo in front of blind people, process it, and then send it to blind individuals using stereo earbuds utilizing fuzzy clustering techniques to produce real-time visuals. After the image has been analyzed, stereo sound is generated. The system comprises many image processing techniques that automatically use advanced decision-making tools called fuzzy logic systems to determine position and location automatically.

III. METHODOLOGY

The diagram below displays the separate components of the system, such as Face recognition with Emotion Detection, Walk-lane detection, Road sign detection, and Obstacle detection with pothole detection using the smart cane, and how these components function according to the priority assigned to each function in different scenarios faced by the user. In addition, each function has been associated with an audio/vocal instruction or command feature. The information processed by the system is to be relayed to the user via headphones so that the user can make decisions based on the information provided. Several methods are being used in the research to provide accurate and effective results. First, the system acquired real-time video data through the camera, analyzed the video frames, extracted the required object, and subtracted the background to enhance the image's foreground. Then, the filtered object runs through a feature detection program to identify the object type, ranging from road signs, obstacles, walk-lanes, or individuals.

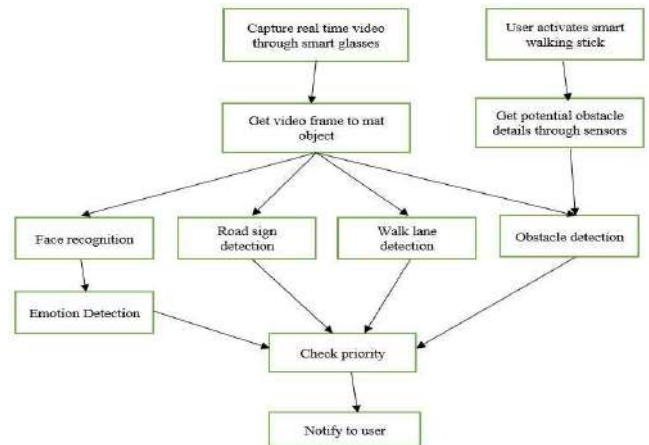


Fig 1: System overview diagram

A. Walk-lane Detection

The proposed walk-lane detection system can detect the walk-lanes on the main roads. In addition, the System can guide the user through the lane using audio commands, and if the user tends to move outside from the walk lanes, Users will be warned with voice commands to walk through the lanes. Walk-lane detection is carried out through the smart glass gadget. The Raspberry Pi is used as the microcontroller in the System. The Pi camera module sends a live video feed from the environment to the microcontrollers for processing. Since the System utilizes eyeglasses, the video feed from the eye level is acquired.

This System is developed with a simple neural network to extract features from the lane classification dataset for the classification of walk-lane detection. In walk-lane detection, convolution layers are used in place of the classification network's completely connected layers to create the detection network. The datasets of walk-lanes and main roads are collected from Colombo's significant cities. TensorFlow lite is used to build the classification model to differentiate walking lanes from the main roads

When the user starts walking on the walk-lanes, this system starts guiding them through the voice commands. The live feed from the camera is processed using the OpenCV library, which is an open-source real-time computer vision library. Initially, the region of interest (ROI) is separated from the live feed. Then, the ROI is converted into gray scale and then gaussian blur is applied to reduce the noise of the frames. After that, canny edge detection is performed to detect the edges of the ROI. To transform the detected clusters of edges into actual lines, the Hough line transform method is used. Each line segment from the output of the Hough transform method contains two coordinates: one indicates the start of the line and the other denotes the end of the line. From these coordinates, the slopes and intercepts of each line are calculated. From the calculated slopes, each line is classified as a left or right line. Average values of slopes and intercepts are acquired from the left and right lines. From the average values, two lines are constructed on the left and right sides to represent the walk-lane boundary. From this, the boundary system can understand whether the user is on the lane or not.

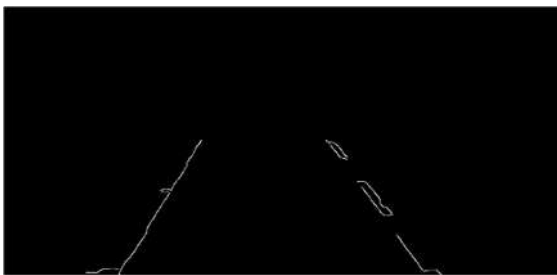


Fig 2: Canny Edge Detection on the ROI

To provide proper navigation, a horizontal line is constructed at the bottom of the ROI, and a vertical line is also constructed in the middle of the ROI. While walking on the lane, from the deviation of the vertical line from the horizontal line, proper voice alerts such as "move left or right" are provided to the users to stay in the middle of the walking-lanes.

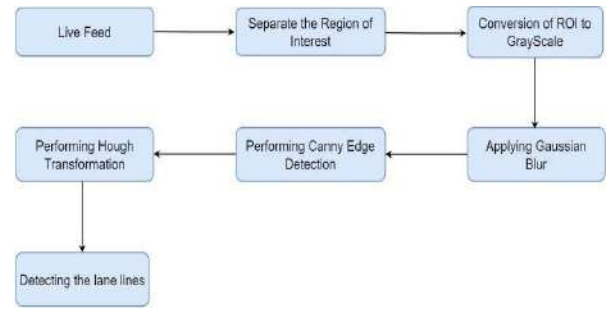


Fig 3: Flow Diagram showing the process of detecting the lane lines

B. Road sign Detection

The system can identify traffic signs and notify the user via headphones. Identifying traffic signs is carried out through a wearable IoT gadget. This gadget mainly consists of an eyeglass and a microcontroller attached to it. Raspberry Pi is used as the microcontroller in the system and using the Pi camera module, the live video feed from the surrounding is sent to the microcontrollers for processing. Since the system is uses eyeglasses, the video feed from the eye level is acquired.

The system locates the sign-containing region of interest using image processing techniques, and then extracts this region for categorization. Searching the area of interest with sliding windows might be time-consuming and result in a lot of erroneous classifications. Hence, the system uses a more straightforward, faster, and reliable approach. Shape identification and symbol detection are the system's two major parts. The sign detector looks for areas in an image that match a sign using a conditional maximum entropy model. The sign recognizer compares the predicted sign regions to databases of sign images. The system then determines whether the most probable sign is accurate or whether the hypothesized sign region does not correspond to a sign in the database. The data sets cover a broad spectrum of variability, such as variations in lighting, orientation, and viewing angle. Once a sign has been found, the user receives this information via headphones. To assess the system's performance, the detection rate of the system was determined using a number of randomly chosen photographs of signs and pictures without them. The effectiveness of user feedback was assessed by having volunteers rate the system's usefulness.

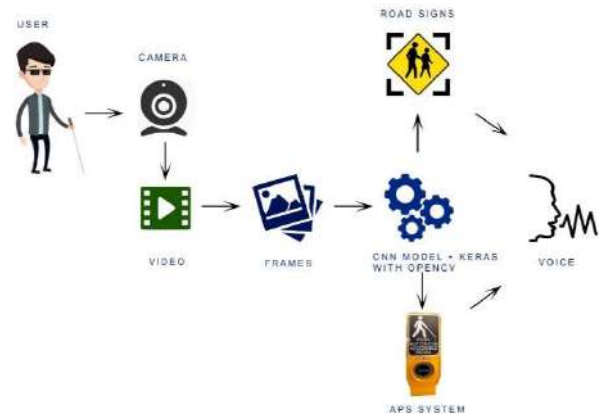


Fig. 4: System diagram for road sign detection

Loads of existing solutions can be used to identify traffic signs or pedestrian crossings. However, crossing the road is an enormous challenge for a visually impaired person, especially considering hurdles like dealing with traffic, finding the edge of the street and crosswalks, and staying on course during the crossing are all essential tasks. Therefore, a simple solution has been developed to help visually impaired people to cross the road safely named APS, Accessible Pedestrian Signals.

1. This will make it easier to find a crosswalk.
2. It will also indicate the best time to start crossing.
3. It can also allow a person to maintain a straight trajectory during the crossing.

The downside of APS is that it is deafening, and the tones can be loud and disruptive to some residents, especially those residing beside traffic signals. The Smart-Nav system will be an alternative solution for the existing APS system.

The system is capable of identifying the pedestrian crossing, reading the countdown pedestrian signals, and making the decision whether the user can cross the road safely or they should wait for the next time. Another raspberry PI is connected to the device near the pedestrian signals. Two devices are connected via Wi-Fi when the user wants to cross the road. The device in the pedestrian signal reads the countdown numbers and notifies the user about the decision via headphones after the processing.

C. Obstacle Detection

Real-time object recognition and object access to previously defined classes are the main goals of this component. The used algorithms are more computationally effective. Many people in the visually impaired community use the white cane – the most successful and widely used travel aid for the blind. This entirely mechanical device is designed to find risks such as steps, holes, uneven ground, and other obstructions. The main problem with this device is that users must be trained in its use for more than 100 hours – a substantial “hidden” cost. The white cane also necessitates an active scan of the restricted space in front of the user. To learn more about them, they must use the cane to touch everything. Detecting potholes using a cane is also not a convenient task.

Nowadays, almost everything has a camera, from pens to cell phones. This has given rise to a new field, computer vision, to get reliable and valuable findings. The technique examines the individual frames of the video to identify the motion of the foreground, extracts the object based on its color, and then subtracts the background to emphasize the foreground. The foreground has been identified using feature detection on the background-subtracted image. Audio analysis has also been started to identify the items that appear in front of the system.

Through the Smart Navigation system, the user can recognize obstacles, walking surfaces, and potholes using the wearable IoT device and smart cane to provide appropriate information through audio and signal for safe navigation.

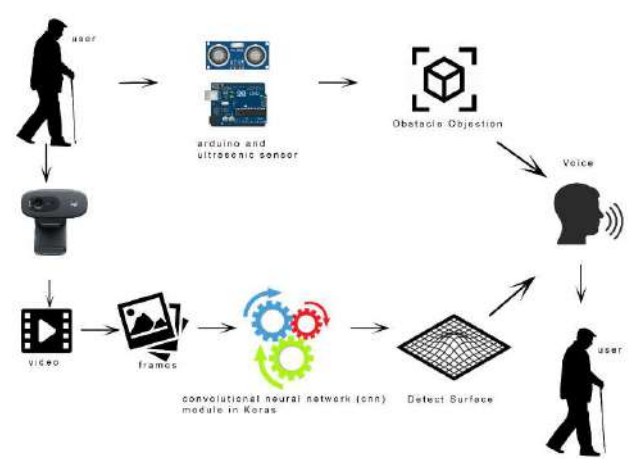


Fig. 5: System diagram for obstacle detection

This method has two main parts: a Smart cane for detecting obstacles and Surface detection and pothole detection using smart glasses.

Obstacles detection using IUP is impractical thing. When it comes to the real world, there will be a time gap between processing time. So, it will be a negative impact on them. Also, the white cane is an inevitable thing for people who are

visually impaired or blind. Regarding features of the smart cane plugin, there are 2m obstacle detection coverage, fast response time, automatic lighting system, and ease to setup (plugin device).

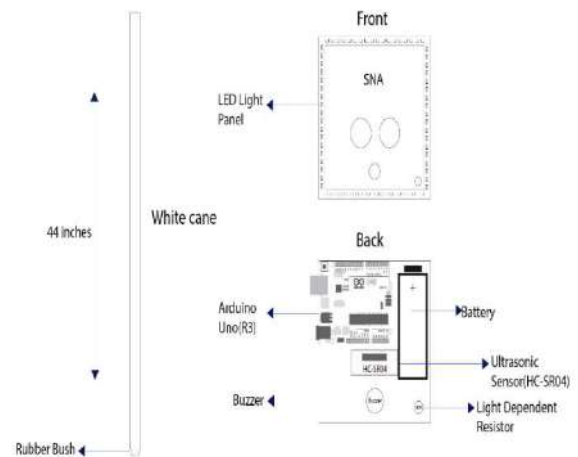


Fig. 6: Blueprint of smart cane plugin

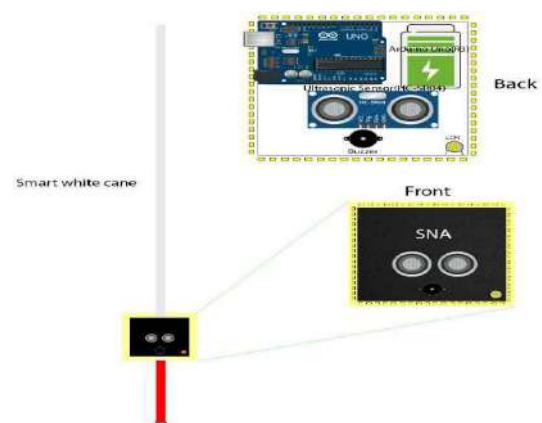


Fig. 7: Prototype diagram for smart cane

There are a few significant phases for surface and pothole detection. Data preparation was conducted in three steps collecting data using Kaggle.com and customer dataset, pre-processing, and data splitting. Then the training and testing model on data was conducted. There, the emphasis was mostly on larger data handling capacity, better accuracy, reduced memory utilization, more efficiency, and faster training speeds. Finally, analysis/evolution of the model, detecting potholes and surfaces from inputs, and voice command implementation was held.

To assess the system's performance, the detection rate of the system was determined using a number of randomly chosen photographs of obstacles and those without them. Additionally, the effectiveness of user feedback was assessed by having volunteers rate the system's usability.

D. Face Recognition with Emotion Detection

One of this system's critical features is to identify family members or friends of the user through accurate time image processing. Where it accesses the user's database and filters through previously uploaded images of critical individuals of the user extracting key structural facial features and provides a confirmation signal when it successfully locates a known individual of the user, this feature caters to the need of the user to identify and interact freely with their social circle. Unfortunately, the current systems that are in place to aid the visually impaired community do not consist of such a feature, and the lack of user-focused features such as these causes the people in the visually impaired community to become one-sided in interacting with society and tends to cause further emotional stress for the user.

Correct face identification of known individuals to the user is coupled with an emotion detection component to enhance interactivity with the person. The emotion detection model utilizes seven basic emotions: anger, disgust, sadness, happiness, surprise, fear, and neutral reactions. The emotion detection model notifies the emotional response of the known person of the user and allows the user to respond accurately to the other individual's emotions, thereby increasing social awareness.

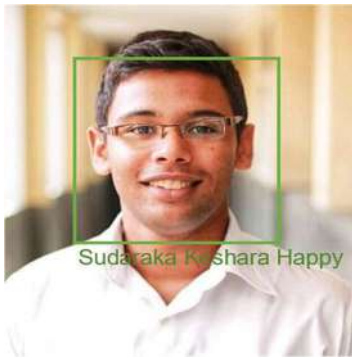


Fig. 8: Face and Emotion recognition of known individual

When considering the various scenarios, waiting for a bus and if the user was to intrude at that moment. It would cause inconvenience to the person; therefore, in the training data set, the algorithm seeks out the background and emotion patterns of the identified individual and compare the characteristics of the input data, and determines whether the individual is in a state of emergency or not and notifies the user accordingly so that the user can decide between wanting to meet that person or to avoid contact depending on the situation.

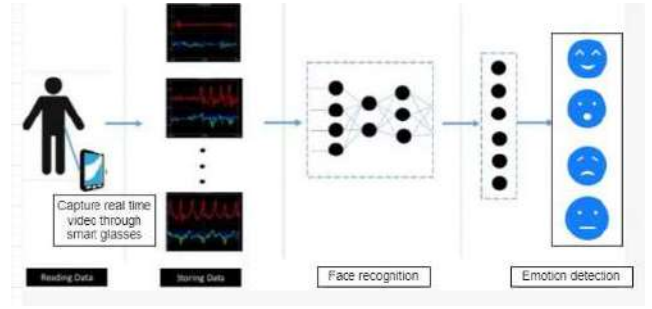


Fig. 9: Face recognition and emotion detection system overview

The system uses image processing techniques to find the facial features of the face that contains unique biometric characteristics to distinctly identify an individual and extract data from the facial region for the classification of emotions displayed by the recognized individual. The process is carried out sequentially, whereas the system only relays the details relating to the emotional aspect only when a person is correctly identified. Once the system has successfully detected a person, this information is communicated to the user via headsets from which the user can choose to interact with the identified person while triggering the emotion detection component to enhance the interaction. The system performance was evaluated considering tests such as successfully identifying individuals in different environments and multiple individuals simultaneously with the same test conditions. Additionally, a performance evaluation of the emotion detection component was done for a variety of emotions that could be shown by one person or many people at once.

IV. RESULTS AND DISCUSSION

Several practical tests have been conducted in various environments and individuals, and the results have been obtained to evaluate the system's performance. The proposed gadget is more convenient for visually challenged people because the microcontroller is attached to the eyeglasses.

The face recognition program could identify individuals in a close range with high accuracy; however, the accuracy gradually declined as the individual was out of clear range for the camera. Nevertheless, upon increasing the camera's quality, the system could identify an acceptable range for the user. The emotion detection model currently works only near the user as facial features should be visible to identify the correct emotion displayed by the individual.

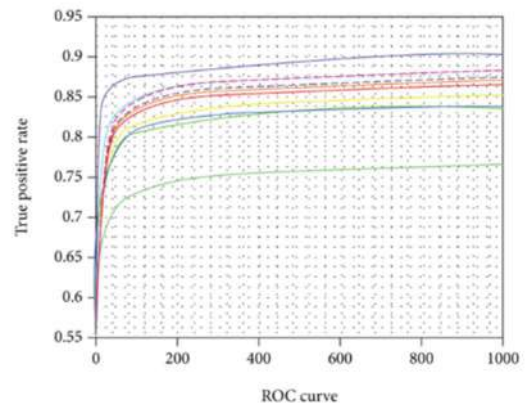


Fig. 10: Face and Emotion Detection Model Accuracy Graph

The walk-lane detection system was created to aid those who are blind or visually impaired when moving through the outside world. Users can securely navigate the walk lanes using an integrated voice command feature. With more than 80% accuracy, this system offers a straightforward, practical, and dependable answer to the users' travel-related issues.



Fig. 11: Visual result of Walk-lane detection

A road sign detection system has been conducted with an 80% accuracy rate with shortcomings in the inability to identify the road sign correctly due to discoloration and improper alignment of signs. Obstacle detection and pothole detection using a smart cane has been conducted using an alarm notification to identify potential obstacles and potholes on roads, but the team intends to switch to voice notifications for further efficiency in the system.

Though the system is accurate, there are some shortcomings, such as being unable to use it in dim light and the nighttime. Another limitation of the system is that except for the smart cane detection of potholes, the user must be facing a particular direction to identify road signs, individuals, obstacles, or walk lanes.

V. CONCLUSION

The Smart Navigation system has been primarily developed using the concept of Image Processing and also uses the concept of IoT for the smart cane implementation. The application has been developed with a user-friendly interface to effectively aid the visually impaired community. Due to multiple features such as face recognition coupled with emotion detection, walk-lane detection, road sign detection, obstacle detection, and pothole detection using the smart cane, the current system presents a unique opportunity for blind people to become more active in a cost-effective manner. The features detected through the camera and its image processing results are converted into audio/vocal commands and relayed to the user. Compared to the other systems in the market, this system consists of several unique features that cater to the multiple needs of blind people and encourage them to lead a more active social life. However, there are several shortcomings in the application. Therefore, overcoming these shortcomings and making this application more usable for the blind community is essential.

ACKNOWLEDGMENT

Completing this research could not have been possible without the participation and assistance of many individuals contributing to this project. First, however, the team would like to express our most profound appreciation and indebtedness to our lecturers and supervisors for their endless support, kindness, and understanding during the project.

REFERENCES

- [1] A. Ramadhan, "Wearable Smart System for Visually Impaired People", MADPI, 2018. [Online]. Available: <https://www.mdpi.com/1424-8220/18/3/843>. [Accessed: 31- Aug- 2022].
- [2] Xu, Shanshan, Jie Ying, and Yanbin Song. Research on road detection based on blind navigation device", *Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2012. [Online]. Available: <https://ieeexplore.ieee.org/document/6319889/>. [Accessed: 12- Feb- 2022].
- [3] C. Ramer, T. Lichtenegger, J. Sessner, M. Landgraf and J. Franke, "An adaptive, color based lane detection of a wearable jogging navigation system for visually impaired on less structured paths", *Ieeexplore.ieee.org*, 2016. [Online]. Available: <https://ieeexplore.ieee.org/document/7523715>. [Accessed: 18- Jun- 2022].
- [4] T. N. A. Nguyen, S. L. Phung and A. Bouzerdoum, "Hybrid Deep Learning-Gaussian Process Network for Pedestrian Lane Detection in Unstructured Scenes", *Ieeexplore.ieee.org*, 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/8998354>. [Accessed: 20- Apr- 2022].
- [5] P. Meijer, "An experimental system for auditory image representations", *Ieeexplore.ieee.org*, 1991. [Online]. Available: <https://ieeexplore.ieee.org/document/121642>. [Accessed: 23- Apr- 2022].
- [6] Manikanta K S, Phani T S S and Pravin A, "Implementation and Design of Smart Blind Stick for Obstacle Detection and Navigation System International Journal of Engineering Science and Computing.", *Ijesc.org*, 2018. [Online]. Available: [https://ijesc.org/upload/6551810b6c41b74e2aea635695ae2034.Implementation%20and%20Design%20of%20Smart%20Blind%20Stick%20for%20Obstacle%20Detection%20and%20Navigation%20System.p](https://ijesc.org/upload/6551810b6c41b74e2aea635695ae2034.Implementation%20and%20Design%20of%20Smart%20Blind%20Stick%20for%20Obstacle%20Detection%20and%20Navigation%20System.pdf)df. [Accessed: 22- Jun- 2022].
- [7] Amjed S. Al-Fahoum, Heba B. Al-Hmound and Ausaila A. Al-Fraihat, "A Smart Infrared Microcontroller-Based Blind Guidance System", *Hindawi*, 2013. [Online]. Available: <https://www.hindawi.com/journals/apec/2013/726480/>. [Accessed: 24- Mar- 2022].
- [8] Ayat Nada, Samia Mashelley, Mahmoud A Fakhr and Ahmed F Seddik, "Effective Fast Response Smart for Blind People", *Second International Conference on Advances in Bio-Informatics and Environmental Engineering - ICABEE, 2015* [Online]. Available: <https://www.seekdl.org/conferences/paper/details/5920#>. [Accessed: 16- Jun- 2022].
- [9] Vijayalakshmi Badre, Romba Chhabria, Tanmay Kadam and Kritika Karamchandani, "Ultrasonic Blind Walking Stick With Voice Playback", *Irjet.net*, 2016. [Online]. Available: <https://www.irjet.net/archives/V3/i4/IRJET-V3I4384.pdf>. [Accessed: 22- May- 2022].
- [10] Somnath Koley and Ravi Mishra, "Voice Operated Outdoor Navigation System For Visually Impaired Persons", *International Journal of Engineering Trends and Technology (IJETT)*, 2012. [Online]. Available: <https://ijettjournal.org/archive/ijett-v3i2p217>. [Accessed: 20- Jun- 2022].