BLIND ASSIST SYSTEM USING MACHINE LEARNING AND IMAGE PROCESSING

Dr. Mahesh Kumar A S[1], Dr. N M Mahesh Gowda[2], Dr. M J Anand[3], Dr. D M Srinivasa[4], Kumar N Krishnamurthy[5]

[1] Associate Professor, Department of Electronics and Communication Engineering, Maharaja Institute of Technology, Mysore, India

[2,3]Associate Professor, Department of Electronics and Communication Engineering, P.E.S College of Engineering, Mandya, India

[4]Associate Professor, Department of Electrical and Electronics Engineering, P.E.S College of Engineering, Mandya, India

[5]Assistant Professor, Department of Electronics and Communication Engineering, P.E.S College of Engineering, Mandya, India

Abstract- This article presents the idea of developing an intelligent framework that may assist vision impaired people with their daily activities. Those who are blind, face several challenges. They frequently require continual support in almost all settings, particularly throughout their everyday activities. The primary challenge is getting from one site to another without assistance. Other challenges include locating obstacles and recognizing people. We propose a "Blind assistance system" in this work to assist prevent this happening. The gadget is a voice-activated system that would help a blind person through daily duties. This method comparison aims to provide color, object, and face detection for smart blind technology. Additionally, this study compares and examines several color models, including RGB, YUV, CMY, CMYK, and HIS, as well as how people see color in different ways. However, this study contrasts the advantages and disadvantages of the face identification techniques Eigenfaces, Fisher faces, Local Binary Patterns Histograms, Scale Invariant Feature Transform, and Speed Up Robust Features. Python and OpenCV were employed as the programming languages for this study. The Raspberry Pi Wide Angle Camera Module, Arduino Uno and GSM, GPS modules are the IoT devices utilized in this work. In summary, our research improves the lives of persons who are blind or visually impaired through improved IoT and AI technology.

Keywords-Face recognition, obstacle detection, route navigation, sensors, voice commands, visually impaired, Raspberry Pi Wide Angle Camera Module, Arduino Uno, GSM, GPS modules

I. INTRODUCTION

Reflecting on our lives, the most valuable gift that we should all undoubtedly appreciate is the capacity to see. Vision allows us to see the world around us and supports us in establishing ourselves in unfamiliar circumstances. According to the World Health Organization, there are 285 million people who are visually impaired. Visually handicapped people with bad eyesight have a lot of difficulty navigating about and may even stumble and fall. People on foot are regularly forced to leave when sidewalks are clogged with animals, vendors, and other impediments. Blind persons frequently use canes; however, a regular cane cannot see anything above the waist. Persons who are visually impaired, or more generally, people who are exceptionally abled, are the ones that struggle the most even to do their everyday activities. Even if they don't want to, most of them must rely on assistance from others. For the benefit of these people, countless technologies are being created or already exist. The most potential remedy is offered by computer vision, one of these technologies. Blind folks have a difficult time navigating the streets. They frequently run the risk of being struck by a car or obstruction since they are unable to view the outside environment. They find it difficult to cross at the zebra since they cannot use the pedestrian traffic signal.

The Internet has evolved thanks to organised technology's quick development to the application of breakthroughs in daily life. Object recognition innovation, sometimes referred to as object detection, is one of the innovations to consider. This phrase refers to the ability to recognise the size, shape, and variety of things, while the device's camera records the items' locations. Object detection is the process of spotting actual objects like cars, bikes, Tele-Vision, flowers, and people in still images or films. It improves our overall knowledge of the situation by enabling us to distinguish, locate, and detect numerous objects inside a picture. Examples of applications include image retrieval, security, and advanced driver assistance systems (ADAS).

The major goal is to create a wearable computer vision system for blind individuals that will aid them in navigating obstacles and seeing their surroundings. By employing vocal commands to identify objects employing image processing technology and providing audio output to the user to guide them past barriers, this method will assist the blind. The system's primary goal is to identify objects and signboards. The vision challenged individual will find it easier to manage daily tasks in this way. As a result, the suggested technique will aid those who are blind in navigating any obstacles and will assist them visualise their surroundings.

II.LITERATURE REVIEW

Using guide dogs to assist them navigate their pathways or walking canes to protect it clear of obstacles is a classic strategy that visually impaired people have utilised for years. They are both affordable and accessible but not prone to mistakes. Being prone to mistakes is desirable for blind individuals since even the smallest mistakes might result in significant harm.

Another technique to cope with this situation is to provide blind folks with intelligent rehabilitative shoes beside the spectacles. Each of these shoes has an ultrasonic transducer on top to detect things at different heights, and the eyeglasses have a pair set centrally over a buzzer and the bridge at one of the extremities. This sensor-based technique has a significant flaw in that it only works well to detect objects rather than identify them; image processing offers a potential solution to handle such circumstances [1, 2, and 3].

To be completed in 2020, Mouna and Riadh's project titled "An Evaluation of Retina Net on Indoor Object Detection for Blind and Visually Impaired Persons Assistance Navigation." The purpose of computer vision in this work is to precisely detect indoor items. By exploring the CNN framework's goals, those who are blind or visually impaired can receive assistance. We must first detect the pixels present in the photos to identify the unique items. It is difficult to capture and accurately identify the items under the improper lighting circumstances. [4, 5, and 14]

2018 will see the completion of Han Hu and Jiayuan's "Relation Networks for Object Detection" project. This study allotted an equal amount of work by considering its features based on the relational models. This eliminates redundancy and achieves accuracy at a certain level. It employs objects instead of words since they are arranged in a 2D scale ratio. The concept is further divided into two parts that belong to the categories of geometrical and original weights. [15]

In 2018, Alice Tang and Zhiyuan plan to finish their project titled "Automatic Registration of Serial Cerebral Angiography: A Comparative Review." Based on this research, which focuses mostly on the medical profession, some adjustments have occasionally been made to the way diseases are diagnosed and treated in order to improve its efficacy and accuracy. Instead of using DSA, image processing methods that have undergone extensive scrutiny are used to evaluate computed tomography (CT) and MRI (magnetic resonance imaging). Even though DSA is used to detect several neurovascular disorders at the time of surgery, it is possible to infer from these factors that the framework was created with patients who had been given an ischemic stroke diagnosis in mind. [6, 7, and 8]

The research on "HCP: A Flexible Convolutional Neural Network (CNN) Framework for Multi-label Image Classification" by Wei and Xia was suggested in 2015. In this study, a CNN model yields the best results for label-free picture categorization. Multi tagging is a difficult problem for training picture layouts because of its intricacy. One picture object is used as the input for the extraction of hypotheses, and CNN is used to generate individual scores by maximum pooling. The various colours used to identify the image's hypotheses can be represented by several clusters. Max pooling makes use of the predicted outcomes produced by the extraction procedure. When analysing the HCP and I-FT approaches, it is shown that the HCP model enhances system effectiveness by 5.7%. [10]

"Indoor Object Recognition in a combination of an RGB(Red-Green-Blue) picture and its associated depth image (RGBD) Images using Complex-Valued Neural Networks for Visually-Impaired People" is the project that Rim and Issam have presented for 2018. In this study, a multi-model is utilised to help persons with visual impairments find things using a multiclass approach inside a building. This model processes many labels at once. The multi-label and CVNN approaches assign labels to the picture that at once relate to groups of objects. The MLCVNN approach may create clusters based on multiple labelling, and it uses image transformation to categorise the issue by ranking solutions. [16]

A project on " Using multi-label classification for acoustic pattern detection and assisting bird species surveys" has been suggested by Liang and Miachel in 2016 [18]. This approach is intended to identify trends in urban settings, such as busy streets, precipitation, dining establishments, etc. By characterising the audio snippets, this approach produces the sequences. This model's key drawback is that it needs a training data set. [9, 11, 12, and 13]

In the study of "A Compressive Sensing Approach to Describe Indoor Scenes for Blind People." [17], people can use a camera to recognise things in a variety of interior settings. This model employs a multiple labelling technique to calculate Gaussian Method and Euclidean distance. It checks if several entities pertaining to the data collection are present. It also shows the locations if the existence of a certain object is discovered.

III. EXISTING SYSTEM

People who are blind or visually handicapped have used many different tactics or methods to deal with this problem. The use of guide dogs or walking canes by visually impaired people has been a conventional strategy for many years. Guide dogs can help blind people navigate through their surroundings. Both would be accessible and affordable yet not prone to mistakes. For blind individuals, being mistake prone is a good thing because even the smallest mistakes can cause significant harm. Giving blind individuals smart rehabilitation shoes in addition to their glasses is another way to address this issue. Each of these shoes has an ultrasonic transducer on top to detect things at different heights, and the eye glasses have a pair set centrally over a buzzer and the bridge at one of the extremities. This sensor based technique has a significant flaw in that it only works well to detect objects rather than identify them; image processing offers a potential solution to handle such circumstances. Sticks are another common tool used in conventional obstacle-finding techniques. These sticks are employed to locate obstacles up front. When the stick contacts the obstruction, the user becomes aware of it.

IV. DRAWBACKS OF THE EXISTING SYSTEM

These blind sticks are meaningless when needed such as crossing traffic lights since they cannot detect whether the signal is red or green. Furthermore, when using a stick, the obstruction can only be identified after the stick contacts the item and a blind stick cannot distinguish between different types of obstacles.

V. PROPOSED SYSTEM

Our approach to the issue is to design a tool that uses a camera to detect obstacles and a voice alarm to indicate whether a traffic signal is red or green. The device runs on a battery-operated Raspberry Pi, a small, compact arm computer. The component's design is portable and tiny, making it simple to transport. The environment will be continually captured on camera by this system, which will then turn the footage into frames. The technology will warn the user about an obstruction or the environment after examining these images. The key benefits include the system's ability to assist those who are blind while being portable, inexpensive, and approachable. The visually handicapped will be able to negotiate any obstacles with the aid of this device, which will also help them visualise their surroundings.

VI. BLOCK DIAGRAM

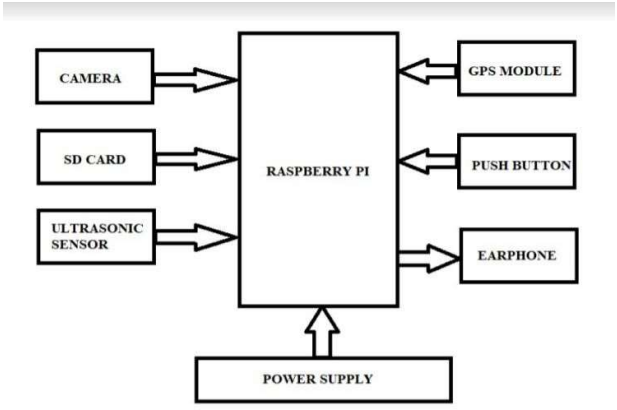

Figure 1: Block Diagram of Raspberry PI

A. RASPBERRY PI

Raspberry Pi is a series of small single-board computers (SBCs) developed in the United Kingdom by the Raspberry Pi Foundation in association with Broadcom. The Raspberry Pi project originally leaned towards the promotion of teaching basic computer science in schools and in developing countries. The original model became more popular than anticipated, selling outside its target market for uses such as robotics. It is widely used in many areas, such as for weather monitoring, because of its low cost, modularity, and open design. It is typically used by computer and electronic hobbyists, due to its adoption of HDMI and USB devices.

B. WEBCAM

A webcam is a video camera that feeds or streams an image or video in real time to or through a computer network, such as the Internet. Webcams are typically small cameras that sit on a desk, attach to a user's monitor, or are built into the hardware. Webcams can be used during a video chat session involving two or more people, with conversations that include live audio and video. Webcam software enables users to record a video or stream the video on the Internet. As video streaming over the Internet requires much bandwidth, such streams usually use compressed formats.

C. SPEAKER OR EARPHONE

Speakers are designed to connect with any kind of sound system, while some can be hooked up only with computers. With the computer speaker, the computer's sound card creates a signal that is used to produce sound. The primary objective of speakers is to offer audio output for the listener. The electromagnetic waves are converted into sound waves through the speaker as they are transducers. The devices, like an audio receiver or computer, give audio input to speakers, which may be in the form of analog or digital.

D. SD CARD

Designed to offer high-capacity memory in a compact form, a SD card (secure digital card) is a tiny flash memory gadget. Users may view and alter the data as they would with any flash storage card after the card is plugged inside the card reader and the card reader is linked to a computer.

VII. WORKING

Utilizing machine learning and image processing, our approach uses two processes to provide blind assistance.

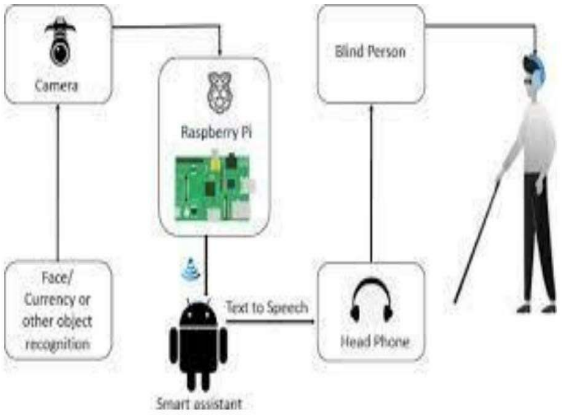

Figure 2: Blind Assist System

The steps entail:

Step 1: Creating and conditioning the neural network to recognize items:

Multiple dense layers and various activation functions, including ReLU and Sigmoid, are used to create convolution neural networks. This neural network was trained using tens of thousands of photos of everyday items. Images used for training are gathered using open-source data sets that are online. The trained neural network is retained and applied to item identification.

Step 2: Audio conversion and object identification:

The initial stage is live video capturing using a camera. To detect objects, video will be transformed into picture frames. The picture is then pre-processed utilizing the open cv package to get the desired resolution. Convolution Neural Network will be used to categorize and identify objects from reprocessed images. A speaker then announces the found item.

Step 3: Panic button and live location retrieval:

When a blind person encounters a problem or is in a critical situation, we have included the panic button which is the safety feature in this device. When the person hits the panic button, the GPS module retrieves the blind person's live location and sends it to the person who is in charge.

VIII.RESULTS

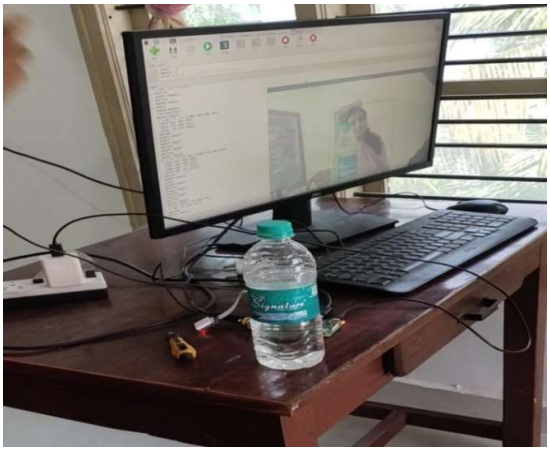

Figure 3: Person Identification

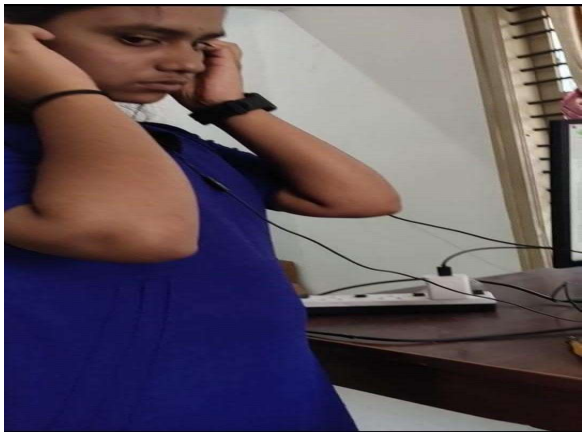

Figure 4: Hearing Audio Output using Earphone

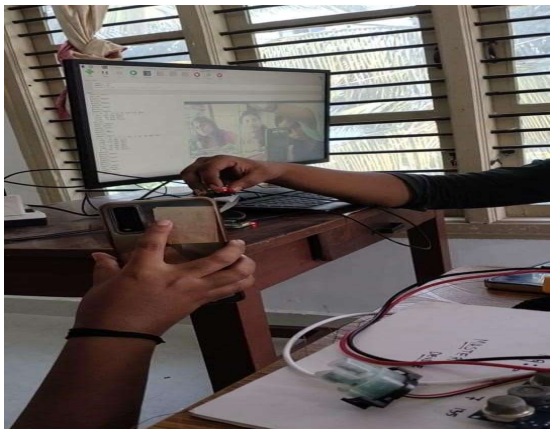

Figure 5: Object Identification

Figure 6: Raspberry pi connection

IX. CONCLUSION

The suggested system has successfully advanced to the point where it can assist persons who are blind in finding their way about and completing daily tasks with ease. It is now a portable, inexpensive, and accessible framework that requires image processing technology. The technology also seeks to assist the blind in navigating their environment by seeing barriers, finding their essentials, and reading signs and messages. Initial tests have produced encouraging results, allowing the user to move about his/her environment securely and freely. By allowing voice to be used as an input to achieve his/her fundamental needs, the structure is implemented considerably more user-friendly.

REFERENCES

[1] J. Connier, H. Zhou, C. De Vaulx, J.-J. Li, H. Shi, P. Vaslin, and K. M. Hou, "Perception assistance for the visually impaired through smart objects: Concept, implementation, and experiment scenario," IEEE Access, vol. 8, pp. 46931– 46945, 2020.

[2] K. Davila, S. Setlur, D. Doermann, U. K. Bhargava, and V. Govindarajan, "Chart mining: A survey of methods for automated chart analysis," IEEE Trans. Pattern Anal. Mach. Intell., early access, May 4, 2020, Doi: 10.1109/TPAMI.2020.2992028.

[3] K. C. Shahira and A. Lijiya, "Document image classification: Towards assisting visually impaired," in Proc. IEEE Region 10 Conf. (TENCON), Oct. 2019, pp. 852– 857

[4] Kavasidis, C. Pino, S. Palazzo, F. Rundo, D. Giordano, P. Messina, and C.Spampinato, "A saliency-based convolutional neural network for table and chart detection in digitized documents," in Proc. Int. Conf. Image Analysis and Processing, Springer, 2019, pp. 292–302.

[5] K. Kafle, B. Price, S. Cohen, and C. Kanan, "DVQA: Understanding data visualizations via question answering," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition, Jun. 2018, pp. 5648–5656.

[6] L. Gomez-Robledo, E. Valero, R. Huertas, M. Martínez-Domingo, and J. Hernandez-Andres, "Do EnChroma glasses improve color vision for colorblind subjects?" Opt. Exp., vol. 26, no. 22, pp. 28693–28703, 2018

[7] Q. Ye and D. Doermann, "Text detection and recognition in imagery:
A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 7, pp. 1480–1500, Jul. 2015.

[8] C. Clark and S. Divvala, "PDF Figures 2.0: Mining figures from research papers," in Proc. IEEE/ACM Joint Conf. Digit. Libraries (JCDL), Jun. 2016, pp. 143–152.

[9] C. A. Clark and S. Divvala, "Looking beyond text: Extracting figures, tables and captions from computer science papers," in Proc. Workshops 29th AAAI Conf. Artif. Intell., 2015, pp. 1–6.

[10] O. Gambino, E. Minafo, R. Pirrone, and E. Ardizzone, "A tunable digital ishihara plate for pre-school aged children," in Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Aug. 2016, pp. 5628–5631

[11] M. Cliche, D. Rosenberg, D. Madeka, and C. Yee, "Scatteract: Automated extraction of data from scatter plots," in Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases. Skopje, Macedonia: Springer, 2017, pp. 135–150.

[12] JinqiangBai, ShiguoLian, Zhaoxiang Liu, Kai Wang, Dijun Liu, "Smart Guiding Glasses for Visually Impaired People in Indoor Environment," IEEE Transactions on Consumer Electronics, Vol. 63, No. 3, August 2017.

[13] Leo, M,Medioni, G.G., Trivedi, M.M., Kanade, T, Farinella, G.M, "Computer Vision for Assistive Technologies," CVIU. 2017, 154, 1– 15.

[14] World Health Organization (WHO) - Visual impairment and blindness.

[15] Bin Jiang, Jiachen Yang, Houbing Song,"Wearable Vision Assistance System Based on Binocular Sensors for Visually Impaired Users," DOI:10.1109/JIOT.2018.2842229, IEEE INTERNET OF THINGS JOURNAL, 2018.

[16] S.Mahalakshmi1, Veena N, Anisha Kumari, "Visual Assistance for Blind using Image Processing", International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 07, July 2020.

[17] Mohamed Lamine Mekhalfi, Farid Melgani, Yakoub Bazi, Naif Alajlan, "A Compressive Sensing Approach to Describe Indoor Scenes

for Blind People", IEEE Transactions on Circuits and Systems for Video Technology , Volume: 25, Issue: 7, July 2015

[18] Liang Zhang, Michael Towsey, Jie Xie, Jinglan Zhang, Paul Roe, "Using multi-label classification for acoustic pattern detection and Assisting bird species surveys", Applied Acoustics, Volume 110, September 2016