Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.98549777

INNOVATIVE SOLUTIONS FOR VISUAL IMPAIRMENT: DESIGN AND IMPLEMENTATION OF A REAL-TIME ASSISTANCE SYSTEM

Rupali Dhaygude[1], Prof. S. K. Waghmare[2]
[1]M.Tech Student and [2]Asst. Prof. of Department of E&TC - VLSI & Embedded System,
G.H. Raisoni College of Engineering & Management, Wagholi, Pune
(An Autonomous Institute, Affiliated to Savitribai Phule Pune University)
*Corresponding authors: - rupalidhaygude12@gmail.com

Abstract: Visual impairment poses significant challenges to individuals in their daily lives, limiting their independence and access to information. This project focuses on the design and implementation of an innovative real-time assistance system to address the needs of visually impaired peoples. The system aims to enhance their mobility, communication, and overall quality of life through technological advancements. The proposed system utilizes cutting-edge technologies such as computer vision, machine learning, and wearable devices to provide real-time assistance to visually impaired peoples. It employs a combination of image processing algorithms and deep learning models to recognize and interpret the surrounding environment, enabling users to navigate their surroundings more effectively and safely. The core functionalities of the system include object detection and recognition, obstacle detection and avoidance, text-to-speech conversion, and indoor localization. Through the integration of a wearable device equipped with sensors and cameras, the system captures and processes visual information in real-time, generating auditory cues and providing spoken instructions to the user. The implementation of the assistance system involves the development of a custom hardware setup, software algorithms, and user-friendly interfaces. The system is designed to be portable, lightweight, and non-intrusive, allowing users to seamlessly integrate it into their daily routines. Additionally, considerations are given to accessibility and usability to ensure the system caters to a diverse range of visually impaired peoples. Extensive testing and evaluation are conducted to assess the system's performance, accuracy, and user experience. User feedback and suggestions are incorporated into iterative design improvements to enhance the system's effectiveness and usability.

The results demonstrate that the proposed assistance system provides valuable support to visually impaired peoples, promoting their independence, mobility, and engagement with the environment. The real-time nature of the system allows for prompt response to changing surroundings and dynamic situations, facilitating easy navigation & interaction in both in and outdoor environments.

Keywords:- Visually Impaired People, Object Detection, Machine Learning, YOLO & CNN Algorithm, Text Classification & Analysis, and Voice Assistance, Image processing etc.

I. INTRODUCTION

A. OVERVIEW:

The According to World Health Organization (WHO) statistics, the number of blind people is estimated at 1.3 billion, of which 36 million were blind in 2018.By 2019 a total of 2.2 billion

people suffers from some form of visual impairment [1]. According to these statistics, the blind and visually impaired community is increasing yearly. Engaging in day-to-day activities without hazel is an extremely difficult task for a visually impaired/ blind person. It becomes more difficult when it requires traveling through unfamiliar locations without a close companion to assist them along the way. Guide dogs are used in assisting visually impaired persons, but it is not easy to get a trained animal due to the high cost. Furthermore, traveling in familiar environments without help could also be challenging since the dynamic situations along the way cannot be predicted earlier, and responding to those situations in real-time is not possible for a blind per Blind people navigate without a clear visual map about the obstacles in their path. Therefore, it is not possible to take precautions to avoid such obstacles similar to a normal person with good vision. This study focuses on improving the independent navigation of the blind and ensuring their safety while navigating.

B. VISUAL IMPAIRMENT

The person having a reduced capability to observe, even with contact lenses, affects badly the right way of accessing information in day to day life. As per the World Health Organization (WHO), there are about one billion people those contain some form of visual disability in the world shows in fig.1, amongst which 36 million are blind, and more than 217 million have reasonable to harsh vision harm. India hosts a third of the world's visually impaired population. But often, they are not equipped with the tools and facilities that can help them shine. Visual impairment is sometime by birth or it's because of some harm that is caused due to accident or disease. Many different terms are used to describe the varying degrees of vision loss an individual may have. All individuals with decreased vision, regardless of the severity of vision loss or blindness can be said as visually impaired.

C. PROBLEM DOMAIN

The question about technology is always that "how the recent technology can help blind so that they can do normal activities without help of the people around them?" During the recent years, Artificial Intelligence has come into existence with goal of answering the quest that arose towards language processing, perception reasoning, planning, natural and ability to move and control object in the real world. Now a day, machine is no more said as equipment, since it always comes with intelligence. Allowing machine to model our world, where aim of researcher is to prepare learning algorithm and make the algorithm understand the scenario present in front of them. To build algorithm seems being simple but, letting algorithm understand the scenario and reaction is quite difficult. It's because, raw input to such learning system is always high dimensional entity. It may be solid object with 3dimensional view. Computer vision and object recognition are the prominent application of machine learning. Image/object recognition is best handled by deep learning technique of AI, which can better serve visually impaired person. Computer vision being field of artificial intelligence, trains computer to interpret and understand the visual world. Many of the communities have come ahead to help the blind by their own application; but still there is need of more accurate system beyond "go-no-go" system. The sensor based technology has proved very effective, but the information about the size and the location of obstacle ahead is not helpful. Research gap identified from the literature is that, there is need to detect drop off on the way, ensure sufficient information of the obstacle ahead and safety of the visually impaired person. Taking a step in

this direction, the paper focuses on using Deep learning technique for assisting blind for day to day activity like roaming around.

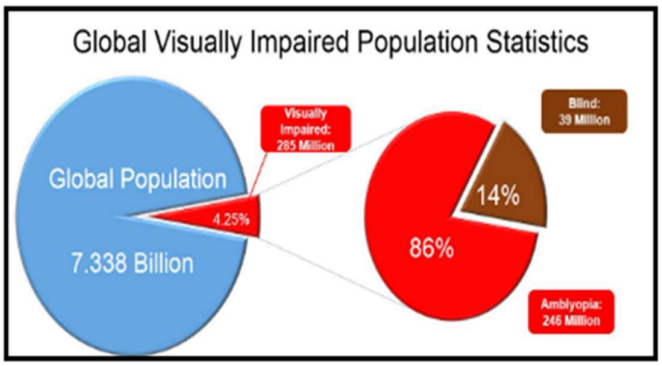

Fig.1: The proportion of visually impaired people in the total world population

II. EXISTING SYSTEM

A real-time system for the benefit of people who are blind or visually impaired must take into account a number of aspects, including hardware, software, and user interface. The system should be designed such that it can provide the visually impaired person with immediate assistance.

To evaluate the collected data and provide the user with the appropriate answer, software components including computer vision algorithms, machine learning models, and tools for natural language processing may be employed. For the program to provide rapid help, it must be efficient and quick to react.

The design of the user interface is equally crucial. The interface should be easy to use and understand, especially for those who are blind. The system should provide the user with audio input as well as numerous non-visual feedback kinds to assist them navigate their environment. Hardware, software, and user interface design must all be carefully taken into account while developing and putting into use a real-time system to aid persons with visual impairments. Making sure the system is effective, responsive, and user-friendly is crucial.

III. RELATED WORK

Visually impaired people face a lot of problem. This paper is to provide assistance for visually impaired. This system combines techniques that are based on sensors along with computer vision concept to result an economically feasible solution. This system uses an algorithm known as novel obstacle based on image depth information and fuzzy logic. Through these techniques, visually impaired people would be able to avoid the front obstacles by detecting the objects in front of them. This system helps in assisting 6 visually impaired people indoor and outdoor. The hardware requirements of the system include camera, GPS, WiFi, microphone, compass, gyroscope and microcontroller. At software level, it requires multisensory data and computer vision approaches.

The system provides 100% obstacle avoidance with 96% detection accuracy. It helps in safe traversal and gives a high performance. The system is considered to be more reliable, simple, wearable and economically accessible. This system has got few disadvantages also which comes into note. The system would not be compatible when it comes to cases like detecting

large objects like walls, doors since their sizes play a vital role as they may not be detected in the frame. So there might be difficulties in detecting by finding the difference between backgrounds and foregrounds. Therefore results can be obtained at the best way using the ultrasonic sensors. Apart from thee, ultrasonic sensors may give a good result in terms of reliability, efficiency in obstacle detection. This helps in measuring the distance of obstacle along with the computer vision techniques thereby increasing the accuracy. [1]

This paper aims in independent movement of people with visual impairment. This is done by designing a portable constrained assistance for the visually impaired people to detect the traffic light. The system is designed based on the algorithm known as AdaBoost. This approach is faster and robust in object detection. This system is enhanced along with a flexible parallel architecture on the field programmable gate array (FPGA) platform. The main images and weak classifier's confidence are calculated in parallel. The weak classifiers parameters are trained using AdaBoost algorithm using MATLAB software and then configured on FGPA platform. FGPA is found to be more flexible and consumes less power.

Experiments show that this system will be able to detect the traffic lights in the video at the rate of about 30 frames per second. This task is done on a real time and hence it detects the position of the traffic light in each of the frames in the video in real time and the system is also considered to be an efficient one. Though the system is efficient, it has got few limitations that it takes large time to train the weak classifiers. The future work should be designed by improving the designed system by introducing more types of weak classifiers, in order to achieve a higher detection rate and accuracy. [2]

To provide navigation and orientation, this paper proposes an electronic device known as NavCane. This device helps in obstacle free path for the visually impaired people both indoor and outdoor. This device sends prior information about the obstacle to the person using it without any information overload and the information is sent to the person through actile and auditory methods. It has many components like ultrasonic and wet floor sensors, GSM and GPS modules, a gyroscope, a vibrator motors radio-frequency identifier, a global system for positioning the module and batteries. This system is assessed by 80 visually impaired people and was successful at various scenarios.

Unlike other electronic devices, this NavCane device detects obstacles present in known indoor settings also. It is considered to be a low power consuming embedded device and a low power system. Analysis says that this NavCane improves the obstacle-free performance much more than the white cane. [3]

This paper aims in developing fast deep neural network for the purpose of real time video object detection by concentrating more on knowledgeable training of data and predicted area of interests. This system develops a framework for training of datasets by deep neural network using limited samples along with cross network knowledge projection which helps in improving the performance and also reduces the computational complexity. Here the training process is regulated by learning projection matrix by projecting the knowledge and visual representation of the teacher-level network from its intermediate layer to the intermediate layer of the child level network. This paper focuses on the following:

1. This system proposes architecture for transferring the data knowledge from large pre-trained teacher level network to thinner child level network.

2. It develops a fast method to find out the candidate regions of interest which contain the targeted objects.
3. It establishes an analytical model to calculate the support regions at each convolution layer and integrates this along with the existing object detection framework using deep convolutional neural networks.

Experiments are done to prove that this system reduces network computational complexity 16 times and improves the network performance at a significant margin. [4] Object detection in real time is very difficult on an embedded device due to limited memory and computational power. Therefore this paper proposes a design without reducing detection accuracy. In this work, a light weight object detection method called as Mini-YOLOv3. For the Darknet based backbone, depth separable convolutions and point-wise group convolutions to reduce the size of the parameter in the network. The process of first reducing the dimensions and then increasing the dimension is adopted along with the residual structure. Boundary effect of point wise group convolution is suppressed by channel shuffle. A Multi-Scale Feature Pyramid (MSFPN) is proposed for fusing 3 feature maps in the backbone to generate base feature. Finally, Mini YOLOv3 is tested with MS-COCO dataset and found that Mini YOLOv3 gives more accuracy and ½ the detection time and achieves a high speed and very competitive performance in embedded applications. [5]

You Only Look Once, also known as YOLO is one of the well-known methods for fast real time object detection that is being popularly used. For better results in object detection in a particular environment, Tinier-YOLO concepts is used and proposed in this paper for minimizing the size of the system for better detection results, accuracy and performance. To overcome all the problems faced with Tiny-YOLO, network performance, detection speed and accuracy is focused. One of the main challenges while introducing fire modules in Tiny-YOLO V3 is to find out the total number of these modules and their places in the system. Then it is also important note that difficulties arises in the connectivity style between the modules to obtain better results in terms of detection, accuracy and performance. This system adopts dense connections among the fire modules so that the movement features could be strengthened and also to ensure flow of information between the networks is at the maximum. But further, the model size cannot be reduced because if it is reduced, then detection accuracy may be affected on a large scale. Therefore, the pass through component is used in this system in order to face the problem of reduction in model size. These pass through layers helps in combining the feature maps extracted from the front layers to obtain the proper features. Finally, to reduce the computational cost factors, this system proposes to remove batch normalization from the modules along with the overall performance. [6]

IV. PROPOSED WORK

Real time object detection is widely used Open CV concept that mainly concentrates on identifying, detecting objects belonging to different classes in the input image. Detecting the objects along with its unique features and background details is done using different techniques and methods. One among them is creating bounding boxes so that pixels of all the aspects of the objects are found and detection is done accordingly with certain computations.

The proposed system aims to design and implement an innovative real-time assistance system for individuals with visual impairment. The system leverages advanced technologies to provide effective support in various aspects of their daily lives.

The core functionalities of the proposed system include:

1. Object Detection and Recognition: The system employs computer vision algorithms and deep learning models to detect and recognize objects in the surrounding environment. It utilizes cameras or wearable devices to capture images or video streams, and then processes the visual data to identify objects such as obstacles, furniture, or common items.
2. Obstacle Detection and Avoidance: By analyzing the depth and spatial information from the captured images or sensor data, the system identifies obstacles and generates auditory cues or vibrational feedback to alert the user. It assists in navigating around obstacles and avoiding potential hazards, improving the user's mobility and safety.
3. Text-to-Speech Conversion: The system incorporates optical character recognition (OCR) techniques to convert printed text into speech. It enables users to access written information such as signs, labels, or documents by capturing the text using a camera and converting it into audible speech, which is then relayed to the user through headphones or speakers.
4. Indoor Localization: To assist with indoor navigation, the system utilizes techniques like Wi-Fi positioning, Bluetooth beacons, or inertial sensors to determine the user's location and provide guidance. It helps users navigate through complex indoor environments such as buildings, shopping centers, or public transportation stations.

The proposed system involves the development of custom hardware components, software algorithms, and user interfaces tailored to the specific needs of visually impaired individuals. It focuses on usability, portability, and non-intrusiveness to ensure a seamless integration into the user's daily routine.

V. SYSTEM DESIGN

Proposed system focuses on the idea that deep learning technique can assist visually Impaired for thorny times such as roaming around and will ensure safety and independent mobility to visually- impaired people illustrate in fig.2. The system consists of following flow for being operational.

The main contribution of this research is to propose a novel obstacle detection mechanism to assist blind navigation that targets, High accuracy and real-time obstacle detection. A deep neural network is used for obstacle detection. In the proposed system, a novel approach was adapted that uses data generated from a simulation environment instead of real-world data to train the deep neural network. Simulation platforms are used in various domains as a mechanism of data generation. Similarly, a simulation platform was used here and to generate data a 3D realistic environment was created. A common mechanism that can be used to estimate the distance of both ground-level, above ground-level obstacles. A deep learning-based monocular depth estimation methodology is combined with binary threshholding to identify the obstacle with possible threats to the navigator.

Step 1: If the user wants to do outdoor activity, he has to start the system. Once system is switched on, capturing image through camera module, mounted on goggles (in real-time) will start capturing image.

Step 2: The captured Image will be reshaped into some fixed dimension with help of Computer vision by Pre-processing.
Step 3: Linear Convolution will be applied on the output obtained from step 2, which will understand edge, shape, and boundaries in an image. i.e. It will identifies the grid responsible for detecting image, so that major attributes of bounding box can be understood.
Step 4: Output from step 3 is given to fully connect dense layer, which will detect the object/target in an image. At each grid cell, 5 bonding boxes will be detected. The predicted output depends on down sampling the dimension of input image. The function will be Relu, size of hidden layer will be restricted to 2000 images (which are identified on the basis of places where visually impaired regularly visits).
Step 5: Detected image will be initially captioned [4, 5] and caption will be converted into voice command, conveyed to user using natural language Processing.
Step 6: If the user wants to move in the direction depending on caption, Voice based guideline will be provided to help reach destination using Sensor, which tells distance of object from the Device/user.

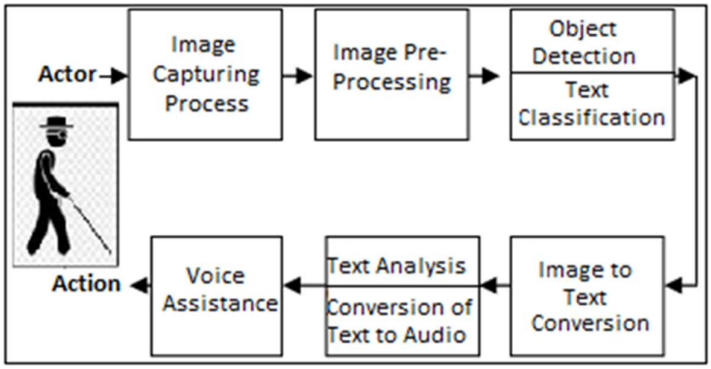

Fig.2: System Architecture Diagram

VI. RESULT ANALYSIS

The experimental results prove that the proposed work can help visually impaired people by notifying them of their surrounding objects and absolute location. Experiments were conducted in many different settings and almost all the objects that were present at that time were detected and notified to the user. The average computational time required for detection was 2000 ms. By conducting these experiments, the computational time varies according to the number of objects present is observed. Computational time increases as the number of objects present increases. More than 75% of the objects are detected and recognized accurately at a time.
Accuracy of the proposed system: The proposed system shows upto 95% accuracy in detection of objects. The model is trained in such a way that it detects objects correctly with the extracted features and boundary boxes. Detection results of the trained YOLO & CNN algorithm and for sample 6 objects are shown in table 1 and 2.

Table.1: Result Analysis

Algorithm	Precision	Recall	F1-Score
YOLO	0.92	0.87	0.89
CNN	0.85	0.92	0.88

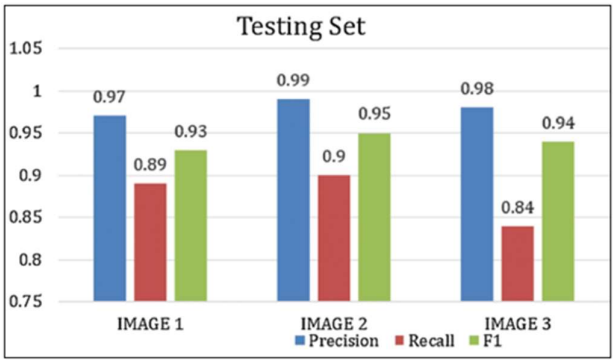

Fig.5: Results of Object Recognition

Table.2: Result Analysis

NAME OF THE INPUT	NO OF INPUTS GIVEN	NO OF TIMES CORRECTLY DETECTED	CLASSIFICATION ACCURACY IN %
Person	30	30	100%
CAR	40	38	95%
BOTTLE	20	18	95%
CHAIR	30	27	96%
POTTED PLANT	20	18	90%
SOFA	30	28	98%

VII. CONCLUSION

The above system implements object detection method by detecting objects with high level confidence. Use of pre trained model is much more efficient since it reduces the computational time and computational costs and gives a more efficient system. In such pre trained model, we load the file, and then preprocessing is done. After this process the input is given in real time and the output is obtained. This system helps in detecting multiple objects in real time. Therefore separate training of images is not necessary since the pre trained model includes all datasets and training of datasets as well. At the end, the detected objects names are displayed as text which is then converted to speech. Since it has an extra voice feature, this system would be useful for visually challenged or deaf and dumb people. Therefore, it has a wide variety of applications.

ACKNOWLEDGMENT

To the researchers and publishers, I'd want to convey my appreciation for making their work available. I also like to thank the college administration for providing the essential materials as well as the guide and reviewer for their informative comments.

REFERENCES

[1] W.-J. Chang, J.-P. Su, L.-B. Chen, M.-C. Chen, C.-H. Hsu, C.-H. Yang, C.-Y. Sie, and C.-H. Chuang, “An AI edge computing based wearable assistive device for visually impaired people zebra-crossing walking,” in Proceedings of the 2020 IEEE International Conference on Consumer Electronics (ICCE), Jan. 2020

[2] Rohilla, Yogesh Parihar, Vipul and K Rohilla, Kusum. (2020). Ultrasonic Sensor based Smart Cap as Electronic Travel Aid for Blind People. 10.1109/IC-SSIT48917.2020.9214226

[3] Y. Zhang, X. Zhang, and W. Lu (2020), "A survey of assistive technologies for the blind and visually impaired," in Proceedings of the 9th International Conference on Information Technology and Applications, Sydney, Australia, July 2020, pp. 349-354.

[4] R. K. Prusty and P. Patra (2019), "Assistive Technology for Blind and Visually Impaired People: A Review," in Proceedings of the 2nd International Conference on Computing for Sustainable Global Development, New Delhi, India, March 2019, pp. 2728-2731.

[5] Sunil S. Khatal, Dr. Yogesh Kumar Sharma (2019), "Healthcare Patient Monitoring using IoT and Machine Learning." IOSR Journal of Engineering (IOSR JEN), ISSN (e:) 2250-3021, ISSN (p): 2278-8719

[6] M. M. Islam, M. S. Sadi, K. Z. Zamli, and M. M. Ahmed, "Developing walking assistants for visually impaired people: a review," IEEE Sensors Journal, vol. 19, no. 8, pp. 2814-2828, Apr. 2019

[7] A. Kumar, S. Kumar, and S. Kumar (2019), "Assistive technology for visually impaired people: A review of recent advances," in Proceedings of the 4th International Conference on Advances in Computing, Communications and Informatics, Jaipur, India, September 2019, pp. 651-656.

[8] S. Degaonkar, M. Gupta, P. Hiwarkar, M. Singh, and S. A. Itkar, "A smart walking stick powered by artificial intelligence for the visually impaired," International Journal of Computer Applications, vol. 178, no. 32, pp. 7-10, Jul. 2019

[9] S. Pruthvi, P. S. Nihal, R. R. Menon, S. S. Kumar, and S. Tiwari, "Smart blind stick using artificial intelligence," International Journal of Engineering and Advanced Technology (IJEAT), vol. 8, no. 5S, pp. 19-22, May 2019

[10] Croce, L. Giarré, F. Pascucci, I. Tinnirello, G. E. Galioto, D. Garlisi, and A. L. Valvo, "An indoor and outdoor navigation system for visually impaired people," IEEE Access, vol. 7, no. 1, pp. 170406-170418, 2019.

[11] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Virtual-blind-road following-based wearable navigation device for blind people," IEEE Transactions on Consumer Electronics, vol. 64, no. 1, pp. 136-143, Feb. 2018.

[12] C.-W. Lee, P. Chondro, S.-J. Ruan, O. Christen, and E. Naroska, "Improving mobility for the visually impaired: a wearable indoor positioning system based on visual marker," IEEE Consumer Electronics Magazine, vol. 7, no. 3, pp. 12-20, 2018

[13] S. J. Rizvi, S. R. Karim, and M. A. H. Akhand, "Assistive Technology for the Visually Impaired People: A Review," International Journal of Computer Science and Mobile Computing, vol. 7, no. 6, pp. 63-75, June 2018.

[14] S. Ren, K. He, R. Girshick, and J. Sun (2018), "Faster R-CNN: Towards real-time object detection with region proposal networks," in Proceedings of the Advances in Neural Information Processing Systems, Montreal, Canada, December 2018, pp. 91-99.

[15] Zubov, "A smart city assistive infrastructure for the blind and visually impaired people: a thin client concept," Broad Research in Artificial Intelligence and Neuroscience (BRAIN), vol. 9, no. 4, pp. 25-37, Nov. 2018.

[16] Z. Wu, Y. Jiang, and J. Yang (2018), "Deep face recognition: A survey," arXiv preprint arXiv:1804.06655, April 2018.