Object Detection System for Visually Impaired

Aditya, Amar Singh, Chirag Pareek
Poornima Institute of Engineering and Technology, Computer Science, Jaipur, India

Abstract - Portable assistive technology systems are developed to enhance the capability of persons with disabilities. One of the most important senses for humans is vision. Vision is the most important sense which helps us in understanding the perception of the surrounding environment. Visually impaired persons face a lot of difficulty in understanding the perception around them, particularly in outdoor environment where objects are continuously changing and moving from one place to another. Object detection solutions would greatly assist visually impaired persons in avoiding from the barriers which they face in their daily routine life. The aim of the object detection system is to provide a simple, user-friendly, handy, economical and efficient solution for the visually impaired persons. Motive of this system is to develop a solution that detects the objects present using camera, as the input device in real time and communicate the same to the user using smartphone through headphones. The system would be using an audio device such as speakers or headphones in providing the information about objects to assist the visually impaired persons. The proposed system helps in identifying and avoiding the objects both in outdoor and indoor environments that affect day to day life activities and occupational performance of the visually impaired persons. The information about the objects in surrounding environment would be very much helpful to the visually impaired persons in their daily life.

Keywords- Object detection system, visually impaired, Blindness, Assistive system, Computer vision, Image recognition

1.INTRODUCTION

There are several organizations that take cognizance of the welfare of the blind and visually impaired. [1] The World Health Organization has reported that there are 2.2 billion people who are either visually impaired or completely blind through a recent study conducted by them. [2] In India, the Ministry of Health and Family Welfare of the Government of India has conducted a survey called “The National Blindness & Visual Impairment Survey India 2015–2019”. It was conducted for approximately 85,000 people. The results of the survey showed the estimation of the number of people with vision problems. The results are presented for the age group of 50 and above. The prevalence for blindness is 1.99% and the prevalence for visual impairment is 13.76%. The survey data supports that there is an astonishingly large number of people with vision related problems. The numbers of people that need support are huge and a simple, efficient system is needed in place for their assistance A successful execution of a supporting system will have a large impact on the ways and livelihood of the people with visual impairment.

The challenges that the completely blind and visually impaired people face directly impact the quality of their lives and also how do they perform their daily activities. The traditional method is that of a supportive guide. A person is needed who can efficiently guide them in travelling from one place to another. The white cane is a helpful aid in navigating with obstacles in near proximity of the person. It can only help identify obstacles that are very close to the person. This brings up the issue of safety of the blind and visually impaired. Independent mobility also is matter of concern. An object detection system is the need of the hour to combat the safety issues and also enable them to have mobility that is not dependent on anyone. The object detection system has to be a simple and most importantly user friendly. It also should be budget friendly so that it can be afforded by them.

We are fortunate to have technologies that can make a huge difference in the life of the blind and visually impaired. Some advanced technologies like computer vision, object detection and object identification are available in today’s world. The camera of a smart phone can be used as the vision apparatus which helps in the object detection system. It then detects the objects that are in front of the camera. Yolo V3 is an object detection model that can be used for detection of objects in a real time video. The detected object can be identified and then communicated as an audio message to the user. It can be done either on the speaker or using a headphone. This application is designed to be used on an Android based smartphone. Smart phones not only help in the regular telephonic

conversations but also help the blind and visually impaired if the smart phones have a camera. The primary goal of the application is to enable the user identify the object in front of the camera. The computer vision techniques supported by the object detection systems can enhance the quality of lives of the blind and visually impaired by many folds. We are proposing a system that is simple and user-friendly. It is not only handy but also economical and efficient. This overcomes the issues of an expensive, additional sophisticated system. The use of smartphones is not just for the super advanced applications. A common man is able to use a smartphone with a camera in a budget friendly manner. The proposed system will not need any additional hardware other than those are already available in a smartphone. Smartphone comes with an integrated camera. It can be used to capture the real time video. This in turn is used to detect the objects. The identified object is communicated to the user as an audio instruction. This paper highlights the computer vision techniques and presents the overview of the object detection system, which can enhance the quality of lives of the visually impaired persons.

2.LITERATURE SURVEY

Various aids have been designed to develop a system that provide visual assistance and which help the persons facing vision problems in their daily routine life by communicating about the obstacles present in front of them. The following literature review presents the outcome and the limitations of some of the existing assistive solutions for blind and visually impaired.

In [3], authors proposed an android based system which captures the images through the smart phone camera and communicates the user about the object detected, the distance and direction of the object to the user via audio instructions using speakers or headphones. The proposed android based system provides the perception of the surrounding environment and enables them to navigate independently. The system has three modules: Object detection module, Distance calculation module and Direction. The object is classified and module isolate different objects present in the image is done in the object detection module. Object detection is done using TensorFlow's object detection API. Open-source computer vision library (OpenCV) and Triangle similarity law are used to determine the distance of the object. Using 'x' coordinates of the bounding box of the object, the direction 'Left', 'Centre' and 'Right' is determined and communicated the same to the visually impaired persons. The proposed system was able to detect with an accuracy of 87%. Some of the limitations identified are the distance calculation functionality is yet to be included in the system. The proposed system can be retrained to detect more objects.

The aim of the system [4] is to minimize the need of dedicated and wearable devices by providing an android smartphone-based application system that assist visually impaired persons in recognizing objects and in moving around. The R-CNN algorithm is used for classification of real time objects. Image processing algorithms are used to process the video captured by the inbuilt camera of the smartphone. After objects are scanned and segmented, Fast R-CNN algorithm is used to identify multiple regions. Later, using Google API text to speech conversion is done to communicate the same to the user. The proposed system revealed that better accuracy is achieved in detecting with low latency by SSDMobileNetV2 model and performed 35% better in speed than MobileNetV1 when paired with TensorFlow Lite. The system has a limitation to focus on better performance in detecting multiple objects with better accuracy.

The goal of the proposed systems in [5–7] is to develop an Android application which provides necessary information about the objects and environment to the visually impaired persons. The proposed system minimizes the usage of the conventional dedicated devices and wearable devices. The paper also provides an overview of the object detection and significance of ORB over SIFTS and SURFS in different cases. The video is captured by the camera then the video is divided into frames. Based on the key matching in the database the image edge detection is processed then text to speech conversion is done using synthesizer which instructs the user.

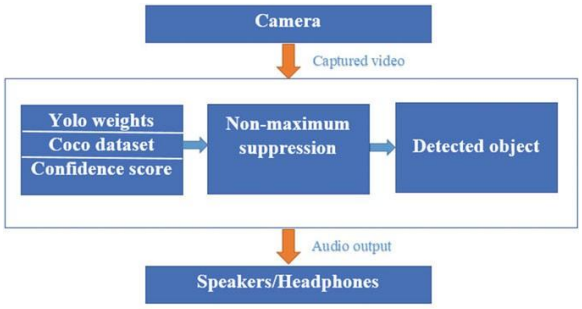
The objective of the papers [8, 9] is to review the different navigation solutions available for visually impaired persons that work both in indoor and outdoor environment. A thorough review of the existing navigation solutions in both indoor and outdoor is made. In Assistive Travel Aids, the author classified available aids into three major categories: electronic orientation aids (EOAs), position locator devices (PLDs) and electronic travel aids (ETAs). Electronic

orientation aids (EOAs)—used in finding the navigation path to the visually impaired persons. Position locator devices (PLDs)—both Global positioning system (GPS) and Geographic information system technologies are used to guide user navigation. Electronic travel aids (ETAs)—help visually impaired persons to avoid obstacles. Some of the limitations observed are in EOAs, it is difficult to add complex computing devices. Using PLDs, the device cannot detect objects and ETAs can be improved in the detection range of the obstacles. The study found lacking of some core features like portability, comfort ability, adaptation time and user learning which are very helpful and important for independent mobility in the existing navigation systems for visually impaired persons. To develop an intelligent assistance system [10] composed of wearable smart glasses, mobile device app, an intelligent walking stick and an online information platform that assist in object detection. The components wearable smart glasses, mobile device app, an intelligent walking stick and an online information platform are composed in the system. Whenever the smart glasses detect objects it reminds user by smart walking stick. If the user falls down then the related information is stored online and shares the same information with their family members. Apart from object detection assistance, the major advantage in using this device is that whenever user falls down or any collision event happens then related information is sent to family members. A few limitations observed are features like recognizing front images like traffic signs and to develop intelligent walking guide are yet to be included.

The aim of the device in [11] is to provide the safe navigation assistance and perception of the environment in the real time. The obstacle avoidance system in the device learns from the RGBD, pilots choice-of-action input and semantic map, it ensures to provide safe and appropriate feedback about the surrounding environment to the visually impaired. A few limitations of the device are the sonar and bump sensors are yet to be incorporated to ensure the safety of the user in exception cases. Matusiak et al. [12, 13] explained about the main features of the smartphone Android application software developed to assist visually impaired persons. The software has three modules, where the main module in the system recognizes and matches the scanned objects with the objects stored in the system database. The other two modules are capable of identifying direction of the major colour's regions in the captured images. The SIFT (Scale-Invariant Feature Transform), a computer vision algorithm is used in the proposed application to describe image features. Devi et al. [14], Bhatnagar et al. [15] explains about determining obstacles and potholes and notify the user through audio output and to share active location of the user with the care taker. The obstacles detection part is carried out by ultrasonic sensors attached to the device interacts with the Microcontroller then the respective voice commands are given to the user. Potholes are identified by the infrared sensors and vibrates the aid to let the user know. GSM gathers the active location of the user [16, 17]. Whenever user press panic button then emergency assistance request is made. Apart from travelling assistance, emergency request is also included. The limitation observed in the system is since all components are placed on stick, weight may be a constraint and repeated hitting of stick on ground may cause malfunction of the device.

3. PROPOSED SYSTEM

Visually impaired persons face many challenges in identifying objects and performing the day to day activities. Most of the time they are dependent on others in moving from one place to other. They face a lot of difficulty especially when moving in outdoor environment, where objects are continuously moving. This object detection system would assist the visually impaired persons by providing the perception of the surrounding environment or position of the objects. This device can assist visually impaired persons in avoiding the obstacles in both indoor and outdoor environment. The proposed system also helps the user in identifying the objects around them. It would minimize the visually impaired person's difficulties and help them lead a quality life.

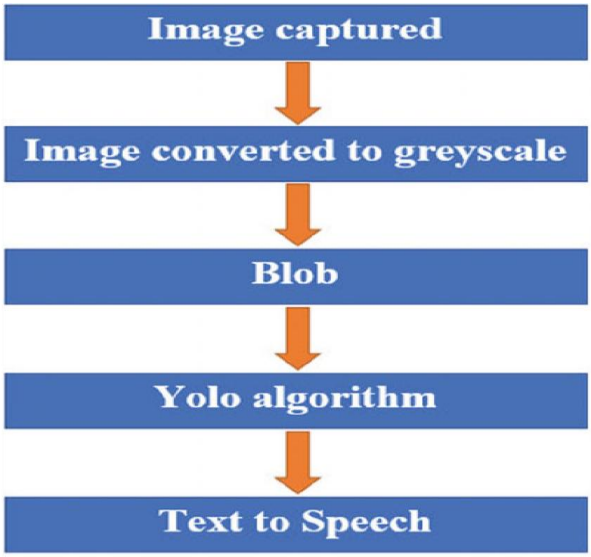


Proposed System

The proposed system aims to be simple, user-friendly, handy, economical and efficient solution. Most of the existing assistive systems are highly sophisticated and expensive. Smartphones are widely used these days and the usage of smartphones by all persons has become very usual in the recent past. Therefore, this system uses all the advanced built-in features of a smartphone. Smartphone's integrated camera is used to capture the real time video to detect the objects present and headphones or smartphone's speaker is used to communicate to the user through audio instructions. The proposed application can easily be accessed by the visually impaired persons. The Yolo V3 (You only look once) algorithm is used to detect the real time objects captured from the continuous streaming by the smartphone camera. It is considered as one of the most powerful real time object detector algorithms. Unlike other algorithms like R-CNN and Faster R-CNN which examine several regions of the image to identify objects, Yolo passes image or video only one time through its network and uses a unique neural network using the characteristics of the complete image to predict multiple boxes containing an object. This is a significant feature in Yolo, which reduces the processing speed when compared with other algorithms. Processing speed plays a key role particularly while detecting objects in real time video stream. OpenCV provides a function that facilitates image pre-processing for deep-learning classification. Group of connected pixels in an image that share common properties (Blob ex: Grayscale value) of each frame captured is identified. Bounding boxes are created for each object identified and based on the Yolo V3 pre-trained weights, confidence score and coco dataset each object is processed and labelled. The coco dataset contains all the objects or class names on which the model is trained. Non-maximum suppression (NMS) uses function called "Intersection over Union (IoU)" which is used to determine the best boxes. In order to select the best box, NMS follows three steps. 1. It selects the box with highest score. 2. Computes the overlap with other boxes, removes the overlap which has more than a certain threshold. 3. Iterates the process until there are no more boxes with a lower score than the currently selected box. The above NMS steps will remove all the boxes that have a large overlap with the selected boxes. Only the best boxes remain. After detecting the object by the proposed model, it produces the text format output. Then using text to speech converter, the label of the object detected is converted to speech and communicated the same to the user through headphones.

4. METHODOLOGY

The object detection, recognition and communicating speed of the object to the user should be very fast. Since the proposed system is intended for visually impaired persons, the response time of the system plays a key role. Delay in communicating about the obstacles to the user will not meet the purpose. To overcome processing speed and delay issues, Yolo V3 algorithm is used in the proposed system. Yolo V3 algorithm is faster than many other real time object detecting algorithms.



Methodology

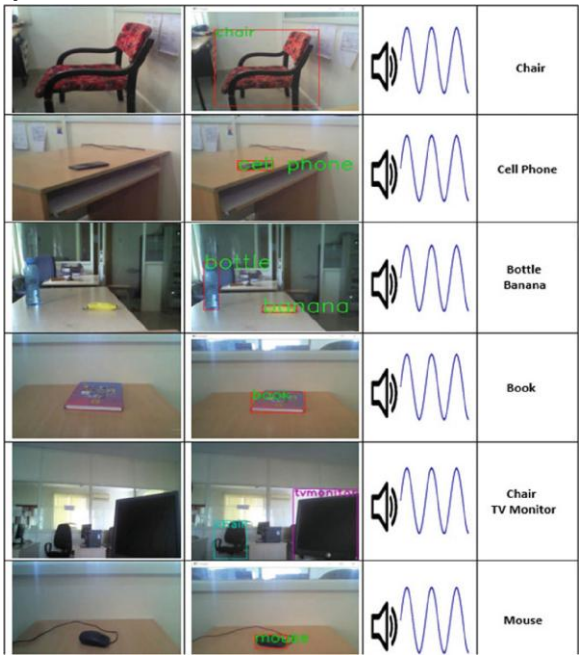
- Captures the objects in the surrounding environment in real time and this is done by using the camera of the smartphone. The video will get divided into frames.
- The RGB colour image is converted into greyscale image.
- Group of connected pixels in an image that share common properties of each frame captured is identified (Blob).
- Based on the Yolo V3 pre-trained weights, confidence score and coco dataset each object is processed and labelled (Yolo V3 algorithm).
- Using text to speech converter, the label of the object detected is converted to speech

5. RESULTS

The main motive of the proposed system is to assist visually impaired persons by providing the perception of the environment, which helps them in avoiding obstacles or barriers and in moving from one place to another. The goal to provide a simple, user-friendly and handy solution is achieved. The proposed system is capable of detecting the objects present in the surrounding environment with good speed and accuracy. It detects objects effectively in both outdoor and indoor environment. The system is able to successfully detect the multiple objects present in the surrounding environment and communicate the same to the user in audio through headphones or speaker. The proposed system is tested in detecting objects in indoor environment, outdoor environment and objects which are more than 10 m from the camera. The system is capable of detecting the objects in the surrounding environment and provide audio output to the user. The performance of the application in the above mentioned three categories is satisfactory. The results captured are presented below.

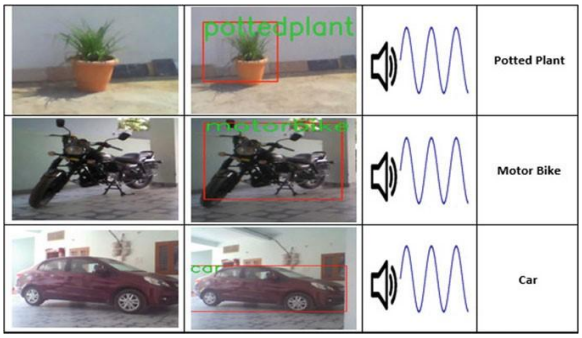
5.1 Indoor Environment

The following figure show the objects captured by the built in camera and the detected objects with the labels and the respective audio output of the proposed system.

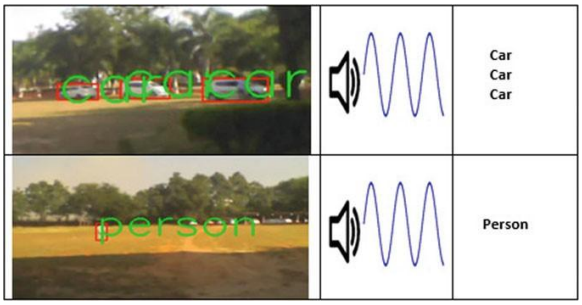


5.2 Outdoor Environment

The following figure shows the objects captured by the built in camera and the detected objects with the labels and the respective audio output of the proposed system.



The proposed system successfully detected the objects which are far away in the outdoor environment with good accuracy. Following are the objects that are detected in the outdoor environment. The system was successful in detecting the objects which were more than ten meters away in the outdoor environment.



6. FUTURE SCOPE

The performance of the system can be further improved by re-training on larger datasets. Providing information about direction of the object will help the user to understand the perception better. The system can be improved in detection time and better accuracy

CONCLUSION

In recent years, there were many assistive solutions developed for the visually impaired persons to provide assistance in detecting objects present in their surrounding environment and in moving from one place to another. But most of the existing solutions are expensive, highly sophisticated, difficult to handle, designed as a dedicated aid, require training etc. Our primary goal is to provide an assistive solution to the visually impaired persons which is simple, user-

friendly, affordable and handy and assist the visually impaired persons in understanding the surrounding environment and help them in moving from one place to another place independently.

The real time video stream is captured by the smartphone camera and then the label of the object detected using object detection algorithm is communicated to the user in speech through the speakers or headphones. The audio output communicated to the user would help them in performing their day to day activities and lead a quality life.

ACKNOWLEDGEMENT & FUNDING STATEMENT

Authors would like to thank the Deanship of Scientific Research at CMR College of Engineering & Technology, Hyderabad for supporting this work under TIDEDST Project No. SEED/TIDE/2018/55.

Conflicts of Interest:

The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] World Health Organization, "World Report on Vision"https://www.who.int/publicationsdetail/world-report-on-vision.

[2] National Programme for Control of Blindness & Visual Impairment (NPCBVI), Government of India "The National Blindness & Visual Impairment Survey India 2015–2019". https://npcbvi.gov.in/writeReadData/mainlinkFile/File341.pdf

[3] Badave A, Jagtap R, Kaovasia R et al (2020) Android based object detection system for visually impaired. In: International conference on industry 4.0 technology 2020, pp 34–38. https://doi.org/10.1109/I4Tech48345.2020.9102694

[4] Anitha J, Subalaxmi A, Vijayalakshmi G (2019) Real time object detection for visually challenged persons. Int. J. Innov. Technol. Explor. Eng. 8:312–314

[5] Jadhav PSP, Tomy S, Jayswal SS et al (2016) Object detection in android smartphone for visually impaired users. 5:332–334. https://doi.org/10.17148/IJARCCE.2016.51171

[6] Khairnar DP, Karad RB, Kapse A et al (2020) PARTHA: A visually impaired assistance system. In: Proceedings of 2020 3rd international conference on communication system computing IT application CSCITA 2020, pp 32–37. https://doi.org/10.1109/CSCITA47329.2020.9137791

[7] Sait U, Ravishankar V, Kumar T et al (2020) Design and development of an assistive device for the visually impaired. Procedia Comput. Sci. 167:2244–2252. https://doi.org/10.1016/j.procs.2020.03.277

[8] Kuriakose B, Shrestha R, Sandnes FE (2020) Tools and technologies for blind and visually impaired navigation support: a review. IETE Tech Rev (Institution Electron Telecommun EngIndia) 1–16. https://doi.org/10.1080/02564602.2020.1819893

[9] Manjari K, Verma M, Singal G (2020) A survey on assistive technology for visually impaired.Internet of Things 11:100188. https://doi.org/10.1016/j.iot.2020.100188

[10] Chen LB, Su JP, Chen MC, et al (2019) An Implementation of an intelligent assistance system for visually impaired/blind people. In: 2019 IEEE International conference on consumer electronics ICCE 2019, pp 4–5. https://doi.org/10.1109/ICCE.2019.8661943

[11] Lin Y, Wang K, Yi W, Lian S (2019) Deep learning based wearable assistive system for visually impaired people. In: Proceedings of 2019 International conference on computer vision workshop ICCVW 2019, pp 2549–2557. https://doi.org/10.1109/ICCVW.2019.00312

[12] Matusiak K, Skulimowski P, Strurnillo P (2013) Object recognition in a mobile phone application for visually impaired users. In: 2013 6th international conference on human system interaction (HSI 2013), vol 17, pp 479–484. https://doi.org/10.1109/HSI.2013.6577868

[13] Matusiak K, Skulimowski P, Strurnillo P (2013) Object recognition in a mobile phone application for visually impaired users. In: 2013 6th International conference on human system interaction (HIS), pp 479–484. https://doi.org/10.1109/HSI.2013.6577868

[14] Devi RA, Uthaman I, Shanthanam RR (2020) An IOT Security based electronic aid for visually impaired detection with navigation assistance system, p 6
[15] Bhatnagar V, Chandra R, Jain V (2019) IoT Based alert system for visually impaired persons. Springer, Singapore
[16] Anand S, Kumar A, Tripathi M, Gaur MS (2019) Human face detection enabled smart stick for visually impaired people. Springer, Singapore