A Study of Object Detection Techniques to Assist Visually-impaired People

Maheshwari Kunte
Maharashtra Institute of Technology, Dr. B.A.M. University
Chatrapati Sambhajinagar, India

Sarita Sanap, PhD
Maharashtra Institute of Technology, Dr. B.A.M. University
Chatrapati Sambhajinagar, India

ABSTRACT

Vision is an integral aspect of human life. A blindness caused due to natural or accidental way stands as a major hindrance in progress of individual. It is of utmost importance to provide the significant help to blind individual for navigation and take informed decision. Several attempts have been made to design system that are helpful in multiple ways to the visually-impaired individuals. Recently, surge in artificial intelligence field have opened new avenues for the advanced research. This paper systematically high- lights the work done in the field. This paper can act as a good starting point for the researcher. It also identifies the gaps and suggests few possible remedies that can help blinds. It focuses on the mobile application made for the assistance of the blind individuals. It summarizes the real time test results and the timing information of the application when detecting the surrounding objects.

General Terms

Deep learning, artificial intelligence

Keywords

Object detection, blind assistance, accuracy, navigation

1. INTRODUCTION

Almost every country has a portion of society that is visually im- paired or blind. Person who is not capable of discerning minute de- tails of the scene being seen through eyes is visually impaired (VI) whereas blind person can not see anything and has zero vision. A navigation becomes major challenge for VI people. It is believed that human receive more than 75% information from the environ- ment through eyes. As a consequences VI people face major ob- stacle during daily routines. They are generally helped by external beings or tools for their normal functioning. The problem worsen with aging. Therefore this area has been of prime importance in re- searchers list. They meticulously look for innovative solutions. Dif- ferent areas of technology have helped researchers to create those solutions that can help VI's to overcome the limitations to certain extent.

The challenges faced by these individuals can be addressed by taking leverage of artificial intelligence (AI) techniques. Figure 1, shows the major challenges faced by the visually impaired individuals Due to enhanced digital platforms and services AI has seen tremendous growth over recent years. This study has gath- ered information of few techniques that have helped VI individuals to overcome their limitations. They always poised with difficulties while moving around. The systems that are available has many advantages and disadvantages. This paper highlights such systems, challenges associated and benefits offered by them. The main ob- jective of this paper is as follows:

—To understand the challenges faced by visually impaired individ- uals.

—To study available system that can help VI to overcome those challenges.

—To discuss the recent trends in technology that helpful for the welfare of VI.

The paper is organized as follows. Section II, discusses the related work, while section III, gives the idea of role of AI in VI assistance. Section IV, concludes the findings of paper.

2. LITERATURE SURVEY

To help VI individual tremendous efforts have been taken in re- cent years. However most of the work orients around in-house navigation assistance. This section explores many systems that have helped VI for their safety and upgrading the living standards. Range of solutions spans a broad spectrum starting from smart-phone technology to cloud-assisted services. State-of-the-art (SOTA) computer vision techniques have also been evolved over the years that are providing upper edge to VI individuals. Proposed systems have associated limitation and advantages with it. This sec- tion critically examines few important studies that have been taken for the VI assistance.

An attempt has been made into [1], to create a device that uses sim- ple components like ultrasonic sensors, Arduino board, buzzer and GSM module. Detection is done by sensor, processed by the com- puting platform, communicated and alerted by GSM and buzzer, respectively. This stick critically provides a solution to the blind in- dividuals. It has been claimed that device provides assistance while crossing roads, provides navigation assistance however comes with certain challenges like it is not self sustained to identify the exact obstacles, simultaneous detection of multiple objects etc. In another interesting study, careful use of Raspberry Pi and its peripherals is done to help visually impaired individuals while navigation and related simpler tasks. Obstacle detection has been tested on cer- tain objects like chair, person and most common object that can be found in daily occurrences. The name given to the system was "SMART EYE" due to its special capability to provide assistance to the visually impaired individuals during the adverse condition and when they are lonely.

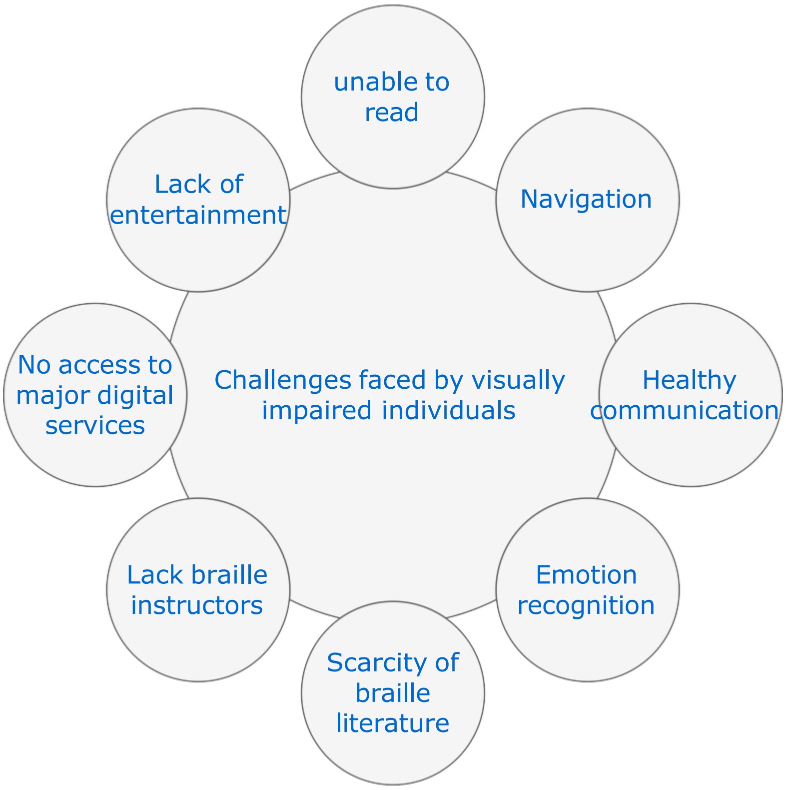


Fig. 1: Challenges faced by visually impaired individuals

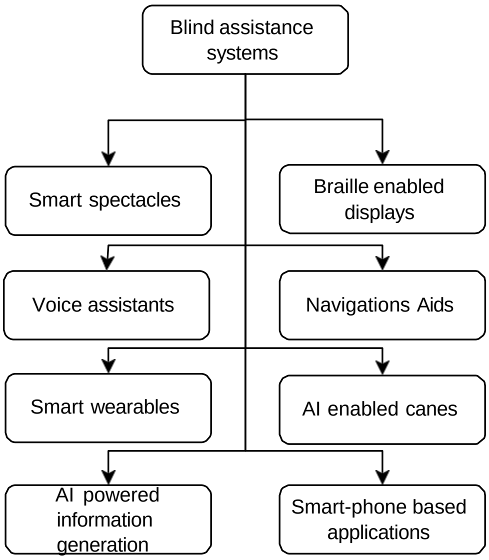
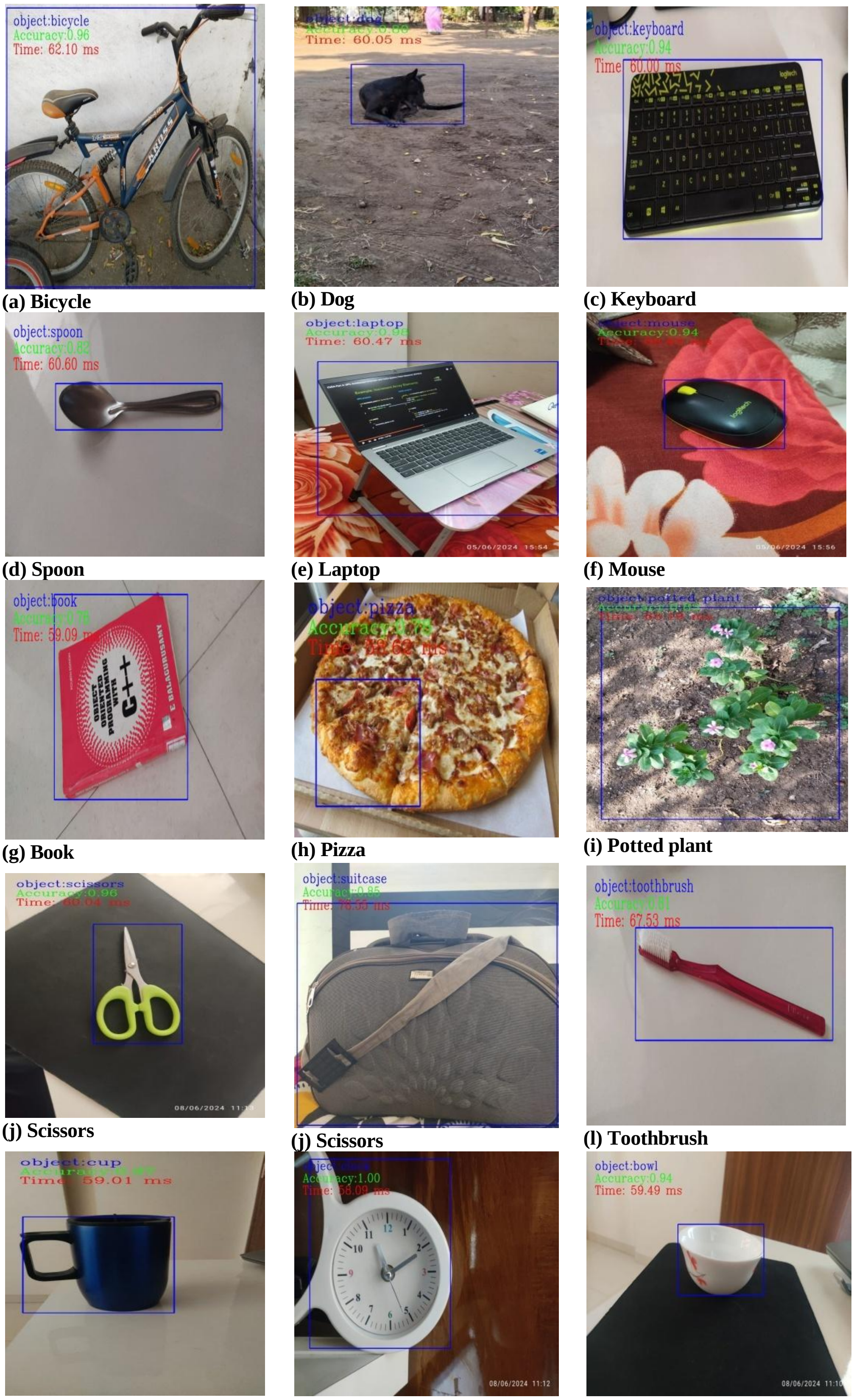


Fig. 2: Various systems proposed for VI assistance



(a) Bicycle (b) Dog (c) Keyboard (d) Spoon (e) Laptop (f) Mouse (g) Book (h) Pizza (i) Potted plant (j) Scissors (j) Scissors (l) Toothbrush (m) Cup (n) Clock (o) Bowl

Fig. 3: Few objects that are detected with the application.

The system was tested both during night and day-time and claimed that system provides acceptable level of performance in simpler navigational circumstances[2]. It is often possible that the obsta- cles may not always be in linear range of person. The objects that are at angular difference from the visually impaired needs precise detection and hence smoother navigation. Systematic used of state- space model and controller based mechanism with multi-sensor inputs is realized in [3] to provide accurate feedback to the VI. It is very essential to understand the system architecture for VI assistance from different perspective. Based on the interfaces involved for the data acquisition systems can be identified as smart/wearable caps, smart-phones, smart canes and other additional devices. If the systems are examined based on the types of data acquisition tools, there can be major devices types as mentioned here. There are water sensor to know the exact level or depth associated with the present water, radar sensors to detect the radio/EM waves being radiated from the objects. Ultrasonic sensors to detect the objects, LIDAR sensors that provide more comprehensive information of the obstacles and finally cameras (may be mobile or standalone). Based on the types of feedback the systems are again classified as speech-based, vibration-based, sound/alarm-based and text-based (text-to- speech synthesizer). In [4], a straight-forward AI based model is used for the object detection and have used speech feedback to the user based on the input received. The field of assistive technology explores many option from the environments. Along with smart-canes and smart application another obvious option available is smart cap. The study proposed in [5] has given a excellent system that has more dynamic features to offers with its counterparts. A low cost and effective solution operates in mixed model and offer services like family-person detection, object-detection and social help detection as well. This system leverage the help form recent AI-trends, also it engages sensors, small computing platform has affordable price range of |5-6K.Figure 2, shows the types of systems used for the VI assistance. Somewhat similar approach was taken in [6] to help VI for comfortable navigation. A system that utilizes RPI, a sensor and camera has been transformed into a de- vice that detects a within the rage of 1.5 m and generates auditory feedback to the person. Tesseract and EAST open source li- braries were utilized for the development. It was claimed that the system is very efficient and cost effective as well. Object detection has been constantly the center point when the AI based VI assistance systems are discussed. It becomes imperative to check the performances and efficacy of the AI model that are being used in the applications. In [7] various models that are trained on COCO dataset are studied. The model then integrated into RPI 4B and fit- ted on the wearable caps. Such systems offer a great deal of navigation assistance, however the impact of wearing such electronically equipped device continuously may cause adverse impacts on the

VI. Though study does not imply to any such hazards it is always helpful to check the aftermath associated with similar systems. An interesting approach proposed in [8] pays attention to the quality of life factor for senior and blind person as well. It has made the arrangement of putting sensors on overall body so that estimation of potholes becomes easier, also it have used the AI based models and integrated with hardware such that the overall cost of the device becomes affordable with maximum possible navigation assistance. Four pin security feature, tracking of location, cloud based assistance and vision API's are salient characteristics of the system devised for the VI assistance. In another study [9], a cost-effective standalone system that is easy for blind people to use in real-time movement situations was devised. The work was part of incremental research of sensor based studies in the similar context. Here, a vision based wearable device and innovative stick was used for the movement assistance. The deep learning model i.e. YOLO model here, is used to identify the correct path for the movement. Ac- curacy of around 89.25% was found in the study. It has also systematically pointed out different algorithms that were used for the object detection. YOLO (you only look once) has been remained point of discussion around the corner since its inception. It is based on the concept of single shot detector that does localization and the classification of image in single step. Over a recent years, var- ious version of YOLO has been introduced, YOLOv11 being the recent one. Another interesting system for converting the digital by means of optical character recognition has been used in [10].The proposed work had four phases and was used for the data entry by means of digital conversion. The ultimate aim was to reduce the mistakes done by manual operation and save time. Similar systems can further be extended into system that can be useful for the blind individuals as well.

3. PROPOSED METHODOLOGY

In this proposed method, a efficient detector based application has been devised that can help visually impaired individual for the nav- igational assistance. The application also offers the audio feedback to the user hence prompting user to take quick decisions.

3.1 Model conversion

The framework used for the training of the model is Tensorflow. It has application programming interfaces for the development of various convolution networks. It offers the degree of easiness pro- viding direct calling of the layers in Python. It is open source and supported by large developer communities. The model is trained on COCO database, it has total 80 categories of objects that it can detect. To actually train the model COCO annotated dataset is fed to the model training pipeline, the tensorflow save model() api by default saves the model in .h5 model. It is frozen graph with stored weights inside it. Figure 4, shows the training and conversion pro- cess of the model used in the detection studies.

Tesnsorflow offers a API to take the inference on the .h5 model however android studio can not have direct inference of the .h5 model. The tflite interpreter facility inside the android studio can directly infer the .tflite model. The conversion process from h5 to tflite is straightforward and offers no any obstacle when all the layers in the network are supported by the current library version. There are multiple quantization and pruning options are provided with library. It is useful for reducing the size of the model in terms of memory consumed. It has the option of float16, int8 quantization that can reduce the memory footprint while maintaining the trade off for the accuracy.

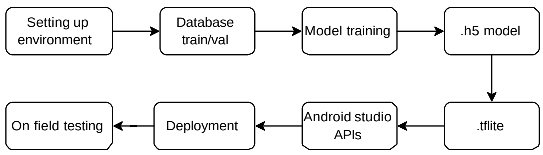


Fig. 4: block diagram Model conversion process to tflite and its deployment

4. RESULTS AND DISCUSSIONS

The current object detection model that was used in the application has been tested in real time scenarios. User was presented with the mobile application that can detect the object captured in the scene and can offer the auditory feedback with the name of the object that has been detected. Figure 3, shows the detection results.

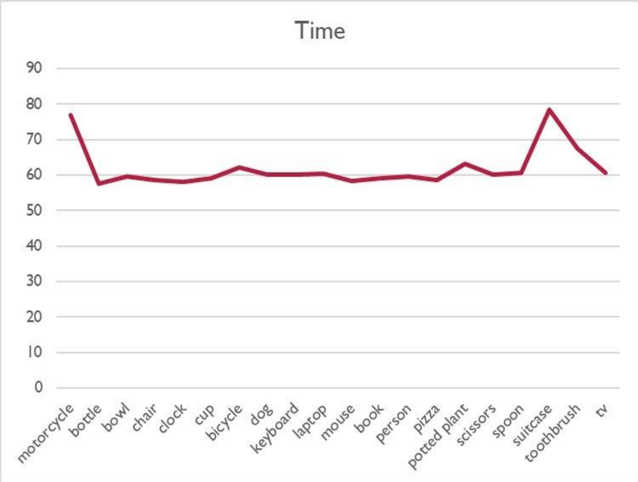


Fig. 6: Comparison of time for various classes

Table 1. : Accuracy and Time Data for Various Objects

Sr. no	Object name	Accuracy	Time (ms)
1	motorcycle	0.98	76.79
2	bottle	0.9	57.69
3	bowl	0.94	59.49
4	chair	0.95	58.51
5	clock	1	58.09
6	cup	0.97	59.01
7	bicycle	0.96	62.1
8	dog	0.8	60.05
9	keyboard	0.94	60
10	laptop	0.98	60.47
11	mouse	0.94	58.43
12	book	0.76	59.09
13	person	0.95	59.61
14	pizza	0.78	58.62
15	potted plant	0.83	63.19
16	scissors	0.96	60.04
17	spoon	0.82	60.6
18	suitcase	0.85	78.55
19	toothbrush	0.81	67.53
20	tv	1	60.59

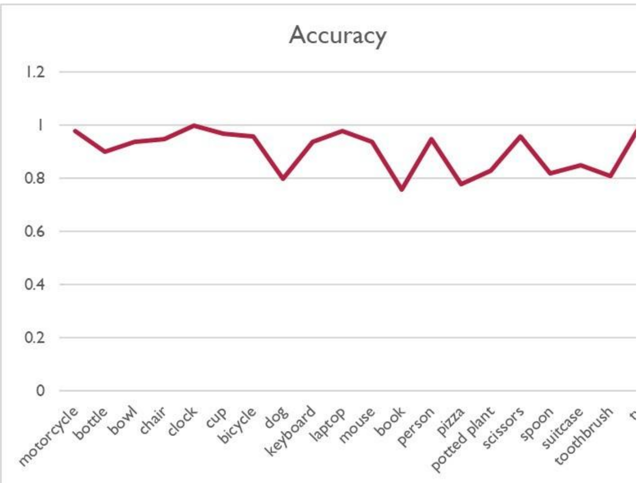


Fig. 5: Comparison of accuracy for various classes

Table 1, shows the accuracy of the model used in the application as well as time required by the model to present the result on the screen. This total time is the addition of the time taken by model for preprocessing, inference and postprocessing.

Avg accuracy: 0.90600
Avg time in ms: 61.9225
Figure 5 and figure 6, shows the comparison curve for accuracy and time, respectively.

5. CONCLUSIONS

In this work, study related to the blind assistance system has been reported. It significantly report which features of the vision systems are explored so that efficient blind assistance system for the visually impaired people can be designed. It has used object detection model based on mobilenet-ssd architecture, that follows the single shot localization and classification of the object. The average ac- curacy of the system crosses the 90%. This work also highlights the different object detection strategies that would help for making robust assistance systems.

6. ACKNOWLEDGMENT

Authors are thankful for the support received by Department of E&C Engineering, MIT, Chatrapati Sambhajinagar, its faculties, staff and all members.

7. REFERENCES

[1] N. Illakiya and V. Loganathan, "Comprehensive Assis- tive Mobility System for the Visually Impaired: The Smart Blind Stick with Ultrasonic Sensor Solution," in *2024 International Conference on Inventive Computa- tion Technologies (ICICT)*, pp. 2088-2094, 2024. doi: 10.1109/ICICT60155.2024.10544554.

[2] B. Pydala, T. Pavan Kumar, and K. Khaja Baseer, "Smart Eye: A Navigation and Obstacle Detection for Visu- ally Impaired People through Smart App," *Journal of Applied Engineering and Technological Science (JAETS)*, 2023.

[3] N. Syazreen Ahmad, N. Lai Boon, and P. Goh, "Multi-Sensor Obstacle Detection System Via Model-Based State- Feedback Control in Smart Cane Design for the Visually Challenged," *IEEE Access*, vol. 6, pp. 64182-64192, 2018. doi: 10.1109/ACCESS.2018.2878423.

[4] D. Denic´, P. Aleksov, and I. Vucˇkovic´, "Object Recogni- tion with Machine Learning for People with Visual Impair- ment," in *2021 15th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS)*, pp. 389-392, Oct. 2021. doi: 10.1109/TEL- SIKS52058.2021.9606436.

[5] N. Bhati, V. C. Samsani, R. Khareta, T. Vashisth, S. Sharma, and V. Sugumaran, "CAPture: A Vision Assis- tive Cap for People with Visual Impairment," in *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 692-697, 2021. doi: 10.1109/SPIN52536.2021.9565940.

[6] D. A. Khan, M. A. Zamir, M. S. Umar, and Z. Haider, "As- sistive Stick for Visually Impaired People," in *2022 5th Inter- national Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, pp. 1-6, 2022. doi: 10.1109/IMPACT55510.2022.10029287.

[7] Raihan Bin Islam, Samiha Akhter, Faria Iqbal, Md. Saif Ur Rahman, and Riasat Khan, *Deep learning based object de- tection and surrounding environment description for visually impaired people*, *Heliyon*, vol. 9,

no. 6, p. e16924, 2023.

[8] Myneni Madhu Bala, D. N. Vasundhara, Akkineni Haritha, and CH. V. K. N. S. N. Moorthy, *Design, development and performance analysis of cognitive assisting aid with multi sensor fused navigation for visually impaired people, Journal of Big Data,* vol. 10, no. 1, p. 21, 2023.

[9] Nitin Kumar and Anuj Jain, "A Deep Learning Based Model to Assist Blind People in Their Navigation," *J. Inf. Technol. Educ. Innov. Pract.*, vol. 21, pp. 95-114, 2022.

[10] J. E. M. Adriano, K. A. S. Calma, N. T. Lopez, J. A. Parado, L. W. Rabago, and J. M. Cabardo, "Digital conversion model for hand-filled forms using optical character recogni- tion (OCR)," *IOP Conference Series: Materials Science and Engineering*, vol. 482, no. 1, p. 012049, Feb. 2019.