iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XII **Month of publication:** December 2024

DOI: https://doi.org/10.22214/ijraset.2024.65823

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XII Dec 2024- Available at www.ijraset.com

A Review on Object Detection with VoiceSupport for the Visually Impaired People

G. Krishna Reddy[1], K. Aasha Mdhuri[2], L. Jhansi Aakanksha[3], P. Lahari[4], S. Mithila[5]
Electronics and Telematics, GNITS, Hyderabad, India

***Abstract:** This project is helpful for VAI (visually acquire information) to detect an object and be informed of text for a more independent accessibility to certain life activities. Using YOLOv8, an efficient object detection model, the system guarantees accurate object recognition; when more info is detected, the gTTS turns that information into clear, human-readable speech immediately.*

A Python-based application with OpenCV library used for image handling, the application run and stability has been tested and provento be efficient in Microsoft Visual Studio Code. Most importantly, it has been designed with flexibility in mind, can be implemented as an application or web-interface, to accommodate different users and their technology usage. It helps visually impaired users to get around spaces and understand text in similar manner to how a sighted person would interpret signs or labels format.

Through the application of modern AI technology the project provides enhanced independence and great improvement of the quality of life for the target population which is visually impaired; thus, the project is an example of progressive inclusion and use of assistive technology.

***Keywords:** YOLOv8 Object Detection, Text Recognition, Real-time Audio Feedback, Visually Impaired Assistance*

I. INTRODUCTION

With the demanding society we live in today, blind and partially sighted people are left with series of hurdles when it comes to mobility on their own. Even basic processes like perceiving an object, a sign, or recognizing to avoid any kinds of barriers may be challenging everyday tasks that demoralise and limit the individual. In managing with these problems, there has been enhanced preference for Computer vision and machine learning technologies in a bid to help the visually impaired identify objects and navigate accordingly. These tools help users interpret stimuli in their environment and respond appropriately and, by offering immediate audio feedback, help them function more safely and independently—the cornerstone of quality of life.

This project focuses on the creation of a wearable camera based object recognition system aimed at helping the visually impaired by detecting objects in their physical space and providing voice instructions.

This build-up is taken to integrate the enhancements achieved in processingof object detection by employing YOLOv8, which stands for You Only Look Once, the current version 8, recognized for its high Real-Time Object Identification performance. Through incorporation of this technology, the system can accurately identify objects due to classification, an aspect considered vital for accurate help. Such a level of precision is crucial to sustaining consumer confidence and use of the services a lot of the time in crowded or delicate settings.

The system is created in Python with the help of powerful libraries such as open-source computer vision OpenCV toanalyse the image and gTSS for voice output that helps to transfer vision into easily understandable audible form. OpenCV effectively manages camera data feed, recognizing the images as well as objects and extracting information in real-time. Then, the data is processed to be given speech using gTTS and the feedback given is comprehensible and natural making the system more user friendly. The integration within the Visual Studio Code environment also provides dependability and convenient adjustments for maintenance of the work while providing the developers a consistent way to enhance"./

This project also has a number of advantages that include flexibility of how it can be implemented. Intended to existas a mobile app as well as a web application, the program is portable and versatile for users with different levels of technological engagement. Apart from its utility, this project highlights a values statement by providing visuallyimpaired individuals with a tool that allows them to better move through the indoor and outdoor spaces independently. Combing the state of the art AI technologies, the concept of accessible design, and the technique of adaptive deployment, this project is mainly ahuge advancement toward independence and improvedquality of life for visually impaired population of theworld.

©IJRASET: All Rights are Reserved

II. LITERATURE REVIEW

This work outlines a portable camera-based object recognition system relevant for blind people done by Vikky Mohane and Prof. Chetan Gode. The system employs a small camera installed at a wearable device in capturing video streams to process real-time images forobject recognition. The system uses MobileNetV2 architecture incorporating real-time efficiencies as its primary concern. The processing device is usually an embedded system such as a Raspberry Pi 4 with 4 GB RAM which deploys the model on an ARM Cortex-A72 processor. The restrictions to this work are: The work uses a MobileNetV2 model, designed for low-power devices and deployed on Raspberry Pi 4, with 4 GB RAM and ARM Cortex-A72 CPU @1.5GHz. The system employs the COCO dataset that is optimized forindoor navigation and scores 80% average precision across the tested objects. Here the response time is 300ms per frame, this makes the program good for real time navigation of the robot. In user testing for indoor testing, an average 85% detection rate of some objects like chairs, tables, doors, among others was observed, although the system's accuracy reduced as the lightconditions changed and in outdoor settings. Analysis ofthe presented results show that this system has a low response time which suggests the possibility of indoor practical applicability; moreover, future works will focus on decreasing the latency and on further extensionof the model for difficult indoor and outdoor conditions.

The 'Project: CNN-Based Object Recognition andTracking System to Assist Visually Impaired People' byFahad Ashiq, Muhammad Asif, Maaz Bin Ahmad et al. then introduces a CNN-based object recognition andtracking system which aims for real-time object awareness regions for visually impaired users, more specifically stressing over low-latency and sustained tracking mechanisms. The system employs YOLOv3 model tweaked for ARM and tracks multiple objects at once with real-time updates for the moving objects. Technical details refer to YOLOv3 with the pruning to favour the embedded optimization, executed on a Jetson by NVIDIA with 128 Cuda cores and 4Gb RAM. The used dataset is COCO; moreover, the dataset has been augmented for static and moving objects The detection accuracy is 92 % with the average IOU 0.75. The system's speed is roughly 100 ms per frame on the Jetson Nano, thus enabling tracking. The recognition accuracy was high and the object was smoothly tracked in differentsituations; and it was tested with still and moving objectsreal-time audio feedback was given. YOLOv3 is employed to minimize its ability to detect both small andlarge objects, but at times instances of object occlusion such as those cases where items can be placed on top of one another led to faults in tracking. The future work contains extending the work by incorporating the stereo vision to get a better depth sense.

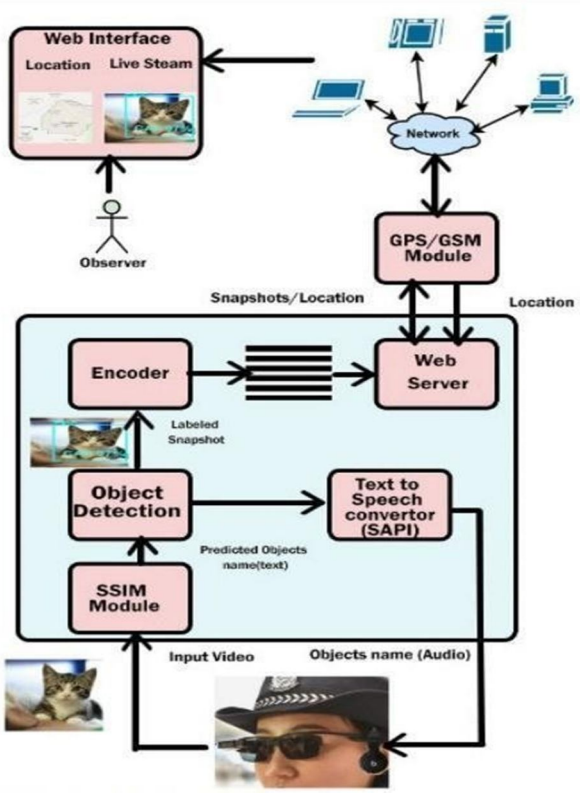


Fig.1. Architecture of the System

The study "Object Detection and Recognition for Visually Impaired Users: On the title “A Transfer Learning Approach” Wei-Cheng Won, Yoke-Leng Yong, and Kok-Chin Khor proposed the use of thetransfer learning in improving their object recognition system training time, computation time while still achieving high accuracy. The system uses transfer learning from the ImageNet dataset with the ResNet50 model; it is fine-tuned for domestics objects and runs ona low-power Intel Movidius Neural Compute Stick 2. The technical specification of this implementation comprises the ResNet50 model, which can be pretrainedvia the ImageNet dataset and fine-tuned for transfer learning, operating on Intel Movidius Neural Compute Stick 2 (Myriad X VPU). ImageNet and the Open Images Dataset V6 are involved as well as the objects inside a home include common household items, with the well-trained model obtaining a top five accuracy of ninety percent on such items. In terms of response time it takes about 150 ms to make one inference. The system was effective in the recognition of items found in a home including tables, chairs and kitchen utensils among others while still ensuring that the computations required were optimized enough to allow the project to be affordably deployed in mobile-platforms. The transfer learning approach shortened the training greatly and also augmented the recognition accuracy, but the variation of light influenced the result negatively. The future is to expand in terms of improvement in the performance of the method through incorporation of light control mechanisms during operation in real time environments.

In the project “Real-Time Object Detection for Visually Challenged People” by Sunit Vaidya, Niti Shah,Naisha Shah & Prof Radha Shankarmani, they have independently built a real time object detection system for the visually impaired with clients telling them which objects are being detected through voice feedback. The system incorporates YOLOv4 which we identified as a fast and accurate object detection algorithm to offer real-time feedback in the identified tasks based on voiceguidance. The Coquette model operates on a mobileprocessor for wearables. Technical requirements consist of: • YOLOv4 model which is suitable for real-time processing; • Hardware: Qualcomm® Snapdragon™ 865Mobile Platform; RAM: 6 GB. A custom-labelled datasetwhich contains the objects of a typical urban environment, including the interior of houses, was used and combined with the realistic lighting conditions and the method provided the mAP of 88% of the primary categories, such as vehicles, people, and obstacles. The current system response time is at around 50 ms per frame, or instant reaction time. site-specific performanceevaluations showed low rates of collisions and high reliability of obstacle and regular objects recognition, increased users’ self-esteem and mobility both indoors and outdoors when comparing the results of the project with the results of their previous experiences. Based on the project presented, spatial detection with YOLOv4 guaranteed fast, high-speed operation providing real time voice feedback. Richness of audio instruction often encountered unwanted noise in urban settings which constitutes areas for improvement in the future which include use of noise eliminating feature for better outdoorperformance.

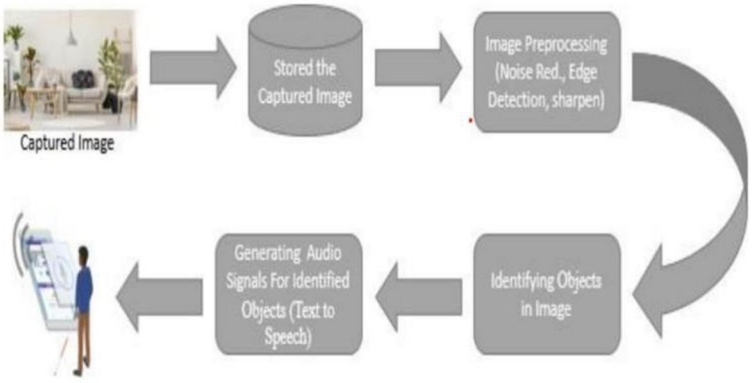


Fig.2 System Design

An object detection system for the visually impaired is the topic of the paper prepared by MissRajeshvaree Ravindra Karmarkar and Prof. V.N. Honmane, which proposed the real-time object recognition system that is combined with voice guidance. The system employs a Fast R-CNN model for detecting objects properly in the infrastructure, a Raspberry Pi 4 (4GB RAM) is employed, and the camera employed records images at a resolution of 1280 * 720 pixels. Theyare able to identify both indoor and outdoor items, for example household items, cars and pedestrians then offeraudio signals to enable the mobility.The model also considers a Raspberry Pi and was trained on the custom dataset extracted from Open Images Dataset V6. The accuracy of the system was at 85% for object detection and the average intersection over union of the system was 0.70. Negotiation of each frame takes approximately 350 ms, somewhat slower than for other approaches but still manageable. The user testing demonstrated that for visually impaired users, the system increased navigation ease; however, outdoor conditions sometimes affected the sound quality. In the upcoming work, it is recommended to integrate the ANC technology into the audio improvement system.

III. CONCLUSION

The introduction of enhanced real time object detection system with voice help service is an important invention in the assistive technology particularly for the visually impaired persons. This project aims to provide a low-latency and efficient method of achieving safe and independent mobility using YOLOv5 for in real-time and fast object detection. The incorporation of this technology gives users instantaneous sounds the details of theirenvironment they need to act proactively.

For the efficiency of the project, the activities will be subject to improvements in the selected areas in the course of their implementation. First, the audio feedback system must be improved so it can operate in necessary noisy urban environments. By doing so the system is able to remain intelligible and avoid the distortion of information that needs tobe passed to the user at a given instance. Furthermore, enhancing the object recognition accuracy in dynamic environmental settings, including; brightness, weather, and varying complexities of the backgrounds, will enhance its reliability in practical applications. Last but not least, integration with depth-sensing could provide more refined spatial interactivity that lets users get not only information about the objects but information about distance and positioning of them in relation to the user.

As future versions of these enhancements are integrated, they will contribute to an increasingly intelligent, adaptable, and user-friendly assistant, always enhancing the safety, liberty andwell-being of those with vision impairment.

IV. FUTURE SCOPE

This system can be greatly expanded in the future for the purpose of increasing assessability and usability for those who are visually impaired. First, it is necessary to refine audio feedback to make it possible dynamically changing the frequency for various noisy environments; it is possible to use noise canceling techniques and adaptive frequency modulations. In order to enhance object recognition in various conditions, including low light, or a lot of objects in a frame, using more prescriptive machine learning algorithms and applying transfer learning with datasets best suited to the content of the context will be effective. There are two improvements that might help deepen spatial understanding: LiDAR and stereo vision, both of which would help users understand distances and avoid hits and bumps more effectively.

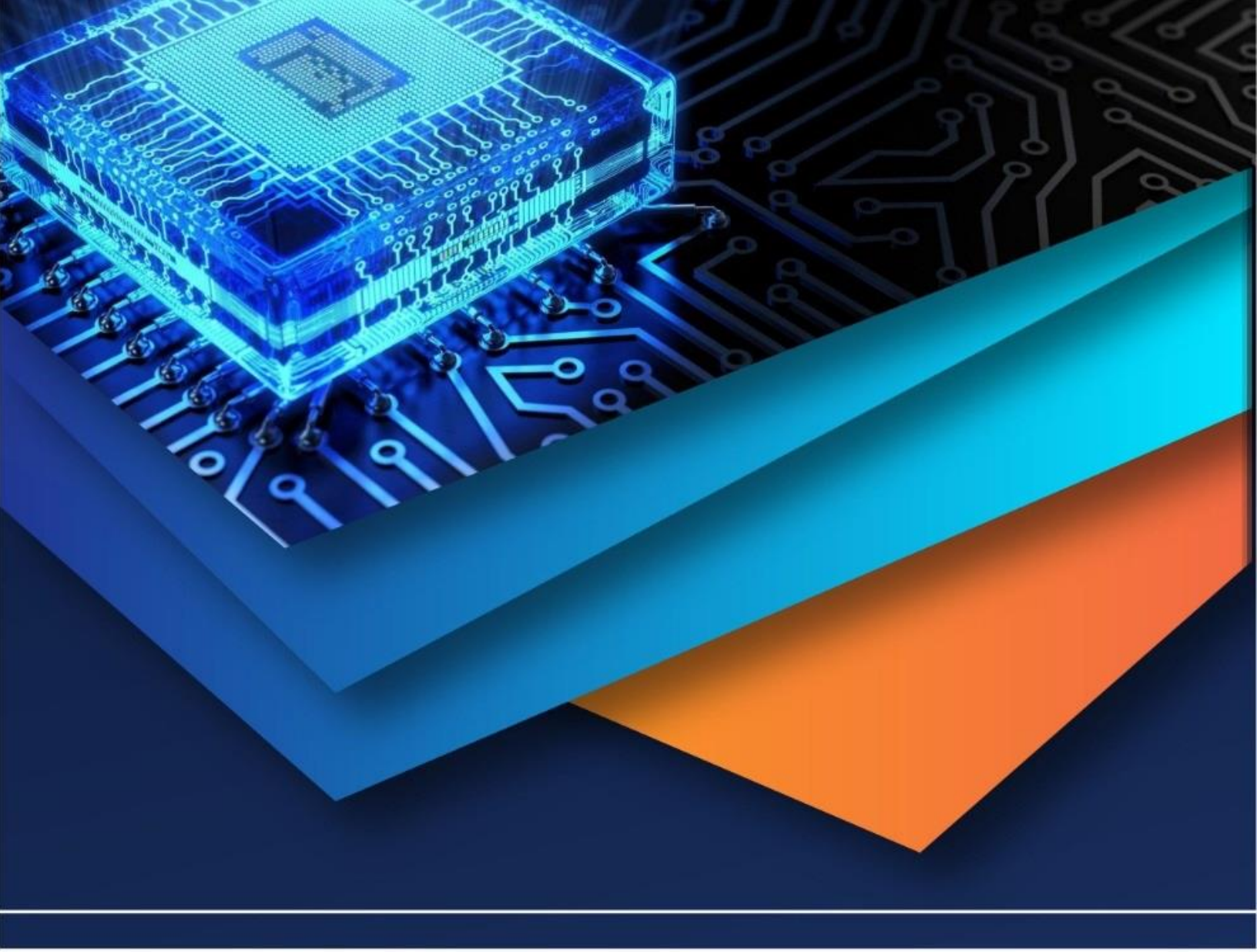
Wearable and mobile optimization we believe is another promising direction, as making the system itself portable in a form of small wearables or mobile applications would further extend its convenience factor. Switching multiple languages might extend the system's coverage, which would be beneficial to the customers of various language origin. Last, but not the least, after real-world testing across different settings and with users and collection of feedback subsequent improvements will be able to incorporate into the system so that it remains effective, user-friendly and safe for the visually impaired in their everyday lives.

REFERENCES

[1] W. Elmannai and K. Elleithy, ''Sensor-based assistive devices for visuallyimpaired people: Current status, challenges, and future directions,'' Sensors, vol. 17, no. 3, p.565, 2017.
[2] L. B. Neto, F. Grijalva, V. R. M. L. Maike, L. C. Martini, D. Florencio, M. C. C. Baranauskas, A. Rocha, and S. Goldenstein, ''A Kinect-based wearable face recognition system to aid visually impaired users,'' IEEE Trans. Human-Mach. Syst., vol. 47, no. 1, pp. 52–64, Feb. 2017.
[3] K. Kumar, B. Champaty, K. Uvanesh, R. Chachan, K. Pal, and A. Anis, ''Development of an ultrasonic cane as a navigation aid for the blind people,'' in Proc. Int. Conf. Control, Instrum., Commun. Comput. Technol. (ICCICCT), Jul. 2014, pp. 475–479.
[4] Choi D., and Kim M. ' , "Trends on Object Detection Techniques Based on Deep Learning,." Electronics and Telecommunications Trends, Vol. 33, No. 4, pp. 23- 32, Aug. 2018.
[5] J. Dai et al.' , "R-FCN: Object Detection via Regionbased Fully Con- volutional Networks." Conf. Neural Inform. Process. Syst., Barcelona, Spain,Dec. 4-6, 2016, pp. 379-387
[6] O. Russakovsky et al., ' , "ImageNet Large Scale Visual Recognition Challenge,." Int.J. Comput. Vision,Vision, vol. 115, no. 3, Dec. 2015, pp. 211- 252.
[7] World Health Organization, "Blindness and vision impairment," World Health Organization, 2019. [Online]. Available: https://www.who.int/news-room/fact- sheets/detail/blindness-andvisual-impairment. [Accessed Apr. 16, 2021]
[8] A. D. P. dos Santos, F. O. Medola, M.J. Cinelli, A. R. Garcia Ramirez, and F. E. Sandnes, "Are electronic white canes better than traditional canes? A comparative study with blind and blindfolded participants," Universal Access in the Information Society, pp. 93 – 103, March 2020
[9] IrisVision Global, "Top 5 electronic glasses for the blind and visually impaired," irisvision.com, 2021. [Online]. Available at: https://irisvision.com/electronic-glasses-for-the- blind-and-visuallyimpaired/. [Accessed Apr. 16, 2021].
[10] HanenJabnoun, FaouziBenzarti,and Hamid Amiri "Object recognition for blind people based on features extraction"In INTERNATIONAL IMAGE PROCESSING APPLICATIONS AND SYSTEMS CONFERENCE 2014
[11] B. Durette, N. Louveton, D. Alleysson, and J. H´erault. "Visuo-auditorysensory substitution for mobility assistance: testing The VIBE". InWorkshop on Computer Vision Applications for the Visually Impaired,Marseille, France, 2008.
[12] A. F. R. Hern´andez et al. "Computer Solutions on Sensory Substitution for Sensory Disabled People". In Proceedings of the 8th WSEAS International Conference on Computational Intelligence, Man-machine Systems and Cybernetics, pages 134–138, 2009.
[13] Jamal S. Zraqou, Wissam M. AlKhadour and Mohammad
[14] Z. Siam Real-Time Objects Recognition Approach for Assisting Blind People 2017 International Journal of Current Engineering and Technology

[15] Aniqua Nusrat Zereen, Sonia Corraya Detecting real time object along with the moving direction for visually impaired people 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)
[16] Rui Li , Jun Yang Improved YOLOv2 Object Detection Model 2018 6th International Conference on Multimedia Computing and Systems (ICMCS)

.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)