International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

# Real-Time Object Detection and Tracking System for Visually Impaired People Using aDeep Learning Model

Mr.S.T.Saravanan
Computer Science and Engineering
*(JNTUH)*
Sphoorthy Engineering College
*(JNTUH)*
*Hyderabad*, India
saravanan@sphoorthyengg.ac.in

Dr.Subba rao Kolavennu
Computer Science and Engineering
*(JNTUH)*
Sphoorthy Engineering College
*(JNTUH)*
*Hyderabad*, India
profrao99@gmail.com

Sushmitha Nelakanti
Computer Science and Engineering
*(JNTUH)*
Sphoorthy Engineering College
*(JNTUH)*
*Hyderabad*, India
nelakantisushmi2001@gmail.com

Varun Sai Gunda
Computer Science and Engineering
*(JNTUH)*
Sphoorthy Engineering College
*(JNTUH)*
*Hyderabad*, India
varonnsai@gmail.com

Manikanta Jakkidi
Computer Science and Engineering
*(JNTUH)*
Sphoorthy Engineering College
*(JNTUH)*
*Hyderabad*, India
manikantareddy8092@gmail.com

***Abstract*—In today's world, technology is increasing drastically, which in turn makes people work efficiently and solve issues effortlessly. But there are some technologies that are way behind. For instance, there are large numbers of visually impaired people (VIPs) globally who are facing difficulties doing their day-to-day work.**

**In this work, a "deep learning" model is developed for object recognition. It is a smart and intelligent system that can be deployed into a device that can be used to assist mobility and ensure safety for VIPs. As VIPs cannot visually see real-time objects with their naked eyes, this work helps in sensing and visualizing the objects in real-time with the system's automated voice. In this work, we also developed a web application to track the device on which the deep learning model is installed.**

**Keywords—Deep Learning Model, Convolutional neural network, Web Application, Systems automated voice, Object Recognition.**

## I. INTRODUCTION

The World Health Organization (WHO) has reported that 285 million of the world's population are eyeless or visually bloodied. Out of these, 39 million are blind [1]. The major conditions that beget visual impairments include refractive error, glaucoma, trachoma, corneal darkness, cataracts, diabetic retinopathy, and unaddressed diplopia[2].

Visually disabled persons(VIPs) face difficulties in performing the conditioning of diurnal living, e.g., openings for work and training, moving in their surroundings, having the capability to interact with the terrain, textual information and translation, and searching for common objects(indoor and outdoor) on their own or indeed with some backing[3].

The main challenges for VIPs are object discovery and recognition, currency identification, textual information (signs and symbols) and restatement, mobility and navigation, and safety. In history, several approaches, systems, biases, and operations have been developed in the sphere of assistive technology to grease VIPs into performing tasks that they were formerly unfit to negotiate[4]. Similar results generally comprise electronic bias equipped with cameras, detectors, and microprocessors capable of giving audible feedback to the user. Numerous other object discovery and recognition systems claim high accuracy but cannot give the necessary information and attributes for the shadowing of VIPs to ensure their safe mobility[5].

Though eyeless people cannot see objects in their surroundings, it would be helpful to know about them. Likewise, there's a need to develop a shadowing system through which family members of VIPs can cover their movements.

Based on the important requirements, this paper proposes a smart system that can perform real-time object localization and recognition. As soon as the system recognizes the object, it sends audible feedback to the user. For illustration, after relating a given object (e.g., an auto), the user will hear the word" auto." Also, the user's position and a shot of the most recent viewed scene are periodically stored on a server that can be accessed by family members using an application to track the user. For object discovery and recognition, the YOLO architecture is used because of its high speed rate and good accuracy.

## II. LITERATURE SURVEY

The experimenters have proposed several ways to develop assistive bias for VIPs. The following technologies have been used: vision-based detectors (camera), non-vision-based detectors (e.g., IR, ultrasonic, inertial and glamorous seeing, etc.), and other technologies similar to low-energy Bluetooth lights, GPS, GPRS, etc.

In this section, the main focus is put on vision-grounded seeing biases due to their applicability to our proposed system. Assistive bias can be grouped into the following three classes according to their functionalities:

- Object discovery bias
- Navigation bias
- Hybrid bias (object discovery and navigation)

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

- The bias can be sub-classified based on their working principle and by means( sound or climate) of conveying information to VIPs.

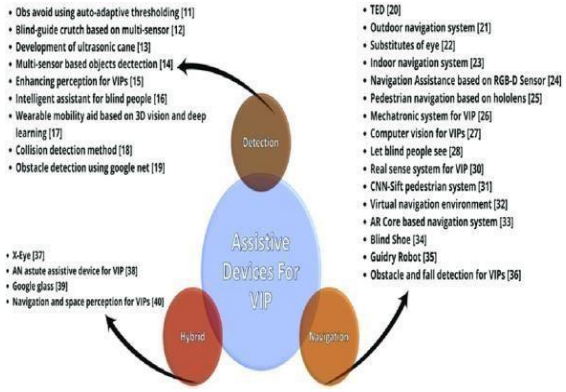


Figure 1: Different assistive devices

Berge et al [6] have introduced a device that helps VIPs and other impaired people through an Android-based system. It's helpful not only for detecting obstacles but also for providing navigation. The system was tested on 15 users. 10 of these were passing disabilities related to visual impairment, and 5 of them were eyeless.

Kammoun et al [7] have constructed a system called NAVIG to aid VIPs. It contains microphones, cameras, headphones, GPS, detectors, and a computer. The device uses a stereoscopic camera located on the top side to capture images of the terrain. Machine literacy methods are used for object discovery.

Chang et al [8] introduced a wearable system that assistsVIPs using artificial intelligence (AI) methods. The system attendants VIPs at zebra crossings. The system achieved accuracy up to 90. Chang et al [9] developed an intelligent obstacle discovery system for VIPs. The system consists of a walking stick, smart spectacles, a cloud-dependent platform, and a mobile application.

Rahman and Sadi [10], who developed a system for VIPs that helped them identify out-of-door and inner objects through audio dispatches, the system is grounded on four detectors that help detect the object from all directions. The overall accuracy of the system is roughly 99.

In former times, multitudinous studies have been performed to address the challenges. But still, there's a need to design and develop intelligent systems that help in object discovery and recognition and secure the mobility of VIPs.

## III. Proposed System

The proposed system comprises a Raspberry Pi DSP (Digital Signal Processing) board with a global positing system (GPS) module and GSM, headphones, and a camera. DSP captures a live feed from the videotape camera and passes it to the object discovery and recognition module. The model predicts objects in the current videotape frame and passes their names to the text-to-speech module (gTTS module), pronouncing their names using headphones. Also, the labelled shot is encoded using the Joint Photographic Experts Group ( JPEG) encoder. The encoded image with the user's precise position is saved in the database. Moreover, a user-driven feature is handed to VIPs that enables family members to track their movement (get the current labelled frame and position) while relaxing at home through a web interface. The user can enable/disable the feature with a single button press.

### A. *Deep Learning Approaches*

Deep learning approaches for object discovery provides a fast and accurate means to predict the position of an object. Deep learning is an important machine learning technique in which the object sensor automatically learns image features needed for discovery tasks. Several ways for object discovery using deep learning are available, such as Faster R-CNN, You Only Look Once (YOLO) v2, YOLO v3, YOLO v4, YOLOv5, and Single Shot Detection (SSD).

### B. *Object Detection and Recognition*

The system employs YOLOv5 architecture for Object detection and recognition.YOLOv5 is powerful, easy to use, and capable of achieving state-of-the-art results for object discovery tasks. It's also more accurate and easier to train than its forerunners, making it a popular choice for numerous inventors.

### C. *YOLOv5 Architecture*

YOLO is a state-of-the-art, real-time object discovery algorithm created in 2015 and pre-trained on the COCO dataset. It uses a single neural network to reuse an entire image. The image is divided into regions, and the algorithm predicts chances and bounding boxes for each region. YOLO is known for its speed and accuracy. Since 2015, the Ultralytics unit has been working on perfecting this model, and numerous performances have also been released.

### D. *Web Application as an Interface*

The motive behind the development of the web interface is to ensure the safety of VIPs. This interface facilitates the family members of the VIPs to track their movements (location and snap shorts) while sitting at home. The device allows VIPs to visualize the environment and ensure their safety.

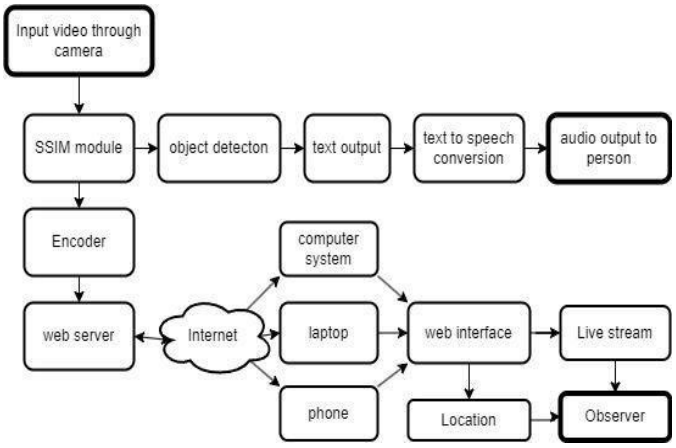


Figure 2: The architecture of the proposed system

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

## IV. Technical Evaluation

This section describes the evaluations of a prototype device designed to assist visually impaired individuals by detecting nearby objects and providing information about them through voice.

To train the object recognition model, a dataset with descriptions of different objects was used to extract their features. The COCO dataset, which includes 80 objects, was utilized to develop the YOLOv5 model, a novel convolutional neural network that can detect objects in real-time with high accuracy and speed.

YOLO can recognize objects and process frames at up to 230 FPS for small networks, making it one of the best models in object recognition. Additionally, a tracking system was developed using Flask, a popular web framework for building web applications. The device provides object representation and type information through voice, and periodically sends the user's position to a web server. However, user testing and feedback are necessary to evaluate the device's effectiveness and ensure it meets the needs and preferences of visually impaired individuals.

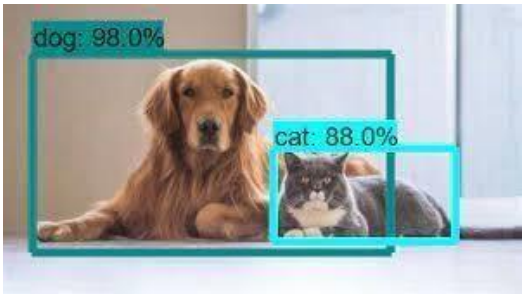


Figure 3: Object Recognition

## V. Conclusion

This paper presented a smart and intelligent system for VIPs to help them with mobility and ensure their safety. The proposed system is grounded in the day-to-day conditions of VIPs. It assists them in imaging the terrain and furnishing a sense of the surroundings. They can sense objects around them and smell the natural terrain using CNN-based YOLOv5 architecture. Furthermore, a web-based application is being developed to ensure the safety of VIPs.

The user can turn on the on-demand function to share their location with the family. It's beneficial to their family as they can cover the movements of VIPs and track their position using the live feed from the camera.

Analysis shows that this work provided satisfactory results. Further, after launching the system, surveys and user feedback will also be gathered to enhance its performance.

## References

[1] C W. Elmannai and K. Elleithy, ‘‘Sensor-based assistive devices for visually impaired people: Current status, challenges, and future directions,’’ Sensors, vol. 17, no. 3, p. 565, 2017.

[2] Work Sheet. Accessed: Mar. 7, 2020. [Online]. Available: https://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impairment

[3] R. Velázquez, ‘‘Wearable assistive devices for the blind,’’ in *Wearable and Autonomous Biomedical Devices and Systems for Smart Environment.*Berlin, Germany: Springer, 2010.

[4] L.-B. Chen, J.-P. Su, M.-C. Chen, W.-J. Chang, C.-H. Yang, and C.-Y. Sie, ‘‘An implementation of an intelligent assistance system for visually impaired/blind people,’’ in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2019, pp. 1–2

[5] C. Shah, M. Bouzit, M. Youssef, and L. Vasquez, ‘‘Evaluation of RU-netra□tactile feedback navigation system for the visually impaired,’’ in Proc. Int. Workshop Virtual Rehabil., 2006, pp. 72–77.

[6] A. Berger, A. Vokalova, F. Maly, and P. Poulova, ‘‘Google glass used as assistive technology its utilization for blind and visually impaired people,’’ in Proc. Int. Conf. Mobile Web Inf. Syst., 2017, pp. 70–82.

[7] S. Kammoun, G. Parseihian, O. Gutierrez, A. Brilhault, A. Serpa, M. Raynal, B. Oriola, M. M. Macé, M. Auvray, M. Denis, and S. J. Thorpe, ‘‘Navigation and space perception assistance for the visually impaired: The NAVIG project,’’ IRBM, vol. 33, no. 2, pp. 182–189, Apr. 2012.

[8] W.-J. Chang, L.-B. Chen, C.-Y. Sie, and C.-H. Yang, ‘‘An artificial intelli gence edge computing-based assistive system for visually impaired pedes□trian safety at zebra crossings,’’ IEEE Trans. Consum. Electron., vol. 67, no. 1, pp. 3–11, Feb. 2021.

[9] W.-J. Chang, L.-B. Chen, M.-C. Chen, J.-P. Su, C.-Y. Sie, and C.-H. Yang, ‘‘Design and implementation of an intelligent assistive system for visually impaired people for aerial obstacle avoidance and fall detection,’’ IEEE Sensors J., vol. 20, no. 17, pp. 10199–10210, Sep. 2020.

[10] M. A. Rahman and M. S. Sadi, ‘‘IoT enabled automated object recognition for the visually impaired,’’ *Comput. Methods Programs Biomed. Update*, vol. 1, May 2021, Art. no. 100015.