Vietnam Journal of Computer Science

Trinetra v1.0: An Aid to Visually Impaired Persons

Tejaswini Borekar*, Shravani Gavli, Aditi Khambe

UG Student, Department of Information Technology,
Vidyavardhini's College of Engineering & Technology, Vasai
University of Mumbai, India

Ashish Vanmali

Associate Professor, Department of Information Technology,
Vidyavardhini's College of Engineering & Technology, Vasai
University of Mumbai, India

Received (Day Month Year)
Revised (Day Month Year)

In this paper, we introduce a revolutionary solution named Trinetra designed to navigate the challenges faced by visually impaired persons in their daily lives. Trinetra represents a significant leap forward in providing a seamless and accessible one-touch solution for the vision loss community, aimed at fostering independence and self-reliance. This innovative system combines strengths of Internet of Things (IoT) and Artificial Intelligence (AI) to identify the surrounding objects and their distances from the person and translates the same as an audio output. The core functionality of Trinetra involves capturing real-time images through an integrated camera, coupled with the implementation of the cutting-edge YOLO (You Only Look Once) algorithm for efficient object detection. An onboard depth sensor measures the distance of the object. This information is then translated into audio output using text-to-speech conversion, providing visually impaired persons with valuable insights about their surroundings.

Keywords: Artificial Intelligence; Depth Sensor; Object Detection; Pi-camera; Raspberry Pi; YOLO.

1. Introduction

As per World Report of Vision of World Health Organization (WHO), approximately 40 million people in the world are blind. Additionally, another 250 million people suffer from different forms of visual impairments [1]. In India, these numbers are esteemed to be around 5 million and 70 million respectively. These people face lot of challenges in their day-to-day life such as reading, cooking, and navigating unfamiliar environments. This also takes toll on their mental health leading them to frustration, anxiety, and depression.

*Department of Information Technology, Vidyavardhini's College of Engineering & Technology, Vasai, University of Mumbai, India. E-mail: teju2612006@gmail.com

Visually impaired individuals encounter significant challenges in their daily lives, hindering their independence and safety while navigating their surroundings. Traditional assistance methods are inadequate in delivering timely, dependable information, limiting their capacity to interact confidently with their environment. With the advancements in embedded systems and machine learning, many solutions are proposed in recent times. The readers can refer work of [2], [3], [4], [5] for more details.

The Trinetra project is like a guiding light for visually impaired individuals, designed to make their lives easier and more independent. The visually impaired persons can move around independently, reducing your reliance on others for assistance. It?s like gaining a newfound freedom, as you can explore your surroundings with greater ease and safety.

2. Previous Work

There have been many attempts in the literature to provide aid to visually impaired persons. The methods vary based on type of camera used, either monocular or stereo vision, depth sensing mechanism, the type of processing system used, type of object detection algorithm and so on. This section provides few glimpses of these methods.

The mechanism will greatly depend on the camera system used. The cameras can be monocular, stereo vision or can be combination capturing visible-infrared spectrum. The depth sensing mechanism will depend on the type camera used. Few variations of these work are presented below:

- **Marzullo et al., "Vision-based Assistive Navigation Algorithm for Blind and Visually Impaired People Using Monocular Camera" [6]:** Marzullo et al. employed a three-stage image processing method using a monocular camera. They used first order Hough transform to detect regions like walls, doors, windows etc. A feedback using physical medium is provided to the user for necessary corrections of trajectory.
- **Liu and Aggrawal, "Local and Global Stereo Methods" [7]:** Liu and Aggrawal has presented a detailed overview of concept of stereo vision. They reviewed different local and global algorithms used for depth estimation. Local methods are faster compared to global method which are based on optimization. However, global methods give more accurate disparity map as opposed to local methods.
- **Adi et al., "Blind People Guidance System using Stereo Camera" [8]:** Adi et al. introduces a supportive solution that employed stereo cameras to help those who are impaired vision identify items and barriers in their immediate surroundings. It makes use of ZED stereo camera and a computer to precisely compute the difference between nearby obstacles and conveys this information to the user through stereo sounds. The system is claimed to have an accuracy of 83.16
- **Zhong et al., "Real-time depth map estimation from infrared stereo images of RGB-D cameras" [9]:** In this paper authors used RGB-

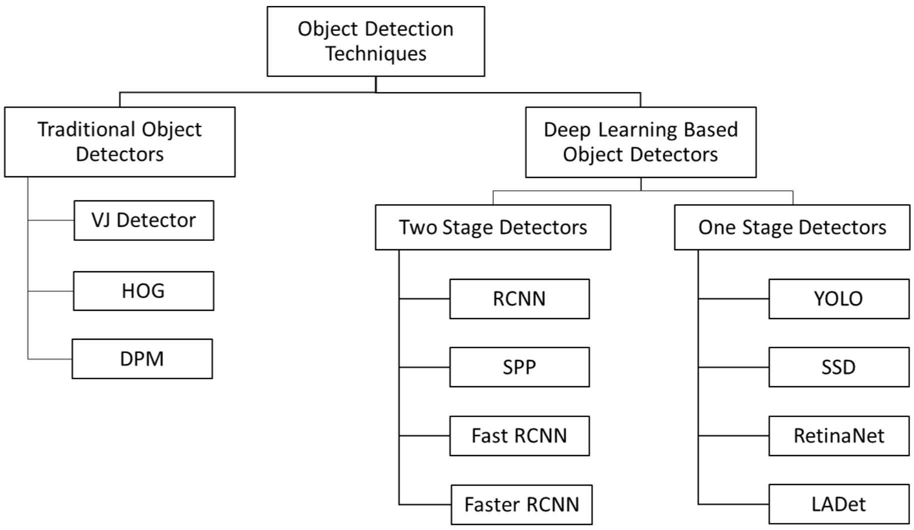

Fig. 1. Object Detection Methods

D cameras for 3D depth perception instead of using regular stereo cameras. They presented a robust and effective coordination formula based on semi-global synchronization principles to generate real-time accurate and comprehensive depth maps that considers the specific characteristics of infrared speckle image. It also used the idea of block matching to improve the system performance.

- **Vanmali et al., "A Novel Approach For Image Dehazing Combining Visible-NIR Images"** [10]**:** Vanmali et al. used a combination of visible and near infrared (NIR) images for depth map estimation followed by dehazing of images.

Another crucial part of the project is object detection. There is a variety of methods available for object detection. A quick glance of these methods is shown in Fig. 1 (Reproduced from [11]). Few commonly used methods are explained below:

- **Redmon et al. "You Only Look Once: Unified, Real-Time Object Detection."**[9] YOLO offers a unique perspective on object tracking by diverging from conventional methods and avoiding the need to adapt prior approaches. Instead of separate steps, YOLO treats object tracking as a regression problem while simultaneously creating classifiers for identification tasks. This approach involves making predictions of bounding boxes with associated class probabilities within a single neural network. YOLO excels at accurately estimating both the bounding boxes and class probabilities for entire images in a single pass. The advantage of this approach lies in its

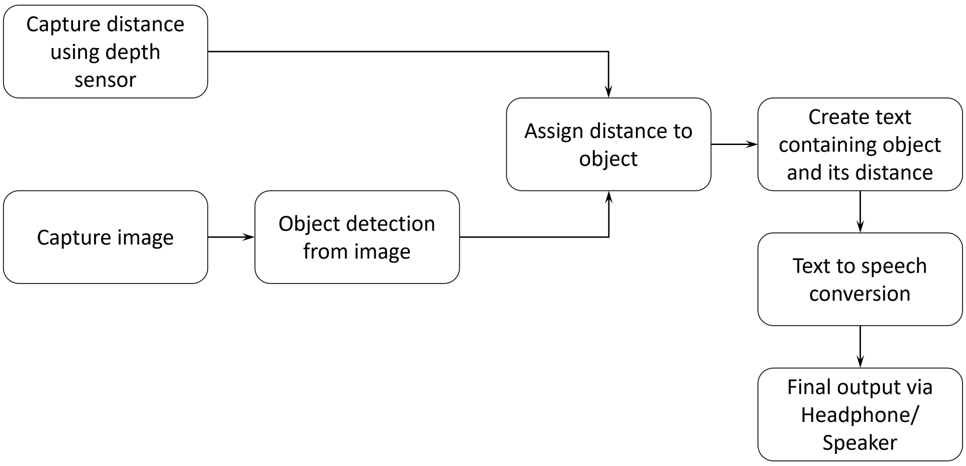

Fig. 2. Block Diagram of Proposed Solution

unified network, which streamlines the entire detection process.

- **Atitallah et al., "An effective obstacle detection system using deep learning advantages to aid blind and visually impaired navigation" [12]:** Atitallah et al. proposed an obstacle detection system based on a modified YOLO v5 neural network architecture. They tested the system on IODR datasets and MS COCO datasets for indoor and outdoor operations.

3. Proposed Solution

The proposed solution combines the strengths of IoT and AI to give a real time solution that will aid the visually impaired persons. The block diagram of proposed solution is given in Fig. 2. The corresponding interfacing of IoT system is given in Fig 3.

The key hardware components of our IoT system are:

- **Ultrasonic Sensor:** Various types of distance sensors are available for IoT systems which include ultrasonic, IR proximity, laser distance, etc. The ultrasonic sensor emits high-frequency sound waves towards the target object, and records the round trip time from the target for the reflected waves to calculate the distance. We have used ultrasonic sensor HC-SR04 for our system which has a range of 2cm-450cm.
- **Raspberry Pi:** Raspberry Pi is a series of small single-board computers (SBCs) developed by the Raspberry Pi Foundation and forms the heart of our IoT system. Raspberry Pi boards are small, compact, easy to install and maintain. They are highly energy-efficient and provide support for various types of sensor and actuators. They also have interface camera which is the

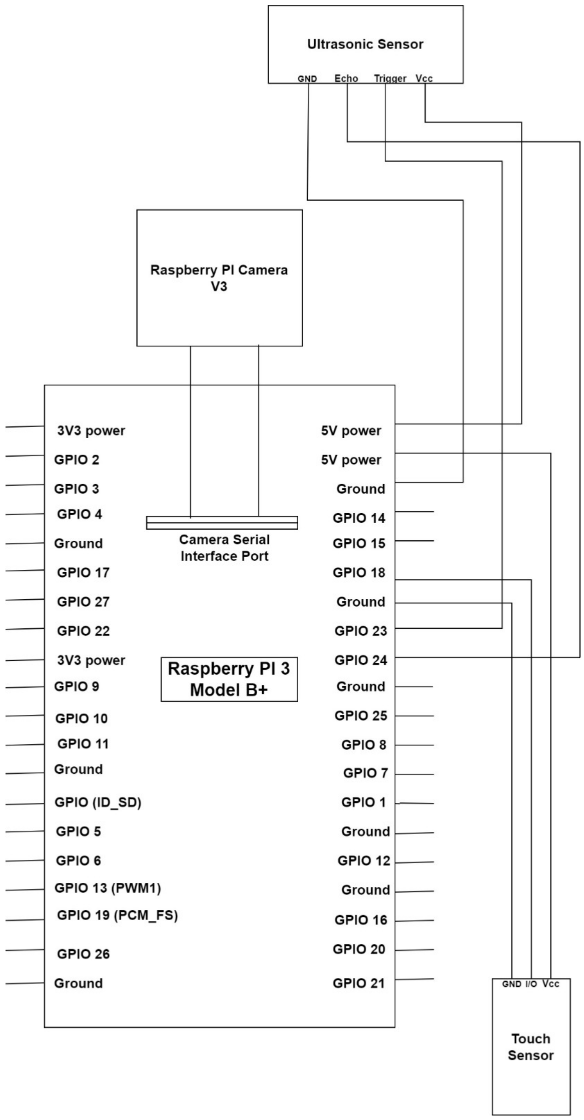

Fig. 3. Interfacing of IoT System

main requirement of our system. We have used Raspberry Pi 3 Model B+ for our system.

- **Raspberry Pi Camera:** Raspberry Pi has a ready made support for the Raspberry Pi camera modules which can be connected with flex cable. We have used Raspberry Pi Camera v3 module which can record full HD video at 50fps with an autofocus functionality.
- **Touch Sensor:** Touch sensors or tactile sensors are can detect touch and

Step 1: Initialize the system
Step 2: Check the status of touch sensor
Step 3: If touch is recognized on the touch sensor

(a) Capture a frame using Raspberry Pi camera
(b) Perform object detection on the frame using YOLO v3
(c) Measure the distance of the object using ultrasonic sensor
(d) Create a text containing object name and distance
(e) Convert the text to audio output using gTTS
(f) Feed the audio output to speaker/headphone

Step 4: Go back to Step 2

Fig. 4. Main Steps of Operation

operate as a switch. We have used TTP223 1-Channel Capacitive touch sensor module in our system which works as a trigger for the system.

The key software components of our system are:

- **OpenCV:** OpenCV is an open-source computer vision and machine learning software library. It provides a comprehensive set of tools for image and video processing, It performs tasks such as image and video analysis, object detection,face-recognition etc.
- **YOLO v3:** WE have used pretrained model of YOLO v3 for object detection. The model is trained for 80 different objects and is available at [13].
- **Text To Speech:** Text to speech (TTS) conversion provides real-time auditory feedback by converting information that is text-based, like the name of the detected object and its distance into spoken words. We used gTTS (Google Text-to-Speech) library [14] in our project.

The main steps of operation of the Trinetra system are depicted in Fig. 4.

4. Results

The photograph of real time implementation of our Trinetra system is shown in Fig. 5. The corresponding results of various objects detected with the confidence level and their distances are shown in Fig. 6.

It is observed that the system gives precise distance for the objects which are in line-of-site of the ultrasonic sensor. We have used 50% as a threshold level for the confidence level of object detection using YOLO v3. Hence, only these objects above threshold will be detected by our system.

Fig. 5. Real Time Implementation

Fig. 6. Outputs for Different Objects

5. Conclusion and Future Scope

The Trinetra system marks a significant milestone in addressing the needs of visually impaired individuals. By integrating object detection, depth estimation, and voice feedback technologies; it has created a comprehensive solution for enhancing safety and independence of visually impaired persons. The system's performance in identifying objects, providing real-time distance measurements, and simulating object interactions has yielded promising results.

The Trinetra system can focus on addressing existing challenges, improving system functionality, and expanding its capabilities to further benefit visually impaired individuals. Here are some potential areas for future development:

- Current system can give the distance of only one object which is in line-of-site of ultrasonic sensor. One can use other alternatives like stereo vision to find the distances of multiple objects.
- Current system works well for static objects. One can improvise it for moving objects as well.
- The overall system costs approximately Rs. 3500 which can be brought down to make it more affordable.
- Integration with wearable devices like smart glasses, which can provide a hands-free and intuitive means for visually impaired individuals to access Trinetra's information and feedback.

As the project continues to evolve and address its challenges, it holds promise as a transformative tool for the visually impaired community.

References

1. World Health Organization, World report of vision Available at: `https://www.who.int/publications/i/item/9789241516570`, Last accessed on - Feb. 2024.
2. E. D'Atri, C. M. Medaglia, A. Serbanati, U. B. Ceipidor, E. Panizzi and A. D'Atri, A system to aid blind people in the mobility: A usability test and its results, in *Second International Conference on Systems (ICONS'07)*, 2007, pp. 35–35.
3. L. Dunai, G. P. Fajarnes, V. S. Praderas, B. D. Garcia and I. L. Lengua, Real-time assistance prototype - a new navigation aid for blind people, in *IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society*, 2010, pp. 1173–1178.
4. N. Tyagi, D. Sharma, J. Singh, B. Sharma and S. Narang, Assistive navigation system for visually impaired and blind people: A review, in *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, 2021, pp. 1–5.
5. M. Bousbia-Salah and M. Bettayeb, A navigation aid for blind people, *Journal of Intelligent & Robotic Systems* **64** (2011) 387–400.
6. G. D. Marzullo, K.-H. Jo and D. Cáceres, Vision-based assistive navigation algorithm for blind and visually impaired people using monocular camera, in *2021 IEEE/SICE International Symposium on System Integration (SII)*, 2021, pp. 640–645.
7. Y. Liu and J. Aggarwal, 3.12 - local and global stereo methods, in *Handbook of Image and Video Processing*, ed. A. BOVIK (Academic Press, Burlington, 2005) pp. 297–308, Second Edition edn.

8. I. P. Adi, H. Kusuma and M. Attamimi, Blind people guidance system using stereo camera, in *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2019, pp. 298–303.
9. J. Zhong, M. Li, X. Liao, J. Qin, H. Zhang and Q. Guo, Real-time depth map estimation from infrared stereo images of rgb-d cameras, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **V-2-2021** (2021) 107–112.
10. A. V. Vanmali, S. G. Kelkar and V. M. Gadre, A novel approach for image dehazing combining visible-nir images, in *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, 2015, pp. 1–4.
11. Object detection methods Available at: `https://media.springernature.com/lw685/springer-static/image/art%3A10.1007%2Fs11042-022-13153-y/MediaObjects/11042_2022_13153_Fig3_HTML.png`, Last accessed on - Feb. 2024.
12. A. Ben Atitallah, Y. Said, M. A. Ben Atitallah, M. Albekairi, K. Kaaniche and S. Boubaker, An effective obstacle detection system using deep learning advantages to aid blind and visually impaired navigation, *Ain Shams Engineering Journal* **15**(2) (2024) p. 102387.
13. YOLO v3 Available at: `https://github.com/ultralytics/yolov3`, Last accessed on - Feb. 2024.
14. gTTS Available at: `https://gtts.readthedocs.io/en/latest/`, Last accessed on - Feb. 2024.