e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:02/February-2022 Impact Factor- 6.752 www.irjmets.com

REAL-TIME DYNAMIC OBSTACLE DETECTION FOR VISUALLY IMPAIRED PERSONS

Tejas Hari Aher*1, Govind Karvande*2, Tejas Uttam Aher*3, Amey Jadhav*4, Prof. Dr. S. V. Gumaste*5

*1,2,3,4,5Department Of Information Technology, MET Institute Of Engineering Nashik-422203, India.

ABSTRACT

In this paper, We describe a live object identification system that may be used by visually impaired people as a blind assist. People who are sight challenged rely largely on their other sensations including touch and hearing to understand the environment around them. Blind people often find it very difficult to determine what objects are in front of them without touching them (by hand or with another tool). The object detection system for sighted People is the most realistic technology that can assist blind people to experience their surroundings most effectively. This system assists them in Recognizing the objects in front of them and move from one destination to another. It also computes the gradient in between disabled person and indeed the obstacle via recognizing every obstacle. The reliability of obstacle detection approaches has substantially improved in terms of overall speediness and precision, leading towards subsequent improvements in deep learning. As a result, modern desktop computers are able to perform highly accurate object detection in real-time. In recent years, Among embedded technologies, there has been a surge for demand in implementing compact as well as speedier deep neural network layouts. This system analyzes the appropriateness of running Obstacle detection and classification with the help of the YOLO (You Look Only Once) algorithm on the Raspberry Pi 4 and Measuring the Distance of an object with voice feedback, a popular embedded computer board.

Keywords: Object Detection, Yolo, Cnn (Convolutional Neural Network), Computer Vision, Raspberry PI.

I. INTRODUCTION

Globally, almost one billion people have impaired vision, according to the World Health Organization(WHO).Blind Impairment and loss of vision, are caused primarily by refractive errors and opaqueness.

As a result, visually impaired persons always require a helper on a daily basis. This makes them incapable of living independently. Thus, a study on visually impaired people is becoming a real-time necessity. Hence, we are going to ease their lives in various ways.

In this study, We Introduced a unique approach for recognizing and detecting moving objects which will empower visually blind people in navigating comfortably. Moreover, We prefer complicated sequences with dynamic movements, preferably with such a moderate magnitude and sharp environmental changes generated mostly by the displacement of obstacles. A computer vision system allows the devices to examine and comprehend pictures and videos in addition to seeing them. Computer vision is among the most common applications of machine learning, in which a computer detects and identifies objects in real-time video. Currently Deep learning and convolutional neural network (CNN) improvements have enabled machines to recognize and trace objects in live footage with remarkable precision.. The distance of an object from the camera is measured by using the camera focal length, and it gives voice feedback. This system can be very helpful for people who are blind. When using this system, they can easily identify the objects or obstacles in front of them. The purpose of this report is to explore the feasibility of developing obstacle detector can be implemented on a solitary device computer, on Raspberry Pi B+, that can keep up with real-time frame rates calculated in FPS(Frames Per Second) while retaining extreme accuracy. An object- detecting system is hampered by a lack of computing power. On the basis of detection precision, training time, and output, One of the most efficient object detection algorithm is chosen and studied (FPS).

PURPOSE:

The major purpose of this study is to create an effective navigation aid for people with visual impairments which can recognize and detect the moving objects which will empower visually blind people in navigating comfortably. In addition to that, the main focus is on enabling visually impaired individuals to Move Freely in

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:02/February-2022 Impact Factor- 6.752 www.irjmets.com

the environment and live Independently. In the paper, a regression-based algorithm called YOLO (You Look Only Once) is recommended for Object Detection.

OBJECTIVE:

The following are the goals:

1) To develop a device that can assist blind people with recognizing their surroundings and moving freely throughout them.
2) To design a device capable of detecting dynamic objects.
3) In order to develop a device that will have greater accuracy than the current technology.
4) To provide better performance with minimal equipment.

II. LITERATURE REVIEW

This section of this study explains the project's conceptual justification, commencing with a description of object detection and moving on to evaluations in machine learning, deep learning, relevant surveys, and challenges faced by sight-challenged individuals. Various papers are available that contain information on object detection. Some referred papers are mentioned here.

1. The authors offered an approach of establishing a vision sight-based prosthetic and portable device which empowers people with blind sight with navigating activities amid in unfamiliar indoors situations in this work. As a result of this system, A route is conspired to enable a blind person to achieve the specific goals throughout secure and comfortable way by detecting open spots, impediments, and objects of intrigue, such as entryways, recliners, floors, and machines, among others. Specifically, Six modules compose up the whole system: Flooring diversification, intensity-based grid generation, object recognition, obstacle avoidance, motion tracking, and delivering sensory valuable feedback to the blind person are all part of the process. As a result of the stereo cameras collect a vast quantity of data generate, the actual facts that individuals move in dynamic environments, and the need for immediate feedback for the blind person (User),

The author defined two constraints: -

i. Using a general-purpose platform graphics processing unit (GPGPU), the profundity and tone data must be examined in real-time.
ii. It ought to be possible to process images on a wearable computer, which is a light, portable processing device. Furthermore, the Author has discussed tradeoffs between sensing, processing, and storage systems that should be able to fulfill these requirements appropriately in terms of system usability. The author intends to continue implementing and optimizing the core components based on our provisional outcomes in flooring classification and establishing a 2D tenacy grid.

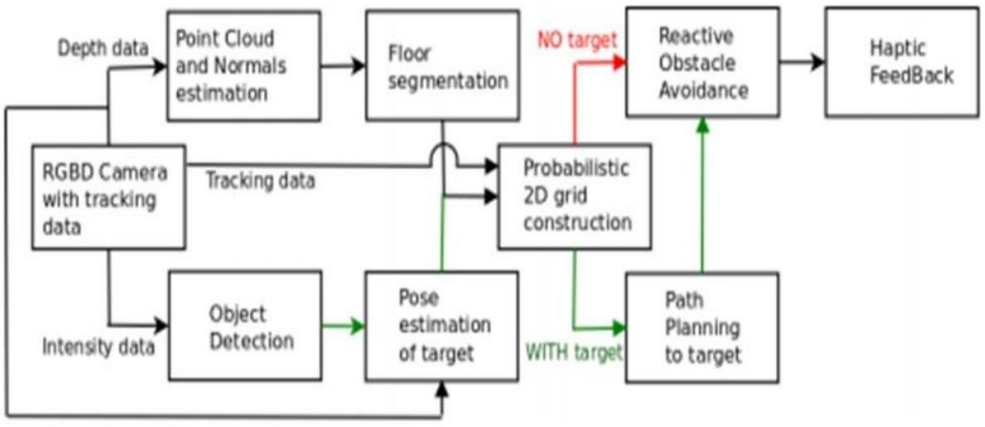

Figure 1: Overview of Author's System Architecture

Based on the results of this research, the author concluded that a system to assist blind people to navigate purposefully has been implemented and is showing progress. As part of the system, both depth and RGB information are provide by stereo cameras, as well as pose measurements since the camera is built with an optical-inertial algorithm for SLAM. The Jetson TX2 development kit is a supercomputer capable of parallel

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:02/February-2022 Impact Factor- 6.752 www.irjmets.com

programming as a co-processor. A haptic belt encoding motion commands are used to provide feedback through four vibrating motors fitted to the belt. [1]

2. Since the 1970s, technological advances in fields such as science, information technology, and communication have made it possible to develop advanced technology aids for the eyeless that surmount the challenges of relying on dogs and canes. Electronic substitution systems take a low-resolution image of a visual situation and turn it into a sensorial mode (tactile or auditory). Objects can be identified and located by assisting the eyeless in travel, navigation, and object identification, so enabling mobility.

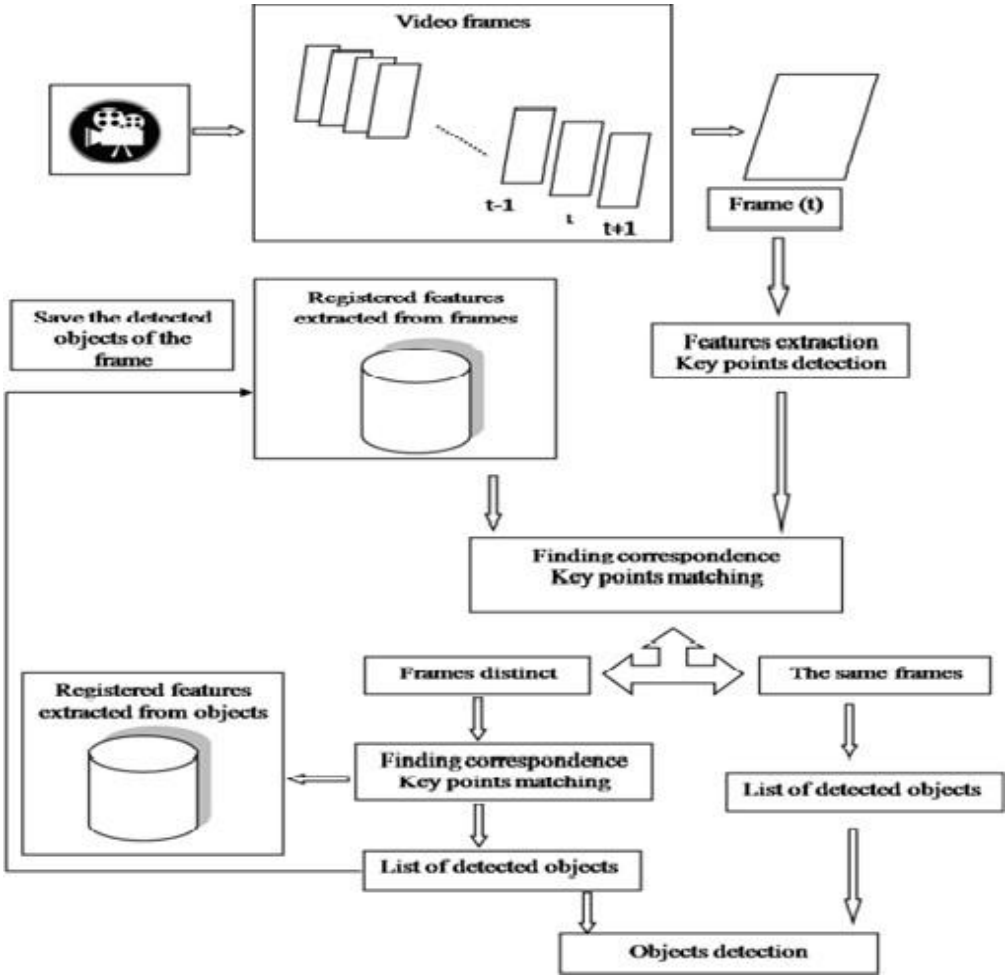

Figure 2: The object detection in video process

In this paper, the author proposed a system of substituting visualize information that restore a important function of the visual system's, namely, identifying objects in the environment. As a result, authors are interested in calculating fast and algorithms for robust for recognizing and locating obstacles in picture. Depends upon the robotic system to recognize obstacles in a video clip and the goal of this study is to build a technology that will help blind or visually impaired people detect their surroundings using an auditory system.

3. In this project, the author attempted to turn the visual world into an audio one so that blind people can be informed of objects and their locations. An object's name is converted to speech when detected from the scene. By using 3D binaural sound simulation, The 2-channel audio encodes their spatial placements. The authors' system has a number of modules. Video is recorded on the client-side and sent to the domain controller for real-time object identification (YOLO) using a portable camera device. The bounding boxes and the position of the identified objects are used to determine the 3D positioning of objects. The binaural digital sound with positions encodes is then rendered using the Unity game engine by a 3D sound creation program. Wireless headphones will be used to transmit digital sound.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:02/February-2022 Impact Factor- 6.752 www.irjmets.com

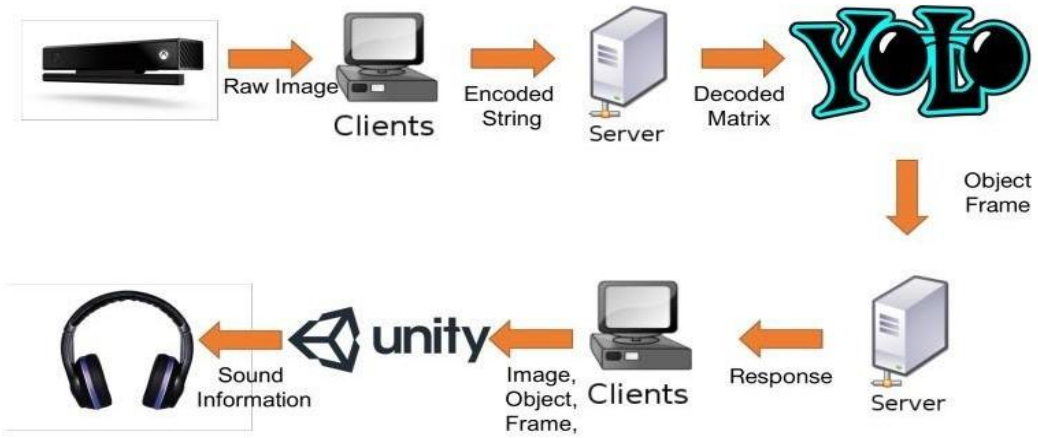

Figure 3: Initial Data flow pipeline

The author investigated the needs of persons who is blinded or visual impaired in this project. Based on CNN motivation, the author developed a blind simulation method that assists blind people in learning about their environment. The system has a compact, real-time interface. An author outlined a system that supports portable cams, a high-speed video connection, and a powerful server to produce 3D sounds. The YOLO algorithm combined with a wireless transmitter provided accurate real-time objective detection in 30 frames, 1080P resolution, and a live stream at a high speed.[3]

4. The authors of this study wanted to design a device that will help people who are blind or have lost their vision. The device is provided to help eyeless persons in overcoming difficulties that most people would dismiss as minor. Advanced Technology in computer science such as deep learning, wireless sensor networks, and computer vision were used to construct the device. A device would be created to help a blind person with daily tasks by recognizing and analyzing objects and providing audio feedback. The device should be able to distinguish multiple items in real-time, as well as provide vocal response on the object class to the user (Blind Person).

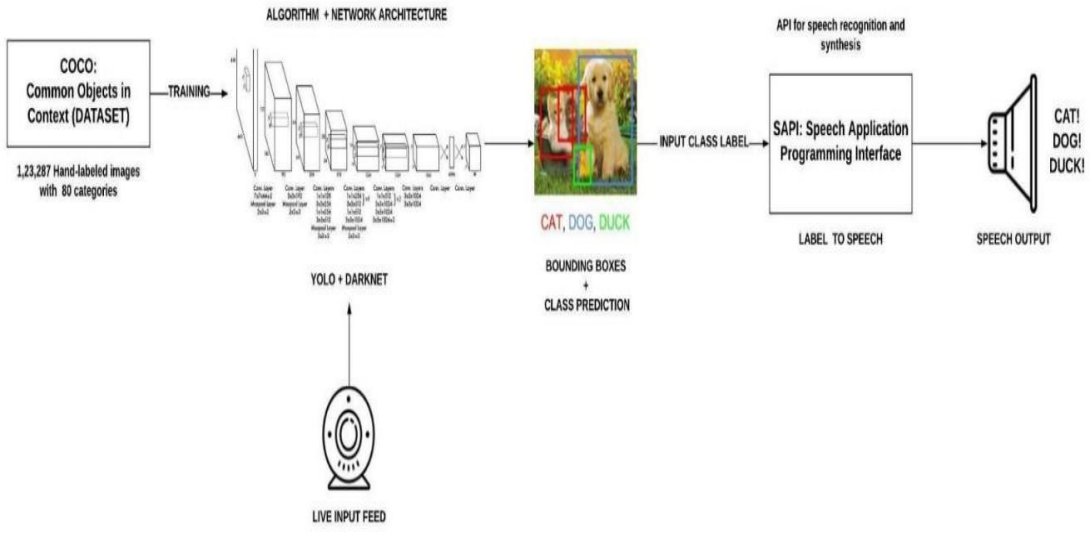

Figure 4: Object Classification & Real Time Detection Process

On the basis of these results, the author concluded that the dataset (COCO training dataset) was efficient when used in conjunction with the YOLOv3-tiny algorithm on the Nvidia Jetson Nano. Using the mentioned algorithm, the author was able to achieve the best FPS value with minor tradeoffs inaccuracy. Moreover, the author suggested that using a fan as a cooling solution for the device could improve its efficiency and performance. [4]

5. In this paper, the Author considered the vision substitute approach in this study among the visibility partial modification and intervention strategies. Vision substitution is a methodology during including one that information being engendered through sensation perhaps through the utilization of a sensorineural prosthesis, or perhaps a composite of both. Especially particularly in correlation with sensation and

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:02/February-2022 Impact Factor- 6.752 www.irjmets.com

perception, visual possesses a broader and deeper assortment of input, permitting the practitioner to handle much more diverse circumstances. As a result, the authors concentrated on the sight permutation strategy for the Visually Sight People. The concepts of eyesight substitution here could be distinguished into the following subsections:

1) Electronic Travel Aids (ETA): This is a strategy for enhancing the travel journey experience of the visual sight challenged people by integrating several kinds of information and being able to produce an outcome via the usage of sensing devices or auditory grants.

2) Electronic Orientation Aids (EOA): This method relates to the system's advice offered to the visually challenged user at every step of mobility.

3) Position Locator Devices (PLD): This approach entails the use of tools to assist with mobility or to monitor the whereabouts of a blind person.

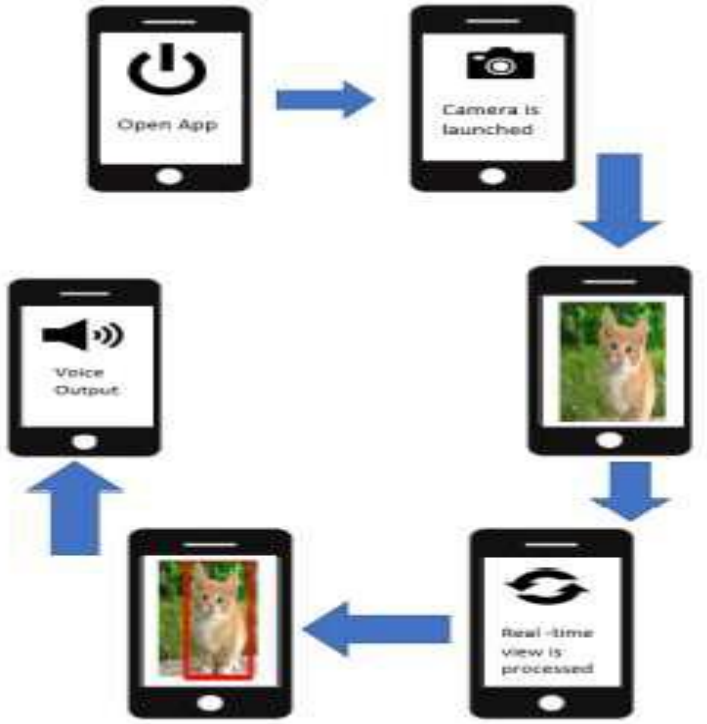

Fig 5: Android Application for Real-Time Object Detection

In this study, the author focused on the system's object detecting capacity in order to give users with efficiency.[5]

PROBLEM STATEMENT:

Visually impaired individuals are a growing segment in this era. Developing tools for vision-impaired people is not a recently emerged problem. The development of computer-aided tools is still a developing field for people with visual impairments. The challenges facing visually impaired individuals will be numerous, but one of the most common difficulties they face is when they involve self-navigating in an environment that is strange for them. In fact, physical movement is one of the biggest challenges for them. In addition, when they are traveling around or walking in a crowded place, they face many challenges. As visual impairments limit the ways the person can interact with others, therefore, a need for assisting aids with multitasking features to cope with different situations is an important issue. Currently, the available aids and technologies on the market are too expensive for the majority of users, especially those with normal income levels. Therefore, new devices with similar functions and cheaper prices are needed.

MOTIVATION:

Coming up with great ideas, trying new skills, finding ways to better something, understanding difficult and simple findings, and getting done in a team are all skills that can be learned. In our thoughts, the most main consideration is to concentrate on a topic that is significant to the daily needs of visually impaired persons.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:02/February-2022 Impact Factor- 6.752 www.irjmets.com

- Current Scenario:

1. A disabled person is dependent on another person.
2. Despite the blind man's stick, the stick is unable to detect obstacles on the path.
3. When someone uses a wheelchair, their safety is not guaranteed.

By keeping this in mind, our chosen topic is "Real-time dynamic obstacle detection for visually impaired people".

SYSTEM ARCHITECTURE

The following figure fig. [6] shows the system architecture of the System. Based on the figure, the architecture of the system is mainly composed of three components - a camera, a Raspberry Pi, and headphones. As soon as the system is turned on, the user will switch on the camera attached to it. With the camera getting started, Live Video Streaming has been taken. Immediately after taking the Live Streaming, the object will start being captured. Then, YOLO API is used to detect the object pattern from the captured scene using its pattern matching functionality. In order to figure out what type of object it is, the system uses a feature extraction process. Feature Extraction identifies the object using its features. By Using the focal length of the camera, the system will calculate the distance of the detected object and Print the Name and Distance in textual form. The Python Audio Library can be used to generate audio signals for the identified objects. As an Audio Output, the Generated Audio signal is sent to the Earphone, which is connected to the User's Ear.

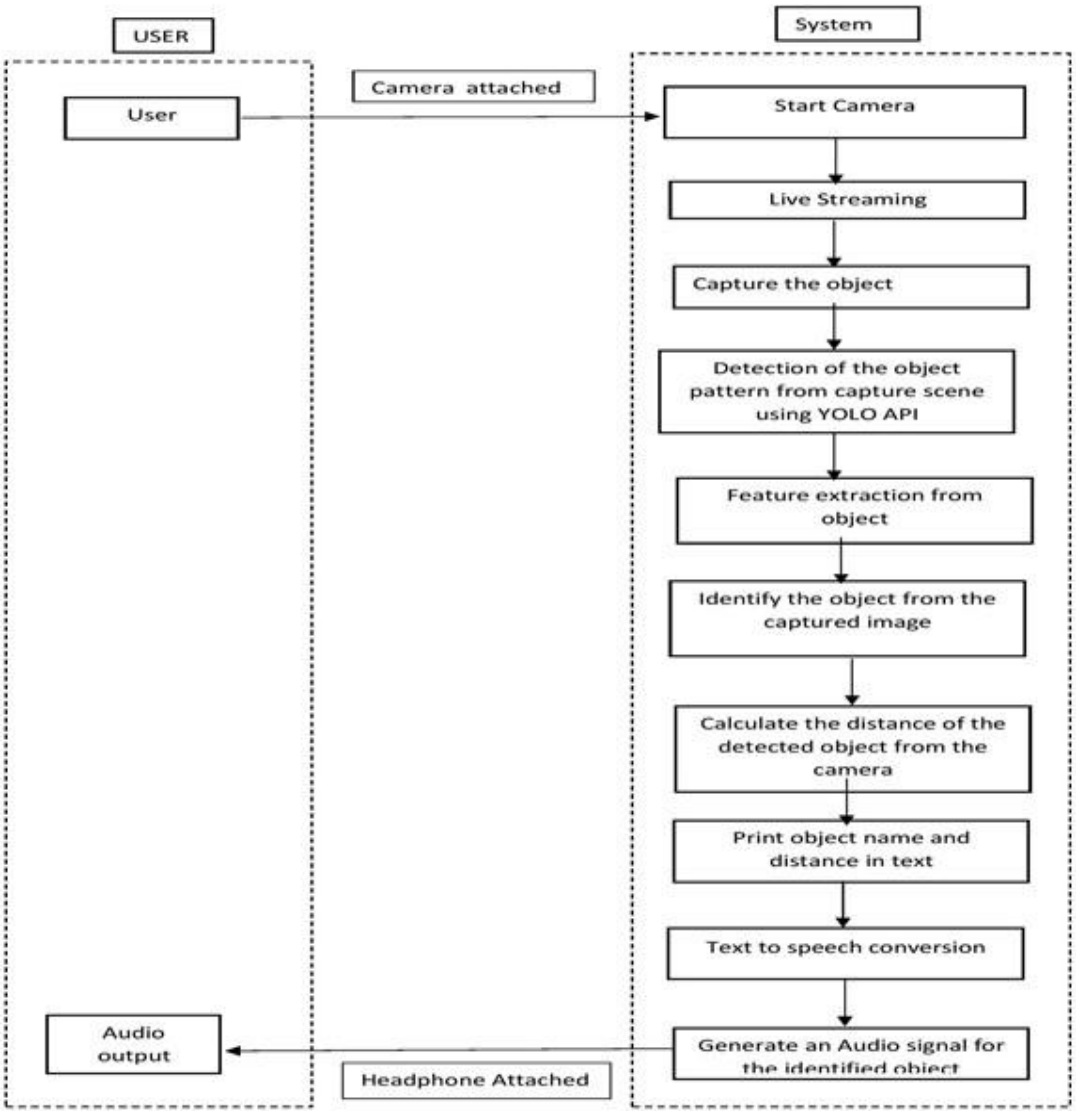

Figure 6: System Architecture

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:02/February-2022 Impact Factor- 6.752 www.irjmets.com

III. METHODOLOGY

SYSTEM IMPLEMENTATION:

We provide a method for aiding blind persons in safely exploring new outdoor situations until they find an object of interest. Raspberry Pi serves as the Processing Device, along with a camera providing frame rates that are high. The little device assesses if there are no barriers in the user's route in real time, recognizes things of interest, and computes a trajectory for reaching the objective in a safe way using a light and a small Processor speed. A combination of algorithms from machine learning and computer vision is used in the system. In order to calculate the distance, we have proposed that the focal length of the camera be used as a measurement. Moreover, as the environment can be dynamic, Priority is given to the time limitation.

ALGORITHMS:

For successful detection of surrounding objects, we investigated into a few current detection algorithms that could classify and analyse things in different parts of a picture. For images, the YOLO (You Look Only Once) algorithm generates possible bounding boxes using region proposals. In order to classify each box, it applies several Convolutional neural networks.

Furthermore, we discussed the YOLO algorithm –

1) YOLO (You Look Only Once):

YOLO detects objects in real-time using neural networks. Speed and accuracy make this algorithm very popular. The technology is used to identify traffic signals, pedestrians, parking metres, and animals in a variety of applications.

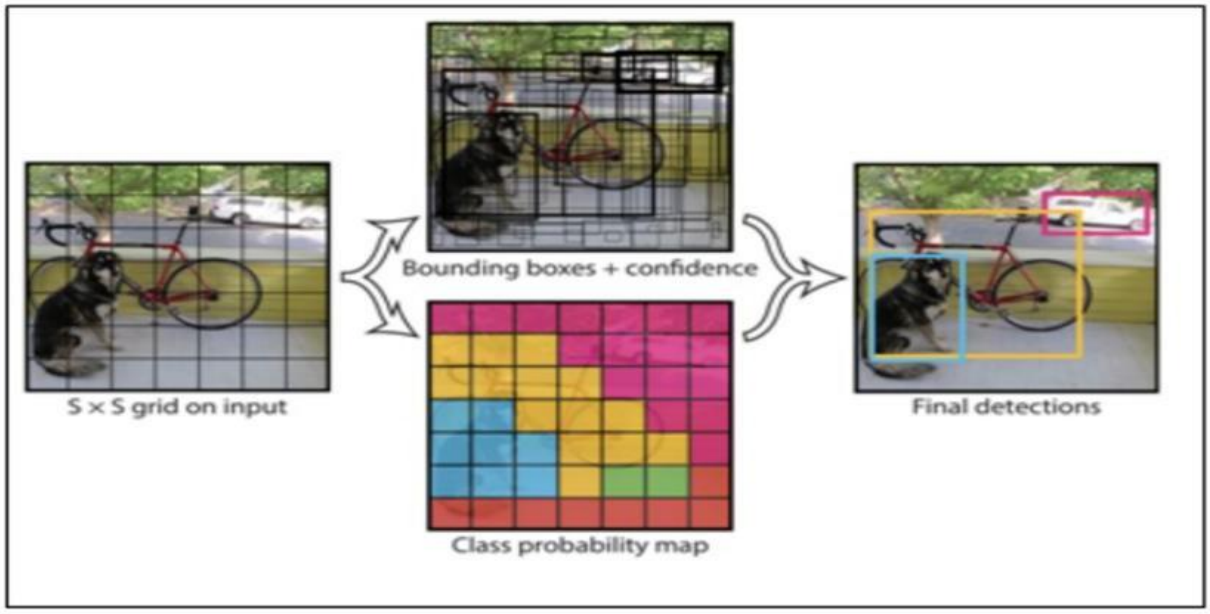

Figure 7: The YOLO Model

Steps-

1. A YOLO network is built by dividing an input image into a grid of SxS cells in which it marked m bounding boxes. The network predicts class probability and offset values using these bounding boxes.
2. Networks predict class probabilities and offset values in these bounding boxes.
3. These bounding boxes exceed the threshold value for class probability and are used to locate the object within the image.
4. The intersection over union (IOU) approach is used for this purpose.
5. YOLO detects objects in the image at a rate of 45 frames per second faster than the rest of the algorithms.

ADVANTAGES:

1. Instead of hundreds of evaluations for just one image, it can predict with a unified network evaluation. In comparison with other object detection algorithms, such as R-CNN and Fast R-CNN, the algorithm is 1,000

times and 100 times faster.

2. With a processing speed of 45 frames per second, YOLO is incredibly fast.
3. Based on one single network, a prediction (object locations and classes) is made. The model can be trained end-to-end to improve its accuracy.

IV. CONCLUSION

In this paper, we conclude that no system exists that acknowledges all of the appropriate features with the optimum level of precision and efficiency. so each system is distinct and individualized in its specific way, and no single system can suit all of the necessities of people who are visually impaired. As a result, this comes as a challenge for us to develop a system that determines all of the necessary factors, such as dynamic obstacle detection, text generation, and converting that text into speech, along with considering indoor/outdoor situations, angle of elevation, and daytime conditions. Furthermore, the perfect system must be designed so that the visually challenged people should be assured certain of its authenticity, wireless features, overall performance.

V. FUTURE SCOPE

- Providing a GPS locator
- Wireless charging to recharge the battery
- Detection of objects in full dark mode

VI. REFERENCES

[1] Andrés A. Díaz, Sixto E. Campaña, Eduardo F. Caicedo, “Methodology to Build a Wearable System for Assisting Blind People in Purposeful Navigation”, 2020 3rd International Conference on Information Technology and Computer Technologies (ICICT).

[2] Hanen Jabnoun, Faouzi Benzarti, Hamid Amiri, “Visual substitution system for blind people based on SIFT description”, 2014 International Conference of Soft Computing and Pattern Recognition.

[3] Rui Jiang, Qian Lin, Shuhui Qu, “Let Blind People See: Real-Time Visual Recognition with Results Converted to 3D Audio-”, 2016 Research Gate.

[4] M. Vaidhehi, Varun Seth, Bhavya Singhal, “Real-Time Object Detection for Aiding Visually Impaired using Deep Learning.”, 2020 International Journal of Engineering and Advanced Technology (IJEAT).

[5] Bhagyashree vidhate, Bhushan Bari, Utkarsha Kale, Vandan Vadnere, “Virtual Sight Mobility for Visually Impaired – A Survey”, 2021

[6] Sunit Vaidya, Naisha Shah, Niti Shah, Prof. Radha Shankarmani, “Real-Time Object Detection for Visually Challenged People”, 2020 Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020).

[7] S. TOSUN and E. KARAARSLAN, “Real-Time Object Detection Application for Visually Impaired People: Third Eye," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, 2018, pp. 1-6. doi: 10.1109/IDAP.2018.8620773

[8] B. Jiang , J. Yang, Z. Lv and H. Song, "Wearable Vision Assistance System Based on Binocular Sensors for Visually Impaired Users," in IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1375-1383, April 2019.

[9] TH. Wang, R. Katzschmann, S. Teng, B. Araki, L. Giarre, and D. Rus, Ā Enabling independent navigation for visually impaired people through a wearable vision-based feedback system,ā in 2017 IEEE International Conf. on Robotics and Automation (ICRA), 2017

[10] Ross Girshick. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, pages 1440–1448, 2015.

[11] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640, 2015.