

Dark Assistant: A Raspberry Pi Based System for Object detection and Recognition for Blind

Praharsh Verma¹, Yogesh Sondulkar², Sameer Zaki Diwan³, Bikas Majumdar⁴, Swati Saxena⁵

U.G. Student, Computer Science & Engineering, ITM Vocational University, Vadodara, Gujarat, India^{1,2,3,4}

Assistant Professor, Computer Science & Engineering, ITM Vocational University, Vadodara, Gujarat, India⁵

ABSTRACT: The World Health Organization (WHO) fact reported that there are 285 million people with visual impairments worldwide. Of these people, there are 39 million blind people in the world. [7] Blind people face a number of problems in their lives, one of which is the most important problem is finding obstacles when traveling. Visual information is the basis of many navigation activities, so visually impaired people are at risk because the required information about their surroundings is not available. With the recent advances in integrated technology it is possible to extend the support provided by visual impairments during their travels. In this context we propose a program, called Dark Assistant, whose purpose is to give blind users the ability to navigate in unfamiliar environments, whether inside or outside, through a user-friendly interface.

KEYWORDS: Raspberry Pi, YOLO(You Only Look Once), Tensorflow, Object Detection

I. INTRODUCTION

People are witnessing a new era of deep learning and embedded tools. While image recognition, driverless cars and excellent preventive health care are all in the domain, real-life applications are on the horizon today or later. The primary motivation for leading our team is how to build a real-world application that benefits the "blind" of a section of society. These people face many problems in their life, and the most important obstacles they face are when they walk. World Health Organization (WHO) reports that 285 million people worldwide have vision impairments. There are 39 million blind people in the world. 82% of people 50 and older are blind. The main causes of vision deficiency are undetected refractive errors (43%) and cataracts (33%); Cataract (51%) is the first cause of blindness. 2010. Visual impairment is a major global health problem: preventable causes account for up to 80% of the total global burden. [7] Increased demand for auxiliary equipment for navigation and orientation. The simplest and most affordable navigators and tools available are trained dogs and white cane. Although these devices are very popular, they do not provide all the information and features for safe mobility available to those with blind vision. Therefore, we are trying to make a device that can find the object in real time using the Raspberry Pi machine. The library "Tensorflow" and Yolo (you only see one) algorithm and speak to the blind through a Bluetooth speaker connected to the device. It is important to create a system based on identifying objects to meet the major challenges of the local population in the needs of three main categories: localization, policy and object recognition [12].

II. BACKGROUND STUDY AND RELATED WORK

Object detection is a computer vision technique whose purpose is to detect objects such as cars, buildings and humans. Objects are usually identified from image or video feeds. Object detection is widely applied in video surveillance, self-driving cars and tracking of objects / persons. How Object Detection Works: Object Detection Detects the presence of an object in an image and draws a bond around the object. It usually consists of two processes; Sort and type the object, and then draw a box around the object. In the computer vision community, developing visual aids for the disabled is one of the most active research projects. The purpose of mobility aids is to describe the environment to the person who appreciates the surrounding objects.[8]

The most commonly used algorithms for object detection are:

- R-CNN
- SSD(Single Shot MultiBox Defender)
- YOLO(You Only Look Once)

YOLO(You only look once):

We offer a new method of object detection, Yolo. Restores classifiers to do previous work on object detection. YOLO models process 45 frames per second in real time. YOLO sees image discovery as a regression problem,



which simplifies its pipeline. Because of this simple pipeline, it is very fast. It can process streaming video in real time with less than 25 seconds delay. During the training process, YOLO looks at the whole picture, so context can be included in object detection.

In YOLO, each border box is evaluated by the attributes that come from the full image. There are 5 estimates for each boundary box; x, y, w, h, faith. (x, y) represents the center of the boundary box relative to the boundaries of the grid cell. W and h estimate the full width and height of the image. This model is implemented for the traditional neural network and is estimated in the Pascal VOC Discovery Dataset. The network's convolutional layers are responsible for capturing the features, while the fully connected layers coordinate and predict the output potential.

The network architecture for this model is inspired by the Google Net model for image classification. The network consists of 24 convolutional layers and 2 fully connected layers. One of the main challenges of this model is that it can only accommodate one class and it does not work well on smaller objects such as birds. This model achieves an average accuracy of 52.7%, however, it can go up to 63.4%.

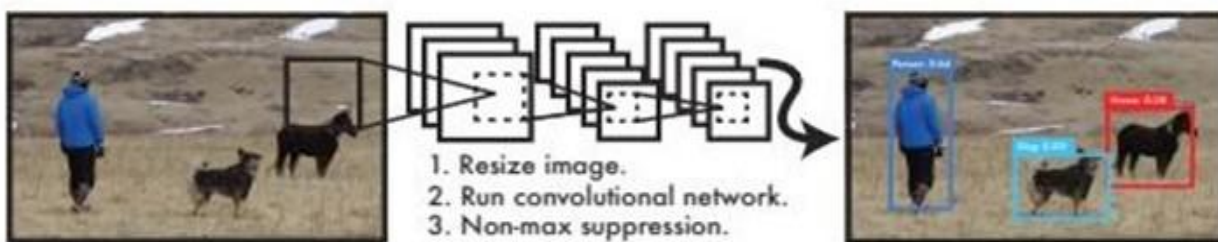


Figure 1: The Yolo Detection System. Processing images with Yolo is simple and straightforward. Our system (1) resizes the input image to 448 * 448, (2) runs a single convolutional network on the image, and (3) threshold the resulting detections by the model's confidence.

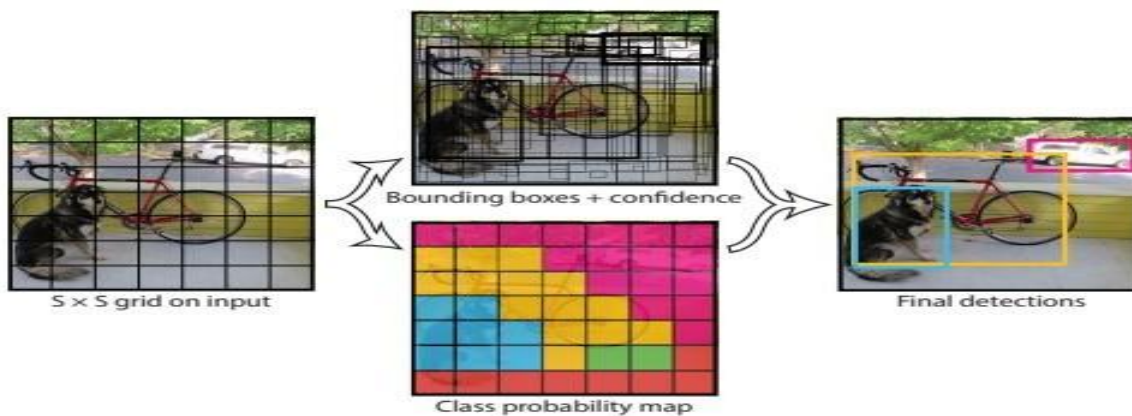


Figure 2 I/O System [6]

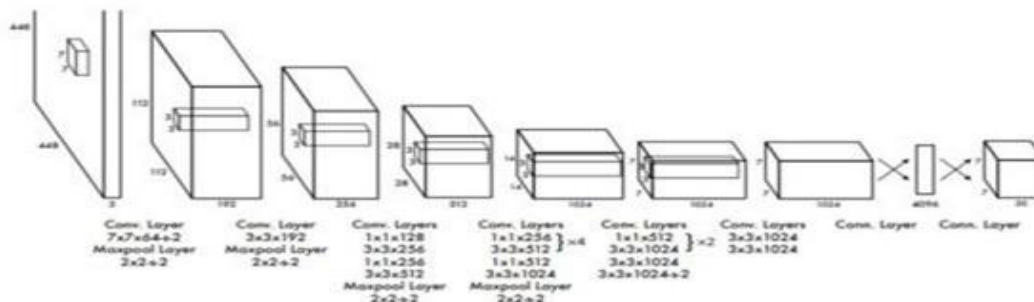


Figure 3: Architecture. Our Discovery Network consists of 24 convolutional layers followed by 2 fully connected layers. 1 * 1 convolutional layers substitute the space of properties from previous layers. We train the convolutional layers of the imagenet classification task at half resolution (224 * 224 input image) and double the resolution. [6]

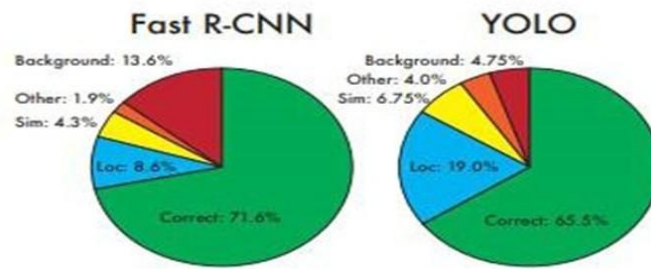


Figure 4: Error Analysis: Fast R-CNN vs. Yolo These maps show the percentage of localization and background errors in the top N detections for different categories (N = # of objects in that category).[6]

III. LITERATURE SURVEY:

Seema Udgirkar et. al [1] describe an Android app built for object detection using the YOLO (see once) algorithm for real-time objects, but the only difference is that the technology they use is the Android app and our system is a free hardware device. This may be an error, because the Android app can run multiple processes at once, which drains the battery. If calling on a cellphone, the application is unable to run, which is again an error as the blind cannot call the app at once. Although our system is a free hardware device, the above disadvantages can be solved.

Walking Kane is a simple and automated tool dedicated to finding permanent barriers, uneven surfaces and holes in the ground through simple touch-force feedback. This device is lightweight, limited in non-portable range and is not suitable for protection from obstructions near the head area [1].

Another option that provides great travel assistance for the blind is Guide Dogs. Based on the symbiosis between the disabled owner and his dog, training and contact with animals is critical to the success of this method. The dog can detect and analyze complex situations: cross-walking, stairs, potential danger, knowing ways and more. Much information is exchanged through feedback that touches the handle mounted on the animal. The user can feel the attitude of his dog, analyze the situation and give appropriate directions. Guide dogs are still far from affordable, with the cost of a good car, and their average working hours are limited to an average of 7 years [1].

Rohini Bharti et. al proposed [2] that they use Raspberry Pi, which is the technology we use, but the algorithm they use is Fast R-CNN (convolution neural network), which is not a good detection method, but great for objects that appear due to background patches in an object. Instead, the YOLO (see only once) algorithm uses less than half of the background errors compared to the fastest R-CNN. It is all the boundary boxes of all classes at once for the image. YOLO is faster in object detection algorithms.

Patil Nandini Jalandar et. al [3] proposes object detection and processing of images using the SURF and SIFT algorithm, but in some cases the only error is the error, and the distance between them cannot be measured. In our project we use Augmented Reality to measure the distance between objects and the blind.

In [4] Yiwei Wang et al suggest an algorithm to isolate moving objects in the video, and then introduce a rule-based tracking algorithm. Preliminary test results show the effectiveness of the algorithm in some complex situations, such as new testing, discontinued tracking, and track collision. The tracking method that does not capture the background is discussed in [5].

The author introduces video tracking in computer vision, including design requirements and learning the patterns of texture and dynamics, from simple window tracking to tracking complex, clumsy objects [9].

IV. PROPOSED METHODOLOGY

Materials used: Camera Pi, Raspberry Pi 3B, YOLO Algorithm, ImageNet(Predefined model), Tensorflow for training our own model, and a power bank to use Raspberry Pi in real time.

Our project could be explained by the given image below. The Project Dark Assistant, as explained by the image given below is for blind person whose specs has been integrated with the the cameraPI. The cameraPI records the videos in the real time. Those videos are sent to the RaspberryPi.

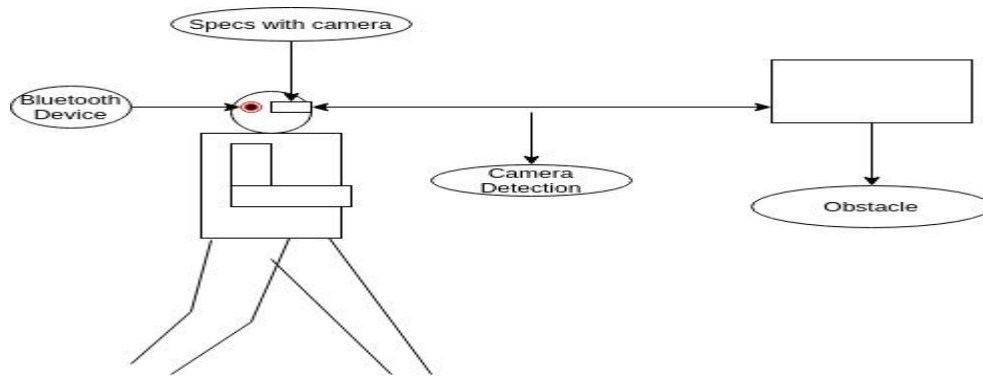


Figure 5

The Raspbian is loaded as an operating system in RaspberryPi. Tensorflow is installed on the operating system. Our code is loaded into the OS for object detection.

The code works as follows:

1. The recorded video is divided into frames.
2. Compare these frames with the pretrain model ImageNet, which can now detect 74 objects.
3. We used the YOLO (you only see one) algorithm comparing the frames to 45 frames per second.
4. If the match is found, the corresponding keyword is sent as output.
5. Then the text output passes through the text to the speech module.
6. Then the output is sent to the Bluetooth headphones with the blind. So the blind man can know the obstacle ahead of him.

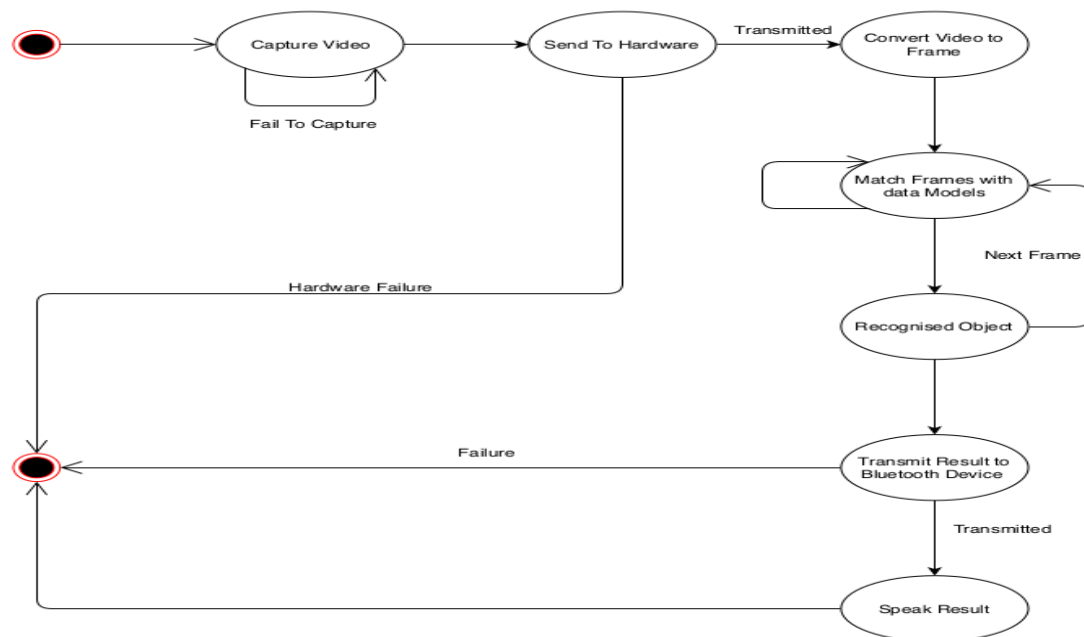


Figure 6: State Diagram of the project “Dark Assistant”

V. CONCLUSIONS

The specific system design is simple, easy to install, and somewhat reliable. The use of "rasberry pie" optimized the system and significantly reduced the cost compared to conventional systems. The YOLO algorithm minimizes fast object detection algorithm with object detection time, which is precisely designed, coded and tested on a laptop, minimizing errors and delays. Above all, the advanced system develops a technical approach to timely addressing the real-life problem.

REFERENCES

1. Object Detection System for Blind People by Prof. Seema Udgirkar et. al; on International Journal of Innovative Research in Computer and Communication Engineering(IJIRCCE) on Vol. 4, Issue: 9September'2016.
2. Object Detection and Recognition for Blind Assistance by Rohini Bharti et. al; on International Research Journal of Engineering and Technology (IRJET) Volume: 06 Issue: 5 May'2019.
3. Object identification using picture for blind peoples by Patil Nandini Jalandar et. al; on International Journal of Advance Research, Ideas and Innovations in Technology(IJARIT) (Volume 4, Issue3).
4. Yiwei Wang et al; "Moving Object Tracking in Video", in proceedings of 29th applied imagery pattern recognition workshop, ISBN 0-7695-0978-9, page 95,2000.
5. Bhavana C. Bendale et al; "Moving Object Tracking in Video Using MATLAB", and International Journal of Electronics, Communication and Soft Computing Science and Engineering ISSN: 2277-9477, Volume 2, Issue1.
6. "You Only Look Once: Unified, Real-Time Object Detection" published in CVPR 2016. This paper is "You Only Look Once: Unified, Real-Time Object Detection" published in CVPR2016.
7. Pascolini D, Mariotti SPM. Global estimates of visual impairment: 2010. British Journal Ophthalmology Online First published December 1, 2011 as 10.1136/bjophthalmol-2011- 300539.
8. "Visual substitution system for blind people based on SIFT description" by H. Jabnoun et. al;at International Conference of Soft Computing and Pattern Recognition in Tunisia, August 2014.
9. "Video-Based People tracking"by Marcus A. Brubaker et al; in hand bookof ambient intelligence under smart environments 2010.
10. "Computer visionbased object recognition for the visually impaired in an indoors environment": a survey by Iafri et al; 2013 in The Visual Computer, Springer Berlin Heidelberg,
11. H. Bay et al; "SURF: Speeded Up Robust Features"Computer Vision and Image Understanding (CVIU), 110:346359, 2008.
12. Jabnoun H. et al; in 2014. "Visual substitution system for blind people based on SIFT description"in International Conference of Soft Computing and Pattern Recognition. Tunisia.