

# Smart Guiding Application for Visually Impaired People Using ML

Shailesh Yadav, Rahul Pal, Rishabh Saxena, Shobha Sharma

[shailesh.19b311096@abes.ac.in](mailto:shailesh.19b311096@abes.ac.in), [rahul.19b311035@abes.ac.in](mailto:rahul.19b311035@abes.ac.in),

[rishabh.19b311013@abes.ac.in](mailto:rishabh.19b311013@abes.ac.in), [shobha.sharma@abes.ac.in](mailto:shobha.sharma@abes.ac.in)

Department of Electronics and Communication Engineering

ABES Engineering College, Ghaziabad, Uttar Pradesh

**Abstract--** This work provides solutions for people suffering from partial blindness due to eye diseases and accidents. Many people are born blind. Our project brings blessings to people like them. This project transforms the visual world into the audio world by informing a blind person about the objects. Our project uses the YOLO algorithm. Real-time objects in images are recognized by the names displayed in their bounding boxes, and these names are translated into audio. Speech conversion is done using the e-Speak tool, Google's text-to-speech (gTTS) system. The prototype consists of several modules. A Raspberry Pi Camera Module captures the image and transfers it to the Raspberry Pi desktop. Detection of objects is then performed in real-time using the YOLO network.

**Keywords--**YOLO, E-speak tool, Visual Word, Raspberry pi, Headset.

## 1. INTRODUCTION

Today, 10 lacs of people throughout world experience vision impairment, primarily as a result of age issues, sugar, and accidental issue. One of the top 10 disabilities for males and 4,444 women is blindness or a visual impairment. The inability to carry out home tasks, go shopping alone, or finish simple task to-do lists are difficulties that the partially sighted and visually impaired must overcome. They also experience a loss of flexibility and initiative. It also covers the time and effort needed to use the new technology

that are currently being offered in an effort to assist people lead regular lives. The new technologies are evolving quickly, however some auto-navigation issues still exist because of the social discomfort of. Prior technologies include NAVI, a navigation aid for the blind person, that can distinguish the background and obstacles, VOICE, which allows the system to convert images into sounds to help users recognize obstacles, TVS, Tyflos system, etc. System device on the person's abdomen. Convert depth in the form of vibration. These system are very slow and practical unusable thing. A cane is used as the planned prototype, and photos are captured using a Raspberry Pi camera affixed to the cane. The You Look Only Once (YOLO) technique is used to detect intelligent objects [9], [10]. This is a brand-new, slashing technique for object detection through regression and classification. There are several methods for detecting objects, but they are not as precise as YOLO. The neural network only processes the images once, producing an output that includes all objects that were identified [11]. The next phase entails transforming text into speech and generating human voice using Google's gTTS module. Blind is able to access the information he needs with the aid of a headset. Consequently, you may offer an efficient camera-based self-navigation.

## 2. LITERATURE SURVEY

Similar to a guide dog trained to guide the blind to specific locations, the self learning walking-rod can assist you avoid objects in front of you.

The camera zoom-based system of [1] helping to recognize text on handheld item. The system is made of a photo-clicker that always take picture of things of their interested field and an process that extracts text from the back of it.OCR differs recognized text by her characters. The recognized text is output as speech output using the software developing item or device The easy carrying device of [2] are a system that receives input takes from the device user and recognizes objects. The is equipped with an ultrasonic sensor to help alert the user to objects in their path. Recognize objects using the Haar cascade algorithm. The images are captured with a carry-able camera, the images are converting to gray levels and filtering by GaOssian filtering process. Then binarize with trim. The cut image is passed to Teseract OCR. e-Speak generates an analg signal of the word format and feeds it into the earing device.

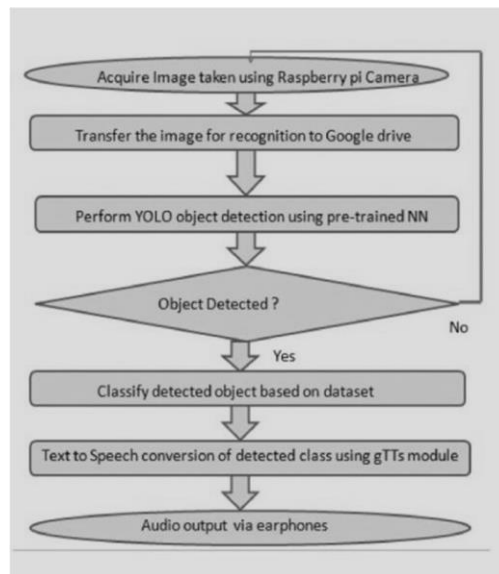
some scientists presented their hard work on "Blind assist Systems Using object Detection" in 2021 In this technical work, the function of Computer vision is the 100 percent accurate detection of indoor items. The visually impaired can be helped by moving through the CNN frames lenses.4,5,14 To check whether or identify specific objects first on priority basis, we need to detect the existing pixels in the image of device . If the lighting conditions are poor, it is difficult to capture and identify the subject with great accuracy. To detect inner objects, the algorithm needs to extract features from the image with a specific class and this can be done by Retina-Net.25 To allow the network to detect small objects by Subject area export 1. : Object detection using machine learning for visually impaired networks which involves down sampling to get image information. Resort with 152 samples achieved average accuracy with 83.12% and Dense Net with I22 samples achieved average accuracy with 79.89%.

other group of scientists proposed "Objects Identification for Visually Impaired person" . Based on relationship patterns, this work was assigned an equal amount of work. considering its characteristics. This eliminates duplication

and achieves accuracy for specific standards. Because objects are aligned to the 2-D scale, it uses objects instead of words.the model is classified into two components of geometric weight and original weight.

### 3. METHODOLOGY

The goal of our project is to always create a low-cost item that allows people with poor vision to spend normal lives without putting more dependencies on other people, in specific cases or environments such as at home or at work. Everything one uses in its daily daily work can be detected by a walking stick with a Raspberry Pi is usually attached. U see the diagram as mentioned.



**Fig-1.** diagrammatic description

Capture images using Raspberry Pi-4 Model B. Captured pics's are replicated to the Google's drive.Our algo networks are used for item identifying process and classifying process. Recognized class testing speech transcribe is performed using gTTS and sent to user's headphones.

### 4. RASPBERRY PI CONFIGURATION

To start our technical process of identifying with your Raspb-Pi, you first get an operating systems.

NOOBS is a pre-configured and simple OS installation managing software for the Raspberry Pi. NOOBS includes the following OS Raspbian, Pidora, penELEC, RaspBMC, RISC OS as follows.

1. NOOBS can be downloaded from the official websites [5].
2. you need to put the SD memory card format software and install it in your Windows OS to ensure compatibility of your SD card for putting this coded-program.
3. Clear all the SD card to FAT format with the above software. Your micro-small SD card is now available to use your software.
4. you need to update your OS with the current-version of your Raspberry Pi. To do this, two commands are triggered: a-`sudo apt-get updating`: b-`sudo apt-get upgrading`:
5. You have successful ready the Raspbian OS by following the steps above.



**Fig-2.** camera setup.

## 5. YOLO ALGORITHM

This process[6], [8] is mainly used to accurately predict bounding boxes from images. The images are classifying into  $N \times N$  grids, and bounding box prediction and class probabilities are performed for each grid. YOLO is one of the foremost and vivid popular object detection algorithm and model architectures. These software uses best neural networking structures to produce highly accurate and entire processing speed, which is prime reason why it is so popular. If we discover for article detection algorithm on

Google, the first outcome will be related to YOLO. The YOLO algorithm aims to predict a class of the object and the bounding box defines the position of the object on the input image..

Images can vary in complexity and can be split in number of rasters. After segmentation of image, both classifying and localizing processes are performed for each neural containing objects. trust values are calculated for all neural. Confidence values and bounding boxes for each of the grids depend on whether an items is detected. If there is not object, the values will be shown as 0 and if the object exists, the value will be shown as 1. The bounding box values indicate the security of the network, that is, how well the detected objects match the observed objects. Bounding box projections are shown below.

## 6. EXPERIMENTAL PROCEDURE

Process started by taking the data base, learning the YOLO network after that providing test images to test the output. The network recognizes five classes: C1 for bottles (class 1), C2 for watches (class 2), C3 for keys (class 3), C4 for pens (class 4), and C5 for glasses (class 5). trained to do so.

### *Collection of Data Base*

The Real life relatables of a data set for a An Class 5 pen is made. Similar datasets were taken for other rooms.



**Fig-3.** data-base for pens

## *Object Identifying Techniques Using HandEngineering Features.*

They are having five basic ways to item identifying process, named as:

### *a. On basis of algo match:*

This way is used in goal recognising platforms. Whenever we give an image in the form of input to our software-system, it is placed to check similarity to Sample images are stored to identify objects in the fig.[4].

### *b. On basis of colour*

These procedure employs colour properties in the way of an extra trait as well as shape-only features for object detection. He provides better results when detecting objects using shapes only feature failure. Object are discovered by representing and image similarity base-on-colour histogram. [5].

### *c. On the scanner methodology*

Methodology include scanning-process of images to locate in an image. However passive scan-process is never involved in the extracting of patterns at the time digitization [6].

### *Two-passive scanning methodology:*

i) Slid-window technique: always checks if any item has present or no at all coordinates of equally spacing-grid. One test-sample are checked at every grid, after that Classifying-process is accomplished according to which if it is an solid-item or back of an image[7].

ii) Partiality based reach: this define point of notice in a Picture. It find-out the values of interest for local input sample in all coordinates in a grid are evenly spaced, after that preferences point, a newly created local sample evaluation is checked if this has belonging to the obstacle or background [8]. Active scan-process allows localized samples that will be used as a guide for scan-process. At this place the sampled picture is taken out and matching out to an offset-vector that tells the sign of another scanning coordinate.

Consecutive sample are selected, ignore areas that are un not likely to accumulate objects. These are computation and mathematical task and tells better and enhance attribute detect.

## **7. TEXT TO SPEECH**

Getting speech output of the identified text is the project's primary objective. Text recognition makes use of the Tesseract OCR optical character recognition engine. As a result, the Tesseract text files must be translate to audio form, which is done using the Python gTTS module, a text-to-speech engine that turns text words extracted using Tesseract OCR into phonetic form. Running as a small hardware implementation, the entire process is carried out on the Raspberry Pi. 85% of the time, text-to-speech and text-to-image conversions are accurate. Google Text to Speech, a Python interface to Google's Text to Speech API, is utilised (Application Package Interface). MP3 files are made using the gTTS module. gTTS is substantially faster at converting audio.

## **7. RESULT**

The all these network was given proper training to recognize 5 custom device. The given items is in fig-4. and the objects identify by the YOLO network are shown in Fig-5.



**Fig. 4.** Manyitems in picture.

As we see in Fig. 5, the Yolo v3 custom way was used to recognize V classes of it was trained on. Since the network was trained on epoch 10000, the confidence score of the detected objects is relatively high.



**Fig-5. Identified things**

If the loss value of the ideal detector model should be less than 2, then the exponentially decaying loss curve gives a higher confidence value 2

his article presents an overall of the development of technique used for object recognition. Are from manual feature-based classifier for convolution networking classifiers in deep-learning. It also provides datasets are often used for training, testing, and object validation with new models.

## 8. CONCLUSION

Prime goal of our project is to create an object recognition that can identify specific classes. This is created for 80 classes of various products that user is probably going to utilise every day. A camera was used to create the suggested system, together with a raspberry pi for the interface and Bluetooth for audio output. The walking stick is coupled to an ultrasonic sensor for obstacle detection. Yolo causes run times to be shorter [7] and the IoU result measure is above 91%. Futuristic plans call for building cards and GPUs into hardware and running the network on the cloud, providing partially sighted persons complete independence both inside and out. included.

## REFERENCES

[1] Samruddhi Deshpande and Revati Shriram “Real time text detection and recognition on hand held objects to assist blind people” International Conference on Automatic Control and Dynamic

OptimizationTechniques)International Institute of Information Technology (IIT), Pune, 2016.

[2] Deepti Jain BS, Shweta Mitra Thakur and K V Sureshi “Visual assist for visual impaired using image processing” *International Conference on Communication and Signal Processing*, April 3-5, India, 2018

[3] Jing Tao, Hong bo Wang, Xinyu Zhang, Xiaoyu Li and Huawei Yang, “An object detection system based on YOLO in traffic scene” , *6th International Conference on Computer Science and Network Technology (ICCSNT)*, 2017

[4] Rajesh , Bindhu Rajan, Ajay Roy, Almaria Thomas K, Ancy Thomas, BincyTharakan T Dinesh C, “Text recognition and face detection aid for visually impaired person using raspberry pi”, *International Conference on Circuits Power and Computing Technologies (ICCPCT)*, 2017.

[5] Installing operating system in Raspberry pi available at <https://www.raspberrypi.org/downloads>.

[6] Geethapriya. S, Duraimurugan and S.P. Chokkalingam “Real-Time Object Detection with Yolo”, *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249 – 8958, Volume-8, Issue-3S, February 2019 .

[7] Deepa, Tamilselvan, Abrar and Shrinivas Sampath “Comparison of Yolo, SSD, Faster RCNN for Real Time Tennis Ball Tracking for Action Decision Networks”, *International Conference on Advances in Computing and Communication Engineering (ICACCE)*, 2019 .

[8] Joseph Redmon \*, Santosh Divvalam †, Ross Girshick and Ali Farhadi “You Only Look Once: Unified, Real-Time Object Detection”, *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[9] Seema Udgir , Shivaji Sarokar<sup>2</sup> , Sujit Gore , Dinesh Kakuste and Suraj Chaskar “Object Detection System for Blind People”, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 4, Issue 9, September 2016.

[10] Zou, X.Y., Gong, W., Fu, W.L., Du, FT, “Applied deep learning in object detection.” *IEEE/ACIS 16th International Conference on Computer and Information Science*. pp. 631–634 ,2017.