ISSN No:-2456-2165

YoloV4 Based Object Detection for Blind Stick

Mahesh Pawaskar
Computer Science & Engineering (AIML)
A. P. Shah Institute of Technology,
Thane, India

Sahil Talathi
Electronics & Telecommunication Engineering
A. P. Shah Institute of Technology,
Thane, India

Shraddha Shinde
Computer Science & Engineering (AIML)
A. P. Shah Institute of Technology,
Thane, India

Digvijay Singh Deora
Electronics & Telecommunication Engineering
A. P. Shah Institute of Technology,
Thane, India

Adesh Hardas
Electronics & Telecommunication Engineering
A. P. Shah Institute of Technology,
Thane, India

Vrushali Devlekar
Electronics & Telecommunication Engineering
A. P. Shah Institute of Technology,
Thane, India

Abstract:- Blind humans face many problems to interact with their close by surroundings. The intention of this paper is to offer a device to help blind humans to navigate in addition to feel the barriers. We plan to propose an operating model that is taking walks stick with in-constructed ultrasonic sensor with a micro controller gadget. Detection and monitoring algorithms are laid out in terms of extricating the capabilities of photographs and motion pictures for protection and scrutiny applications. famous algorithms of item detection consist of You only look once (YOLO), area-based Convolutional Neural Networks (RCNN), quicker RCNN (F-RCNN). RCNN has higher accuracy in comparison to different algorithms, but YOLO surpasses whilst pace is considered over accuracy.

Keywords:- *YOLOv4, Raspberry Pi, RCNN, Blind stick, Object Detection.*

I. INTRODUCTION

Need and Definition of ML based Blind Stick. Eye is the most significant part of the body. The vision helps us to obtain the environmental information. Blindness is a condition in which a person is unable to see and detect things happening in his/her surrounding can may lead to various problems which cannot be solved by medical means. There are many people with severe vision impairment that restricts them from travelling individually over their path. These blind people should have access to a range of tools which will help them travelling independently in their path. One of the oldest tools for blind people have been the walking stick also knows as the wide stick. They proved to be very useful back in time but now it has some significant problems. The rapid growth of modern technology has introduced better systems such as smart guided stick that can provide intelligent navigation to the blind person. One of the most visceral parts of computer science includes computer vision. Artificial Intelligence based smart guide stick, furnished with image detection technologies that take over front side and back side images and use Machine Learning to operate them. India been the most populated country in the world about 20 percent of its population is blind or visually impaired. A tiny rectangular box including raspberry pi, Bluetooth speaker, and battery bank will be designed to be fitted with a cane that is typically roughly 55 inches long. The Raspberry pi model 4 with 4GB RAM is used. Yola v4 algorithm is used for object recognition, to warn the user of any obstructions in the road, a Bluetooth speaker module is integrated. During the navigation phase, a power bank is inserted to serve as the raspberry pi's source of power.

II. LITERATURE SURVEY

A. Selecting a Template

This proposed method uses the Arduino UNO as a controller. The branch is accomplished by sensing all difficulties in front of the user. [6] The instrument stands used to perceive the obstacles at the range of four meters and the infrared instrument is castoff to perceive the nearer complications in front of the blind people.

Stick with in-built ultrasonic sensor with a microcontroller system. [9] The ultrasonic sensor is used to detect obstacles using ultrasonic waves. On sensing obstacles, the sensor passes the data to the microcontroller. The microcontroller then processes the data and calculates if an obstacle is close enough. Design and implementation of an ultrasonic sensor based walking stick for visually impaired person. [3] an ultra sonic sensor module, HC-SR04 is used for obstacle detection in the path of the blind person and a buzzer is use to make the person alert. the proposed system is implemented using PIC microcontroller 16F877A. The project was published which used ultrasonic, [10] infrared and water sensors to detect any objects within 4 meters very quickly. The stick is integrated with various sensors like ultrasonic sensor, [1] water sensor with GPS-GSM module and RF module and with microcontroller. This paper focuses on deep

ISSN No:-2456-2165

learning and how it is applied to detect and track the objects. Deep learning works with the algorithms influenced by the layout and functionalities of the brain. The advantage of working with such algorithms is that the performance increases with increase in data which isn't the case for traditional learning algorithms whose performance stabilizes even with increase in the amount data. [5] Popular algorithms of object detection include You Only Look Once (YOLO), Region-based Convolutional Neural Networks (RCNN), Faster RCNN (F-RCNN). RCNN has better accuracy compared to other algorithms but YOLO surpasses when speed is considered over accuracy. In YOLO, Object detection is implemented as a regression problem and class probabilities are provided for detected images. In this paper system is developed using two different algorithms i.e. Yolo and Yolo v3 and tested under same criteria to measure the accuracy and performance. In Yolo Tensor flow SSD Mobile Net model and in Yolo v3 Dark net model is used in [7] To get the audio Feedback gTTS (Google Text to Speech), python library used to convert statements into audio speech. To play the audio pygame python module is used. Testing of both the algorithms is done on MS-COCO Dataset consist of more than 200 K images. Both the algorithms are analysed using webcam in various situations to measure accuracy of the algorithm in every possibility.

Blind Stick reduces the human effort and gives better know-how of the surrounding. Furthermore, it also gives an opportunity for visually impaired people to transport from one area to any other without being assisted by using others. [8] The device also can be used in old age homes where vintage age people have difficulty in their day after day activities due to reduced vision. With this paper, the intention to useful resource human beings in wants to "see" the surroundings. Since the sector of artificial intelligence is doing awesome progress now and functions like object detection is getting less difficult and computationally feasible, these features are implemented in the paper. The paper makes a specialty of object detection and type on pictures that are captured by the device mounted on a stick whose statistics can then be relayed to the person in approach of sound or speech.

AI based blind stick is an innovative stick designed for visually disabled people in order to provide them improved navigation and helping them in making smart decisions about the selection of path that has no obstacle till a certain distance. [2] Our search space involves searching a best suited path for a blind friend by using three ultrasonic sensors from front, left and right that will search the best path which does not have an obstacle at a certain distance. These sensors sense the obstacles through ultrasonic waves and direct blind friend to the direction that is clear of any obstruction to a certain distance. The knowledge is acquired through three sensors that senses the distance of the obstacle. This sensor feedback is compiled and through audio facility communicated to a blind person which is then used for decision-making in selecting the path having no obstruction.

A smart stick for the blind, equipped with obstacle recognition using AI Technologies adds more virtual visibility in their journey. [4] It shows that such a stick can be a significant boon to the blind.

III. PROBLEM STATEMENT

In line with world fitness organization (WHO), there are over 1.3 billion folks who are visually impaired throughout the globe, out of which more than 36 million humans are blind. India being the second biggest populace in the world, contributes 30 percent of the overall blind populace. Even though there are sufficient campaigns being performed to deal with those human beings, it has been tough to supply all the requirements. it is far from the era of synthetic intelligence, and it has received immense traction because of large number of statistics and simplicity of computation. the use of synthetic intelligence, it is miles feasible to make those humans' life much simpler. The purpose is to provide a "secondary sight" until they have got sufficient assets required to deal with them. human beings with untreatable blindness can use this to make their normal duties tons clean and easy.

IV. PROPOSED SYSTEM DESIGN

System design is the technique of defining the elements of the system consisting of the architecture, module and components and their working and the way the records goes via the machine. The gadget may be categorized into enter unit, manage unit and the output unit. These gadgets are aligned on the blind stick for the precise item and the brink detection. The stick is embedded with sensors, raspberry pi, Bluetooth speaker, and battery bank. If the blind or the visually impaired humans is on foot together with his/her stick into his/her hand they shall recognise the limitations coming in front of them thru camera so one can do picture processing of barriers coming in their route. The sensors like Raspberry pi version four are located over the blind stick which makes its strong. If the stick of the visually impaired individual detects an obstacle within the radius it shall give a beeping sound. A silicon ribbon is hooked up to the pinnacle a part of the stick so that it will not slip. If the blind man or woman is shifting in a particular route and in a sure course, the boundaries coming during the blind person are being acknowledged via the digicam and the audio message might be given to the blind man or woman. The stick is made artificially wise by way of photo processing method and feature extracting method. We have used Google Collab for this venture; the digital camera is initialized by using the use of OpenCV library and the digital camera starts off evolved taking pictures frames. Then the gadget makes use of YOLO v4 that is educated at the COCO dataset and darkish Neural network (DNN) to discover the item kept before the consumer. The object recognized is later transformed to an audio phase the usage of gTTs that is a python library. The audio phase is the output of our system that gives the spatial location and name of the object to the character. Now with the aid of the use of this data the individual could have a visualization of the gadgets round him. The proposed machine will even guard the individual from colliding to the objects

ISSN No:-2456-2165

round will cosy him from injuries deliver the space among got here.

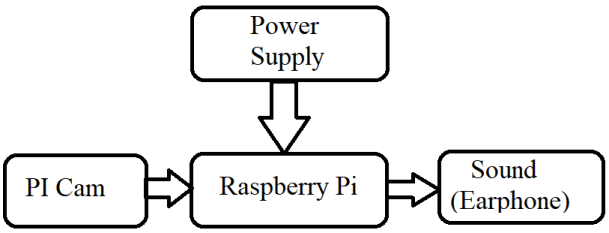

Fig 1: Block diagram of hardware

➢ *Working*

In our model, the visually impaired is helped with the use of our blind stick. It starts with power supplu in this model we are using a power bank which is 5000 maah. Which is better in terms of power and it is also reasonable in terms of the pricing point of the blind stick. We can also increase the battery backup if needed.

The model uses raspberry pi as a backbone of our blind stick, all the data processing is done in it. It makes use of the pi cam, which helps in capturing a picture of the object in front of the blind stick. Which is later processed in the raspberry pi with help of yolo v4 (YOU ONLY LOOK ONCE). The yolo v4 we are using in our model is better and faster than any of the previous models of blind stick. We make use of google collab in the model which gives a constant processing of the image. All the data from the darknet is processed in google collab. It is used in training the dataset we have used in the model.

The object in the front can recognised faster and more accurately with the use of the pi cam. The camera module provides a better resolution which helps in recognising the object faster. The pi cam plays a very important role in the project.

After all the processing and recognising of the object done in the model it is later given out in the form of audio output. Audio output which is given with a Bluetooth headphone, which is very helpful as it wireless and the person can experience less hassle. The visually impaired person gets the audio output in the form of a car or a bike according to the dataset we have trained in our model.

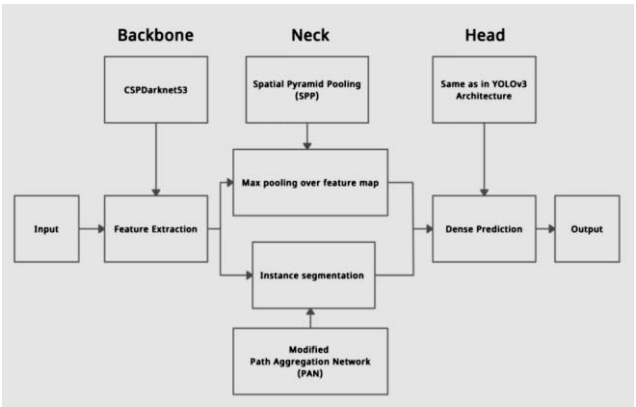

Fig 2: Block Diagram of YOLO V4

Architecture is made up of many parts, and guess what- the first input we have as the training process to feed into the network - they are assembled on the GPU. Next comes Bone and Neck for feature extraction and collection. The sensing neck and sensing head will be sent to the detector together. Finally, the head performs the search/guess. The head is responsible for the search (localization and distribution).

Backbone CSPDarkNet53 is based on the DenseNet design. It combines previous ideas with existing ideas before going into a thick layer - this is called the dense link type.

CSPDarkNet53 has two blocks: *Convolutional Base Layer Cross Stage Partial (CSP) Block* The Cross Stage Partial concept divides the image in the base layer into two parts and combines them with the help of Cross-level hierarchy; multiple gradients to flow between layers alleviate the notorious ”disappearing gradient” problem. The convolutional base layer contains full-size input feature maps.

CSP Block Stacked The two ideas in the instructions next to it will follow directly to the next step without making it from two halves. CSP stores features for better distribution, encourages the network to reuse features, and reduces the number of incompatibilities. The only end-convolutional block in the spinal cord network that can extract richer semantic features is a dense block, because a greater number of densely connected convolutional layers will increase a reduction in speed-seeking.

Neck is the part where assembly is made. It gathers specific information from different levels of the spinal cord, then mixes and combines it to prepare for the next step. In general, the neck has many bottom-up methods and many top-down methods. SPP - Addition Block adds an additional block called SPP (Spatial Pyramid Pooling) between the CSPDarkNet53 backbone and the Feature Collector Network (PANet), this is done to increase the reception area and isolate the most important features and is almost zero. affects network speed. It depends on the last layer of CSPDarkNet’s network connection.

An acceptable area is the area of the image that is subject to a core or filter in a sample. It increases linearly as more convolutional layers are stacked, while it increases exponentially when we stack extended convolutions and introduce a nonlinear state.

Head The main task here is to find and split the connected box. Mark and score the box bounder coordinates (x, y, height and width). Here the x and y coordinates are relative to the centre of the b-box represented by the boundary grid. Width and height are approximate to the entire image.

ISSN No:-2456-2165

V. COMPONENTS USED

- *Yolo V4*

YoloV4 is an important improvement of YoloV4, the implementation of a new architecture in the Backbone and the modifications in the Neck have improved the map (mean Average Precision) by 10 percent and the number of FPS (Frame per Second) by 12 percent. In addition, it has become easier to train this neural network on a single GPU. For getting higher values for precision, YOLOv4 uses a more complex and deeper network via Dense Block.

- *Google Collab*

Collab, or 'Collaboratory', allows you to write and execute Python in your browser, with Zero configuration required, Access to GPUs free of charge, Easy sharing Whether you are a student, a data scientist or an AI researcher, Collab can make your work easier.

- *gTTS (Google Text to Speech):*

A Python library and CLI tool to interface with Google Translates text-to-speech API. Writes spoken mp3 data to a (stout). It features flexible pre-processing and tokenizing.

- *LABELME:*

LabelMe is a python-primarily based open-source photograph polygonal annotation device that may be used for manually annotating snap shots for item detection, segmentation and type. The device is a lightweight graphical application with an intuitive user interface. With LabelMe you can create: polygons, rectangles, circles, traces, factors, or line strips.

VI. RESULT

As a result, system detected automobile vehicles and at the same time provides audio output through the Bluetooth headphones. Number of objects are inputed through camera module. System has correctly identified object surrounding environment. Some of the results are shown in following figures 3.

Fig 3: Car is identified which is captured by Pi camera.

Fig 4: Car is identified which is captured by Pi camera

VII. CONCLUSION

Proposed system is implemented using Raspberry-Pi modules where camera and speaker are interface with it. Yolo-V4 is used in proposed system to identify object in the surrounding environment. After identifying object, system is producing audio of name of the object. This system can be implemented globally to provide blind human beings ease and privateness in daily existence. so as to seriously aid manufacturing and industrial boom in harsh conditions, it is also expected to use for industrial places in which reduced visibility occurs, together with coal mines and sea bottoms. The aim of the observe is to improve the independence of persons with visual impairment, via effectively making use of the proposed system and its associated audio feedback, human with visual impairment may be capable to conquer diverse risks. The camera of the tool may be used to come across item from the surroundings and give output in audio format. Accordingly, assisting visually impaired humans to 'See via the Ears'.

REFERENCES

[1]. M. P. Agrawal and A. R. Gupta, "Smart Stick for the Blind and Visually Impaired People," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 2018, pp. 542-545, doi: 10.1109/ICICCT.2018.8473344.

[2]. Ali, Uruba, Hoorain Javed, Rekham Khan, Fouzia Jabeen, and Noreen Akbar, "Intelligent stick for blind friends," International Robotics and Automation Journal 4, no. 1 (2018).

[3]. N. Dey, A. Paul, P. Ghosh, C. Mukherjee, R. De and S. Dey, "Ultrasonic Sensor Based Smart Blind Stick," 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, India, 2018, pp. 1-4, doi: 10.1109/ICCTCT.2018.8551067.

[4]. Balu N Ilag and Yogesh Athave, "A design review of smart stick for the blind equipped with obstacle detection and identification using artificial intelligence," International Journal of Computer Applications, 182:55– 60, 04 2019.

[5]. N. M. Krishna, R. Y. Reddy, M. S. C. Reddy, K. P. Madhav and G. Sudham, "Object Detection and Tracking Using Yolo," 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2021, pp. 1-7, doi: 10.1109/ICIRCA51532.2021.9544598.
[6]. N. Loganathan, K. Lakshmi, N. Chandrasekaran, S. R. Cibisakaravarthi, R. H. Priyanga and K. H. Varthini, "Smart Stick for Blind People," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 65-67, doi: 10.1109/ICACCS48705.2020.9074374.
[7]. M. Mahendru and S. K. Dubey, "Real Time Object Detection with Audio Feedback using Yolo vs. Yolo_v3," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 734-740, doi: 10.1109/Confluence51648.2021.9377064.
[8]. P Ramachandra Mavarkar and Zarinabegam K Mundargi, "Real time smart blind stick using artificial intelligence," Indian J. Artif. Intell. Neural Netw, 1:9–13, 2021.
[9]. S. Mohapatra, S. Rout, V. Tripathi, T. Saxena and Y. Karuna, "Smart Walking Stick for Blind Integrated with SOS Navigation System," 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2018, pp. 441-447, doi: 10.1109/ICOEI.2018.8553935.
[10]. Nada, Ayat, Samia Mashelly, Mahmoud A. Fakhr, and Ahmed F. Seddik. "Effective fast response smart stick for blind people." In Proceedings of the second nternational Conference on Advances in bio-informatics and environmental engineering–ICABEE. 2015.