IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 9, September 2022

DOI: 10.17148/IJIREEICE.2022.10901

Visualizer: An aid for Visually Challenged

Jayashree K[1], Divyesh Divakar[2], Anantha Krishna Kamath[3]

Assistant Professor, Department of Electronics and Communication Engineering,

Canara Engineering College, Mangaluru, India[1,2,3]

Abstract: One of the most vital senses in the human body is vision, which is also the sense that allows people to perceive their surroundings the best. As a result, thousands of articles on these topics have been published, proposing a range of computer vision services and products by creating new electronic assistive devices for the blind. This work tries to create a system that recovers the identification of nearby objects, a crucial function of the visual system. People who are visually challenged heavily rely on their other senses, such as touch and audio stimuli, to interpret their surroundings. It is quite challenging for a blind person to distinguish what is in front of them without touching it with their hands or any item like walking stick. Physical contact between a person and an object occasionally has the possibility to be fatal. This work employs a Neural Network for the recognition of pre-trained objects. A computer system that has the object identification Neural Network installed to carry out real-time object detection receives input from a camera that is oriented in accordance with the system's predetermined orientation.

Keywords: Computer Vision, Machine Learning, Neural Network, Visually Challenged

I. INTRODUCTION

Millions of people live in this world with incapacities of understanding the environment due to visual impairment. Although they can develop alternative approaches to deal with daily routines, they suffer from certain navigation difficulties as well as social awkwardness. For example, it is very difficult for them to find a particular room in an unfamiliar environment. Visually Challenged peoples find it difficult to know whether a person is talking to them or someone else during a conversation. Computer vision technologies, especially the deep convolutional neural network, have been rapidly developed in recent years. It is promising to use state-of-art computer vision techniques to help people with vision loss. This method aims at exploring the possibility of using the hearing sense to understand visual objects. The sense of sight and hearing sense share a striking similarity: both visual objects and audio sound can be spatially localized. It is not often realized by many people that we are capable of identifying the spatial location of a sound source just by hearing it with two ears [1]. Hence, advanced techniques must evolve to help them.

Continuous effort has been made by science eternity to develop various aids for visually impaired. There is a proposed system in which two cameras are put on the glasses of a blind person. The proposed work has a wearable device consisting of a blind stick and sensor-based detection circuit. It uses an infrared sensor that uses infrared waves to scan the surroundings of a person. It uses object detection and gives them audio information about it. The system must be trained in object information. Feature extraction is also a part of the process [2]. MATLAB Software can be used for signal processing on collected information about what are the different types of obstacles in front of the user, their size, and their distance from the user. Vibrating senses are one of the best sensible for visually impaired and a vibrating motor connected with an ultrasonic sensor will serve the purpose [3].

For less erroneous and accurate systems, multiple object detection algorithm should be adapted. The system should be designed such that 'N' object detectors are trained for 'N' different objects. When an image is sent to the system, all object detectors do their work. If an object is found by a detector, it will mark its boundary and label the object name. After the process complete for all N detectors, the image is displayed with all the tags. Moving a cursor over an object in the image shows the complete boundary of the object with its label beside. This system is a little slower than other systems because a lot of object detectors are working on a single image. The performance can increase by allowing more than one object detectors to run in parallel [4].

Today, Android smartphones are widely used. Using its camera, objects in surroundings can be identified and give an audio signal to the user which helps in obstacle detection and navigation [5]. An electronic device for obstacle detection and face recognition to assist visually challenged users in social interactions is need of the hour. A device in the form of smart glasses that has ultrasonic sensor, pi camera and raspberry pi installed on it can be a great solution for the society. The ultrasonic sensors are connected to the raspberry pi that receives data signals from the sensors for further data processing and detects the obstacles around the user up to 1000cm [6].

We were inspired to create this work by the demand for object detection among blind people and a larger examination of the cutting-edge technologies becoming accessible in the modern world. Technology is something that exists to make

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 9, September 2022

DOI: 10.17148/IJIREEICE.2022.10901

human tasks easier. Therefore, in this study, we employ technology to address the issues faced by those who are visually impaired. The goal of using technology to assist people in object detection is what drives us as engineers to use the tools at our disposal. So, we propose a system built on the breakthrough of image processing and machine learning. The proposed system captures real-time images, then images are pre-processed, their background and foreground are separated and then the DNN module with a pre-trained YOLO (You Only Look Once) model is applied resulting in feature extraction. The extracted features are matched with known object features to identify the objects. Once the object is successfully recognized, the object name is stated as voice output with the help of text-to-speech conversion.

This work introduces a live object recognition system that serves as a blind aid. Visually impaired people heavily rely on their other senses such as touch and auditory signals for understanding the environment around them. The act of knowing what object is in front of the blind person without touching it (by hands or using some other tool) is very difficult. This work tries to transform the visual world into the audio world by recognizing the system more accurately.

Object Detection System using Machine Learning Technique detects objects efficiently based on the YOLO algorithm and applies the algorithm to image data and video data to detect objects. YOLO algorithm is the most sufficient and helpful algorithm for object detection but most of the inventions failed to do it. Initially, blind sticks, trained dogs, and many other techniques were introduced to take care of the blind person such burden is reduced with the help of the recognition system.

The suggested approach produces audio output that a visually challenged user may easily understand. The created programme has the ability to identify nearby items. It can warn the user of any obstructions in his path, making it easier for him to go from place to place without falling. Additionally, it will address the issue of maintaining a particular tool or walking stick. A Raspberry Pi camera that is attached to the Raspberry Pi board is helpful in capturing the image.

II. SYSTEM DESIGN

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

A. Software Overview

I) YOLO:

It is an extremely fast, real-time, multi-object detection algorithm, and it satisfies the basic requirement of our system. YOLO applies a single convolutional neural network to an entire image and divides the image into an S x S grid and comes up with bounding boxes, which are drawn around images and predicts probabilities for each of these regions for object recognition, object localization, and object detection [7].

YOLO predicts multiple bounding boxes per grid cell. For this, we select the highest IoU (intersection over union) with the ground truth. This strategy leads to specialization with the bounding box predictions. Each prediction gets better at predicting certain sizes and aspect ratios. YOLO uses a sum squared error between the predictions and the ground truth to calculate the loss. The loss function comprises: A. Localization loss: measures errors between the predicted boundary box and the ground truth. B. Classification loss: is the squared error of the class conditional probabilities for each class, when an object is detected. C. Confidence loss: detects the objects present in the box. The final loss adds localization, confidence, and classification losses [7].

A person, who is born blind, will not be able to identify structure of an object. This makes our work challenging.

II) OPEN CV:

OpenCV (Open-source computer vision) is a library of programming functions mainly aimed at real-time computer vision. The library has more than 2500 optimized algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, etc. OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To identify image pattern and its various features we use vector space and perform mathematical operations on these features *[8]*.

III) Tensor Flow:

A complete open-source framework for building machine learning applications is called TensorFlow. It is a symbolic math toolkit that carries out several operations targeted at deep neural network training and inference using dataflow and differentiable programming. It enables programmers to build machine learning applications utilizing a range of

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 9, September 2022

DOI: 10.17148/IJIREEICE.2022.10901

instruments, frameworks, and community assets. But the models that TensorFlow produces can be installed on almost any device and used to make predictions.

IV) Dark Net:

Darknet has a different architecture and set of capabilities than other deep learning frameworks and is mostly used for object detection. It is quicker than other NN designs and methodologies, like Faster RCNN, etc. They are at the top of their game in terms of speed and accuracy thanks to their unique framework, the darknet architecture and YOLO. Although this deep learning system is developed in C, once the network has been trained, Darknet is not required for inference. Darknet formats are already supported by OpenCV, therefore both models and training weights can be used wherever OpenCV is installed.

B. **Hardware Overview**

The Raspberry Pi serves as the brain of our operation. Since the output would be in audio form, we chose to use a speaker. Raspberry Pi also supports high-bass headphones. The Raspberry Pi (3 B+) design was employed. We made the decision to employ a power bank as the Raspberry Pi's power source in order to give consumers mobility. With OpenCV on a Raspberry Pi, all the main image processing algorithms and operations may be quickly implemented. The board supports data and related information up to 32 GB with the help a class 10 SD card.

Fig 1 Raspberry Pi Camera

The Raspberry Pi Camera (representative Fig 1) v2 is a high quality 8-megapixel Sony IMX219 image sensor custom designed add-on board for Raspberry Pi, featuring a fixed focus lens. It is capable of 3280 x 2464-pixel static images, and also supports 1080p30, 720p60 and 640x480p60/90 video. It is a custom designed add on module for raspberry-pi hardware it attaches to raspberry-pi hardware through custom CSI interface. The CSI bus is capable of extremely high data rates, and it exclusively carries pixel data to the processor.

III. IMPLEMENTATION

The proposed system provides output in audio form which is easily understandable for a visually impaired user. The system block diagram is shown in Fig 2. The application that was developed can recognize surrounding items. It can warn the user of any obstacles in his way, facilitating safe movement from one area to another. It will also address the issue of using a particular tool or walking stick. A Raspberry Pi camera that is attached to the Raspberry Pi board is helpful in capturing the image. For example, if the obstacle is present while the person is moving, the system suddenly alerts the person by conveying that obstacle is detected and it is present atacertain distance. This Operation is performed by the latest machine learning algorithms. For Image Processing purposes, we used the OpenCV Library, TensorFlow algorithm and YOLO algorithm. If the input image captured in a pi camera is a single image, then the system determines the objects and tells the person what actually the output is with the help of audio output.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 9, September 2022

DOI: 10.17148/IJIREEICE.2022.10901

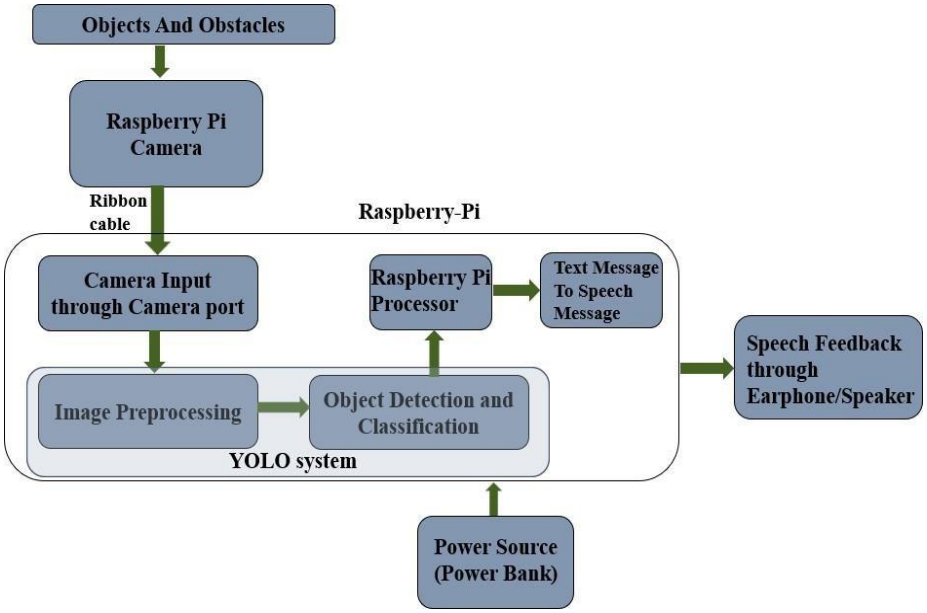

Fig 2 System Block Diagram

A. **Image Processing Stages**

Initially the object is captured with the help of Raspberry Pi Camera which is connected to the Raspberry Pi board. The captured or real time Video is segmented into certain kind of boundaries with the help of latest machine learning technology. The image processing techniques begins by partitioning individual images by identifying it with the help of its name which is already involved in the trained data set. The various stages involved in the process are represented in Fig 3.

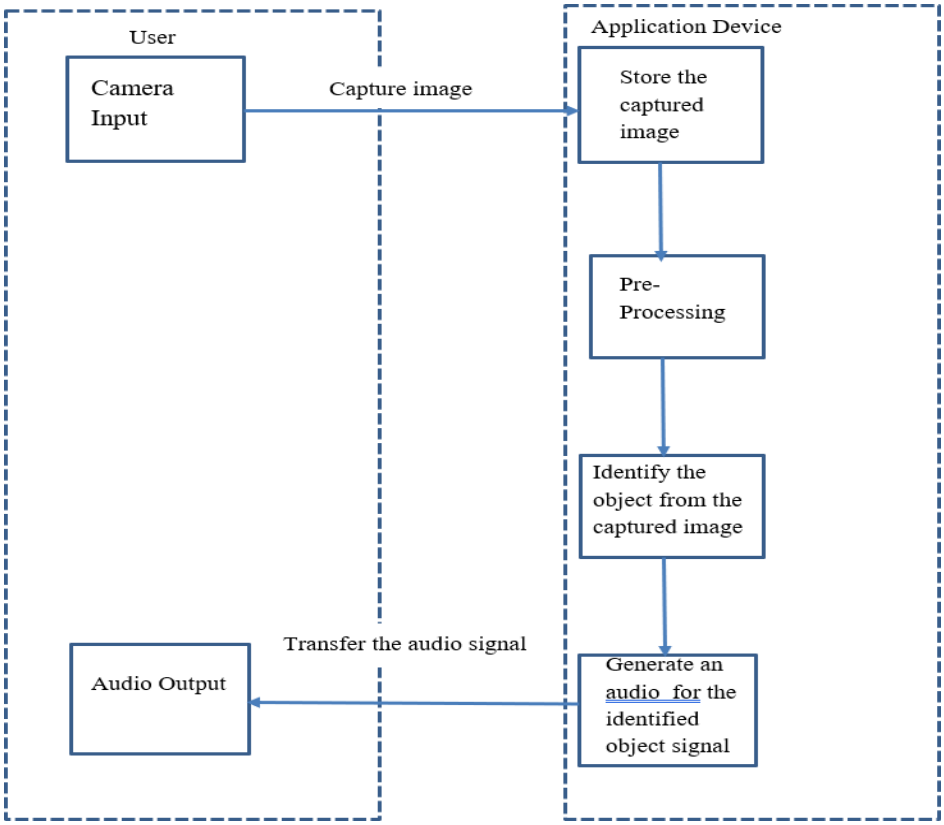

Fig 3 Image Processing Stages

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 9, September 2022

DOI: 10.17148/IJIREEICE.2022.10901

IV. RESULTS AND DISCUSSION

The results obtained were satisfactory considering the fact that the image is captured in the real time, and how object is detected. In the Fig 4, the captured objects inside the view will be identified and have a rectangle boundary drawn around them. This shows the single object laptop with accuracy of 87% and output is determined by text and audio signal.

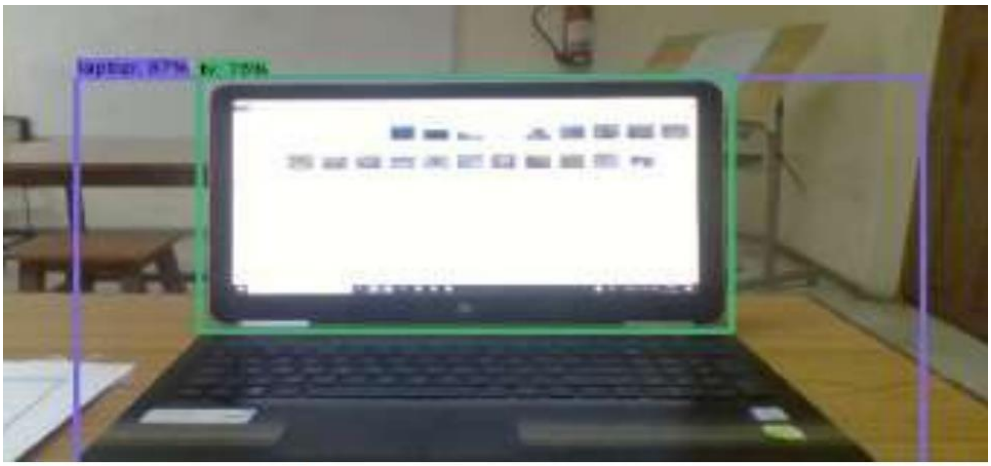

Fig 5 clearly shows the multiple entities in a single image contains book with 50% accuracy, plant pot with 57% and bottle with 56%. These objects are identified based on the location of the object with the help of audio output.
This system uses Artificial intelligence technique for object identification in order to detect objects in different directions. Object detection deals with detecting live images in the real time scenario. The Tensor Flow pre-trained models can easily determine the captured object in an image, because they have very less information aboutthe objects in the surrounding and direction which is essential for travel.

Fig 5 Object Detection (Multiple Entities)

V. CONCLUSION

In recent years, many solutions have been proposed to help visually impaired in recognizing objects but they were not efficient. The purpose of this work was to help blind people in independently identifying objects without any external support. This system named as visualizer, will help the visually impaired by providing virtually visible audio descriptions of their surroundings and help them travel with self-confidence. This system accurately detects the obstacles in the left, right and front side, due to which it is user friendly. This will help to reduce burden of usage of human guides or trained dogs for continuously monitoring the blind.
This work can be improved by increasing the speed of detection of objects in real time by using higher version of Raspberry Pi with higher RAM capacity. There is also a need for calculating the distance of the object from the user.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 9, September 2022

DOI: 10.17148/IJIREEICE.2022.10901

ACKNOWLEDGMENT

This paper and the research behind it would not have been possible without the exceptional support of **Dr. Dayananda G K,** Head of the Department, Electronics and Communication Engineering Department, Canara Engineering College. His enthusiasm and exacting attention to detail have been an inspiration and kept our work on track. We would also like to show our gratitude to **Dr. Ganesh V Bhat,** Principal, Canara Engineering College for sharing his pearls of wisdom with us during the course of this research.
We are also immensely grateful to **Mr. Sathwik Shetty, Ms. Shreya Shetty, Mr. Uddesh Hegde and Mr. Vijeth Kumar K** for supporting us during the entire course of this work.

REFERENCES

[1] S. N , N. Priya M and P. U , "Real Time Object Detection for Blind People," International Journal of Advance Research in Science and Technology, vol. 7, no. 1, 2018.
[2] S. Udgirkar, S. Sarokar, S. Gore, D. Kakuste and S. Chaskar, "Object Detection System for Blind People," International Journal of Innovative Research in Computer and Communication Engineering, 2006.
[3] D. Rajput, F. Ahmed, H. Ahmed, E. Z. A. Shaikh and A. Shamshad, "Smart Obstacle Detector for Blind Person," Journal of Biomedical Engineering and Medical Imaging, 2014.
[4] K. Khurana and R. Awasthi, "Techniques for Object Recognition in Images and Multi-Object Detection," International Journal of Advanced Research in Computer Engineering & Technology, 2013.
[5] E. Peng, P. Peursum, L. Li and S. Venkatesh, "A Smartphone-Based Obstacle Sensor for the Visually Impaired," in International Conference on Ubiquitous Intelligence and Computing, 2010.
[6] R. Rajwani, D. Purswani, P. Kalinani, D. Ramchandani and I. Dokare, "Department of ComputerProposed System on Object Detection for Visually Impaired People," International Journal of Information Technology, vol. 4, no. 1, 2018.
[7] S. Shaikh, V. Karale and G. Tawde, "Assistive Object Recognition System for Visually Impaired," International Journal of Engineering Research & Technology (IJERT), vol. 9, no. 9, pp. 736-740, 2020.
[8] "OpenCV," 2022. [Online]. Available: https://opencv.org/about/. [Accessed 12 August 2022].