ML Based Blind Assistance System

Gudla Vaishnavi[1], Varkala Prakash[2], Sreebhashyam Pranatha[3], Thimmanagari Sampurna[4], B Rekha[5]
[Font: Times New Roman, Size:11]
[1,2,3,4] *Student, Department of Electronics and Communication Engineering, Teegala Krishna Reddy Engineering College*
[5]*Assistant Professor, Department of Electronics and Communication Engineering, Teegala Krishna Reddy Engineering College*

Abstract— Eye diseases usually cause blindness and visual impairment. As per the statistics, there are over 285 million visually impaired people living worldwide. They come across many troubles in their daily life, especially while navigating from one place to another on their own. They often depend on others for help to satisfy their day-to-day needs. So, it is quite a challenging task to implement a technological solution to assist them. Several technologies were developed for the assistance of visually impaired people. One such attempt is that we would wish to make an Integrated Machine Learning System that allows the blind victims to identify and classify real-time objects generating voice feedback and distance. Which also produces warnings whether they are very close or far away from the thing.

***Index Terms*—Object tracking, OpenCV, computer vision, Webcam, NumPy**

I.INTRODUCTION

Over the past decade, machine learning algorithms and applications have contributed to new advances in the field of assistive technology. Researchers are leveraging such advancements to continuously improve human quality of life, especially those with disabilities or alarming health conditions. Assistive technology (AT) deploy devices, present services or programs to improve functional capabilities of people with disabilities. The scope of assistive technology research studies comprises hearing impairment, visual impairment, and cognitive impairment, among others . Vision impairment can vary from mild, moderate, severe vision impairment and total blindness. In the light of the recent advances in machine learning and deep learning, research studies and new solutions for people with visual impairment have gained more popularity. The main goal is to provide people with visual impairment with visual substitution by creating navigation or orientation solutions. Such solutions can ensure self-independence, confidence, and safety for people with visual impairment in the daily tasks. According to estimates, approximately 253 million individuals suffer from visual impairments: 217 million have low-to-high vision impairments, and 36 million are blind. Figures have also shown that, amongst this population, 4.8% are born with visual deficiencies, such as blindness: for 90% of these individuals, their ailments have different causes, including accidents, diabetes, glaucoma, and macular degeneration. The world's population is not only growing, but also getting older, meaning more people will lose their sight due to chronic diseases. Such impediments can have knockon effects; for example, individuals with visual impairments who want an education may need specialized help in the form of a helper or equipment. Learners with visual impairments can now make use of course content in different forms, such as audiotapes, Braille, and magnified material. It is worth noting that these tools read the text instead of images. Technological advancements have been employed in educational environments to assist people with visual impairment, blind people, and special-needs learners, and these developments, particularly concerning machine learning, are ongoing. The main objective of conducting visual impairment research studies is to achieve visual enhancement, vision replacement, or vision substitution as originally classified by Welsh Richard in 1981. Vision enhancement involve acquiring signals from camera which processed to produce an output display through head-mounted device. Vision replacement deals with displaying visual information to the human brain's visual cortex or the optic nerve. Vision substitution concentrate on delivering nonvisual output in a auditory signals . In this paper, we focus on vision substitution solution that delivers a vocal description on both printed texts and images to people with visual impairment. There are three main areas of concentration concerning research on people with visual impairment; namely, mobility, object detection and recognition and navigation. In the era of data explosion and information availability, it is imperative to consider means to information access for people with visual impairment specially printed information and images.Over the past decades, authors have leveraged state of the art machine learning algorithms to develop solutions supporting each of the aforementioned areas. Deep learning has evolved in prominence as a field of study that seeks innovative approaches for automating different tasks depending on input data .Deep learning is a type of artificial intelligence techniques that can be used for image classification, recognition, virtual assistants, healthcare, authentication systems, natural language processing, fraud detection, and other purposes. The study describes an Intelligent Reader system that employs Deep Learning techniques to help people with visual impairment read and describe images in a printed text book. In the proposed technique, Convolutional Neural Network (CNN) is utilized to extract features from input images, while Long Short-Term Memory (LSTM) is used to describe visual information in an image. The intelligent learning system generates a voice message comprising text and graphic information from a printed text book using the text-to-speech approach. Deep learning-based technologies increase image-related task performance and can help people with visual impairment live.

II. EXPERIMENTAL METHODOLOGY

A. Methodology

A Single Shot Detector (SSD) algorithm

SSD is a popular object detection algorithm that was developed in Google Inc. It is based on the VGG-16 architecture. Hence SSD is simple and easier to implement.

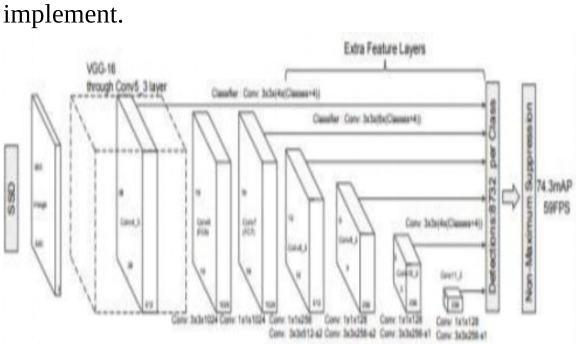


Figure:1 VGG 16 SSD model

A set of default boxes is made to pass over several feature maps in a convolutional manner. If an object detected is one among the object classifiers during prediction, then a score is generated. The object shape is adjusted to match the localization box. For each box, shape offsets and confidence level are predicted. During training, default boxes are matched to the ground truth boxes. The fully connected layers are discarded by SSD architecture. The model loss is computed as a weighted sum of confidence loss and localization loss. Measure of the deviation of the predicted box from the ground truth box is localization loss.

Confidence is a measure of in which manner of confidence the system is that a predicted object is the actual object. Elimination of feature re sampling and encapsulation of all computation in a single network by SSD makes it simple to train with Mobile Nets. Compared to YOLO, SSD is faster and a method it performs explicit region proposals and pooling (including Faster R-CNN).

B. Analysis

Two ways in which the object can be tracked in the above example are: (1)-Tracking in a sequence of detection. In this method a CCTV video sequence of a traffic which is in motion takes place. Suppose someone wants to track a car or person’s movement here, he will take different images or frames at different intervals of time. With the help of these images one can target the object like a car or person. Then, by checking how my object has moved in different frames of the video, I can track it. Velocity of the object can be calculated by verifying the object’s displacement with the help of different frames taken at different intervals of time. This method is actually a flaw where one is not tracking but detecting the object at different intervals of time. Improved method is “detection with dynamics”. In this method estimation of a car's trajectory or movement takes place. By checking its position at a particular time ‘t’ and estimating its position at another time interval let’s say ‘t+10’. From this actual image of the car at ‘t+10’ time can be proposed with the help of estimation.

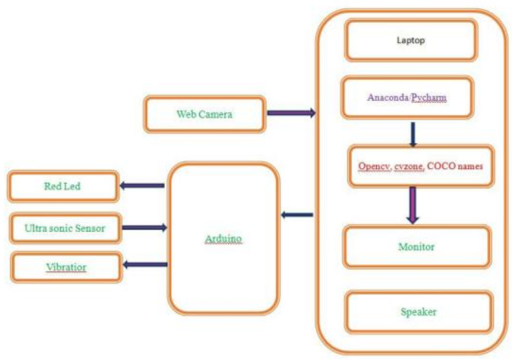
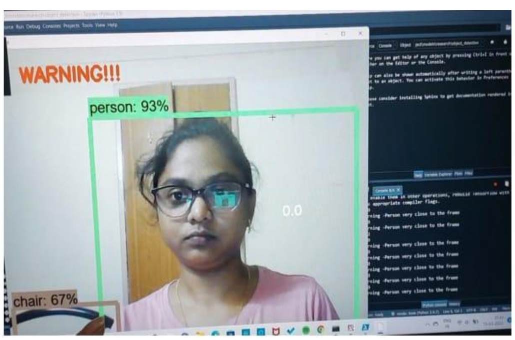
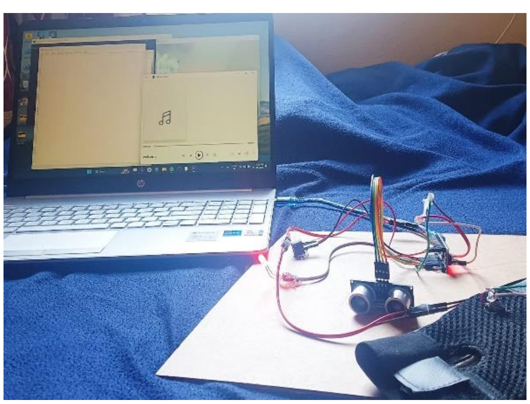


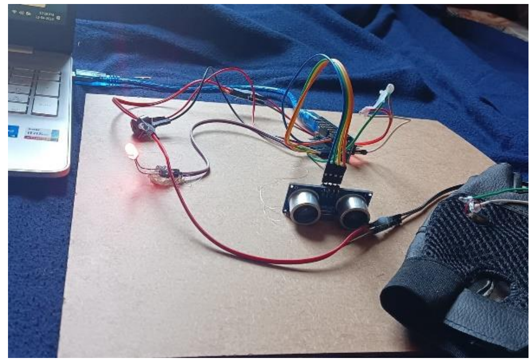
Figure:2 Block Diagram.

IV. RESULTS AND DISCUSSION

Object detection is the emphasis of the proposed system. Wearable and portable technology has been developed. The person's chest is hooked to the system. The Raspberry Pi camera captures video of the scene, which is subsequently translated into frames by the processor. The auditory output from the system directs the user to the object. Figure 6 shows the detection of a person and a thing (blue cell phone).







V. CONCLUSION

The project helps people who are visually impaired to get a better understanding of the surroundings. It speaks to the user what is present in front of them, hence helping them avoid obstacles and giving them the ability to "see" what's around them. The paper aims to make the world a better living environment for people who are handicapped or have a tough time seeing. From the results seen above, it is clear that the user can greatly benefit in terms of knowing what is around them at a cheaper, efficient and easy way. There are wide varieties of objects that the device can detect. Hence it can be used for everyday activities to enhance their experience and create a better place for them. In the future, face detection can be added, so that if there are any familiar faces, they can be recognised. Also text conversions would help the users read books or posters and signs to enhance their understanding even more. With increasing research in embedded systems, more computation in a smaller scale can be expected in the future. This can greatly impact the implementation of AI features like object detection, face recognition etc.

REFERENCES

[1] Muthu krishnan.R and M.Radha "Edge Detection Techniques for image Segmentation" International Journal of Computer Science & Information Technology (IJCSIT) Vol3, No 6, Dec 2011.

[2] Geng Xing, Chen ken, Hu Xiaoguang "An improved Canny edge detection algorithm for color image" IEEE TRANSATION, 2012 978-1- 4673-0311-8/12/$31.00 ©2012 IEEE.

[3] Chen.W, Yue.H, Wang.J, Wu.X. An improved edge detection algorithm for depth map in painting. Optics and Lasers in Engineering 2014;55: 69–77.

[4] Moeslund T. Canny Edge detection. Denmark: Laboratory of Computer Vision and Media Technology, Aalborg University. ⟨http://www.cvmt.dk/ education/teaching/f09/VGIS8/AIP/canny_09gr820. pdf⟩ [March 2009].
[5] OpenCV, 2016. http://opencv.org/. [Accessed December 23, 2016].
[6]AnEmbeddedRealTimeObjectDetectionandMeasurementofitsSize.pdf
[7] Mengel, P., & Wertheimer, R. (2007). D52.300.1 camera requirements and specifications. User Cams project deliverable.