ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of Management** of **SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. T.Sasikala M.E., Ph.D.**, **Dean**, School of Computing , **Dr.S.Vigneshwari M.E., Ph.D., and Dr.L.Lakshmanan M.E., Ph.D.,** Heads of the Department of Computer Science and Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr. A. Jesudoss M.E., Ph.D.,** for his valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **Department of Computer Science and Engineering** who were helpful in many ways for the completion of the project.

TRAINING CERTIFICATE

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited with Grade 'A' By NAAC | 12B Status by UGC |Approved By AICTE
www.sathyabama.ac.in

School of Computing

Department of Information Technology

Certificate of Presentation

This is to certify that

YERRAGOPU VENKATA SAI

of

Sathyabama Institute of Science and Technology

has presented a paper entitled

Blind Assistance System Using TensorFlow

in the AICTE sponsored
International Conference on Artificial Intelligence and Machine Learning (IAIM-2022),
held from 27th January 2022 - 29th January 2022.

Dr. R. SUBHASHINI
HOD [Information Technology]

Dr.T. SASIKALA
Dean [School of Computing]

Dr. T. SASIPRABA
Vice Chancellor

ABSTRACT

There is a ton of classified and earnest data trade here today which is converted into. a desired form with the convolutional neural network to make the blind feel things surrounding him. Efficient Implementation of this Model depends upon the compatibility with python and library installation hurdles. The proposed Blind Assistance ideas were developed to help those visually impaired people around the world. The proposed system has two levels. It works on the SSD algorithm and TensorFlow, in which it recognizes the object, not only the recognition but also localization of an object. It also provides the distance between the person and the object. The Laptop Based Server will be using a pre-trained SSD detection model trained on COCO DATASETS. It will then test and the output class will get detected with an accuracy metrics. In this manner, we have proposed a framework that utilizes AI, which makes the framework more straightforward to use for individuals with visual impedances. After testing with the help of voice modules the class of the object will be converted into a default voice notes which will then be sent to the blind victims for their assistance. Main key aspect of the proposed system is identifying or naming the object detected, prints accurate distance between the user and the objects and voice over using Audio Commands.

ABBREVATIONS

SSD – Single Shot Detector
CNN – Convolutional Neural Network
CSS– Cascading Style Sheets

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Computer vision drew attention by swaying as a data- reliant stratified feature in extraction methods. It is a dream-based vision for the visually impaired. The framework is intended to permit blind individual admittance to programming, which likewise sends time spans to versatile PCs. Working on the capacity of individuals with visual disabilities to make independent savvy glasses can recognize things that are viable with the climate and to know them progressively. People suffering from such defective things, this project helps them to some extent. We need to acquire some updated technologies to finish this project with perfect output. Many projects just take input and display outcome but we need to make a chance of giving voice commands to the system. With the reference of “Object tracking and counting in a zone using YOLOv4, Deep SORT, and TensorFlow”. this paper helped me in many aspects of few steps in my project. In the above paper, they used technologies like YOLOv4, Deep SORT, and tensor flow. Tensor flow helped me to give a nice output to my project. The above-mentioned project mainly focused on detecting multiple peoples and objects like cars bicycles etc. Counts the objects detected, it also captures a video similar to our project. So, in this project only the detection takes place. In our project, we are going to implement new techniques using CNN, SSD, and tensor flow and the addition of voice also takes place. In this project, the user can experience the categorization of images. Detection of images, Localization of images, Voice given by the system single shot detector we come across many phases to detect the object like Anchor box, zoom level, depth level. Anchor box is nothing but a grid like box that surrounds the detected object. Zoom level helps the anchor box in size shape and few more aspects. Depth estimation is used to detect the distance between the two objects accurately

1.1 PROBLEM STATEMENT

It is observed that the performance of the object detection degrades drastically under the course of action. The results are also compared with existing blind assistance approaches to evaluate the accuracy drop. In use assistance systems for blind only tells the object name in front of the lens and nothing more. Main concern with the existential models is they don’t tell the distance between the object and the person, objectifying the problem we have made it possible with neural networks to calculate the distance and will warn the user.

We propose a primary state-of-the-art performance using the solution in terms of using COCO datasets which will identify wide range of objects and residues the distance between object and the end user.

1.2 PROJECT JUSTIFICATION

The necessity of this protrude is extracting the distance between identified object and the users to residue the complexity of calculating distance. Integrated Machine Learning System and convolutional networks which allows the Blind Victims to identify and classify Real Time Based Common day-to-day Objects with the help of TensorFlow APIs and generate voice feedbacks using python modules. It also calculates distance which produces warnings whether the user is very close or far away from the object. The same system can be used for Obstacle Detection Mechanism.

CHAPTER 2

LITERATURE SURVEY

2.1 Object Detection System based on SSD Algorithm

Abstract:

This SSD (Single Shot Multi Box Detector) is an object detection algorithm based on deep learning. As one of the most mainstream detection algorithms, it can greatly improve the detection speed and ensure the detection accuracy. In this paper, the Batch Norm operation is added to the network in order to improve the generalization of the network and speed up network training. The object counting function is added to the image recognition. This paper uses SSD algorithm that incorporates Batch Norm algorithm. The object detection system was built by the Flask framework and the Layui
framework. The system can select the data to be detected on the front-end page, the detection results and the number of each type of object were displayed on the front-end page in real time.

2.2 Object Detection Method Based on Improved One-Stage Detector

Abstract:

We use an efficient way to simulate the two-stage detector, and optimize the algorithm to improve efficiency of object detection as the accuracy of the one-stage detector often lags behind that of the two-stage detector. We have improved the one-stage detector to get higher accuracy on SSD. We mainly adjust the default boxes so that it can well capture the range of the target object in the image, and then the module given the adjusted boxes responds to the alignment of the receptive field of the convolution filter. Test results on the PASCAL VOC dataset show that the final model accuracy can reach that of the advanced detector.

2.3 A Deep Learning based Real Time Assistive Framework for Visually Impaired

Abstract:

Visually impaired people face great difficulties in interacting with unfamiliar surroundings and making use of public amenities. Commuting and using public services without any external help is very challenging to the visually impaired. Over time there have been numerous developments to aid visually impaired but mostly these have been limited the use of sensors to measure distances and send warning signals. The proposed work presents a deep learning technique to aid the visually impaired in real time by identifying nearby public amenities that are used by all in everyday life like Restrooms, ATMs, Metro stations and Pharmacies. This method used faster region-convolutional neural networks (Faster R-CNN) with resnet50 to identify the symbols of public amenities that are the same around the world. This algorithm was trained with a set of over 450 images and was tested on a varying database and achieved an accuracy of 92.13 percent. The experimental results to identify public amenities like
Restrooms, ATMs, Metro stations and Pharmacies are robust, efficient, low in cost and helps in providing sight to the visually impaired up to a large extent.

2.4 Human-in-the-Loop Machine Learning to Increase Video Accessibility for Visually Impaired and Blind Users

Abstract:

Video accessibility is crucial for blind and visually impaired individuals for education, employment, and entertainment purposes. However, professional video descriptions are costly and time-consuming. Volunteer-created video descriptions could be a promising alternative, however, they can vary in quality and can be intimidating for novice describers. We developed a Human-in-the-Loop Machine Learning (HILML)

approach to video description by automating video text generation and scene segmentation and allowing humans to edit he output. The HILML approach facilitates human-machine collaboration to produce high quality video descriptions while keeping a low barrier to entry for volunteer describers. Our HILML system was signifcantly faster and easier to use for first-time video describers compared to a human-only control condition with no machine learning assistance. The quality of the video descriptions and understanding of the topic created by the HILML system compared to the human-only condition were rated as being signifcantly higher by blind and visually impaired users.

2.5 Smartphone-Based Assistance for Blind People to Stand in Lines

Abstract:

We present a system to allow blind people to stand in line in public spaces by using an off-the-shelf smartphone only. The technologies to navigate blind pedestrians in public
spaces are rapidly improving, but tasks which require to understand surrounding people's behavior are still difficult to assist. Standing in line at shops, stations, and other crowded places is one of such tasks. Therefore, we developed a system to detect and notify the distance to a person in front continuously by using a smartphone with a RGB camera and an infrared depth sensor. The system alerts three levels of distance via vibration patterns to allow users to start/stop moving forward to the right position at the right timing. To evaluate the effectiveness of the system, we performed a study with six blind people. We observed that the system enables blind participants to stand in line successfully, while also gaining more confidence.

2.6 Tools and Technologies for Blind and Visually Impaired Navigation Support

Abstract:

The development of navigation tools for people who are visually impaired had become an important concern in the research area of assistive technologies. This paper gives a comprehensive review of different articles published in the area of navigation solutions for people who are visually impaired. Unlike other review papers, this review considers major solutions that work in both the indoor or/and outdoor environments which are based on different technology. From the review, it became clear that the navigation systems proposed for the target users lack some core features that are quite important for independent navigation. Also, there can be instances in which humanitarian conditions also have to be considered in the navigation system design. Based on these findings, a set of recommendations are also given which can be considered in the future design of navigation systems for

blind and visually impaired people.

2.7 Virtual navigation for blind people: Transferring route knowledge to the real-World

Abstract:

Independent navigation is challenging for blind people, particularly in unfamiliar environments. Navigation assistive technologies try to provide additional support by guiding users or increasing their knowledge of the surroundings, but accurate solutions are still not widely available. Based on this limitation and on the fact that spatial knowledge can also be acquired indirectly (prior to navigation), we developed an interactive virtual navigation app where users can learn unfamiliar routes before physically visiting the environment. Our main research goals are to understand the acquisition of route knowledge through smartphone based virtual navigation and how it evolves over time; its ability to support independent, unassisted real-world navigation of short routes; and its ability to improve user performance when using an accurate in-situ navigation tool (NavCog). With these goals in mind, we conducted a user study where 14 blind participants virtually learned routes at home for three consecutive days and then physically navigated them, both unassisted and with NavCog. In virtual navigation, we analyzed the evolution of route knowledge and we found that participants were able to quickly learn shorter routes and gradually increase their knowledge in both short and long routes. In the real-world, we found that users were able to take advantage of this knowledge, acquired completely through virtual navigation, to complete unassisted navigation tasks. When using NavCog, users tend to rely on the navigation system and less on their prior knowledge and therefore virtual navigation did not significantly improve users' performance.

2.8 Learning an Object Detector Using Zoomed Object Regions

Abstract:

The single shot multi-box detector (SSD) is one of the first real-time detectors, which uses a convolutional neural network (CNN) and achieves the state-of-the-art detection performance. However, owing to the semantic gap between each feature layer of CNN, the SSD has a room for improvement. In this paper, we propose a novel training scheme to enhance the performance of the SSD. In object detection, ground truth (GT) box is a bounding box enclosing an object boundary. To improve the semantic level
of the feature map, we generate additional GT boxes by zooming in to and out from the original GT boxes. Experimental results show that the SSD trained with our scheme outperforms the original one on public dataset.

2.9 Real Time Multi Object Detection for Blind Using Single Shot Multibox Detector

Abstract:

According to world health statistics 285 million out of 7.6 billion population suffers visual impairment; hence 4 out of 100 people are blind. Absence of vision restricts the mobility of a person to pronounced extent and hence there is a need to build an explicit device to conquer guiding aid to the prospect. This paper proposes to build a prototype that performs real time object detection using image segmentation and deep neural network. Further the object, its position with respect to the person and accuracy of detection is prompted through speech stimulus to the blind person. The accuracy of detection is also prompted to the device holder. This work uses a combination of single-shot multibox detection framework with mobileNet architecture to build rapid real time multi object detection for a compact, portable and minimal response time device construction.

2.10 A study on object detection method from manga images using CNN

Abstract:

Japanese comics (manga) are popular content worldwide. In order to acquire metadata from manga images, techniques automatic recognition of manga content have been
studied. Recently, Convolutional Neural Network (CNN) has been applied to object detection in manga images. R-CNN and Fast R-CNN generate region proposals by Selective Search. Faster R-CNN generates them using CNN layers called Region Proposal Network (RPN). Single Shot MultiBox Detector (SSD), the latest detection method, performs object classification and box adjustment for small regions in an image. These methods are effective to natural images. However, it is unclear whether such methods work properly to manga images or not, since those image features are different from natural images. In this paper, we examine the effectiveness of manga object detection by comparing Fast R-CNN, Faster R-CNN, and SSD. Here, manga objects are panel layout, speech balloon, character face, and text. Experimental results show that Fast R-CNN is effective for panel layout and speech balloon, whereas Faster R-CNN is effective for character face and text.

2.11 Speed-Up of Object Detection Neural Network with GPU

Abstract:

We realized a speed-up of an object detection neural network with GPU. We