### REAL-TIME OBJECT IDENTIFICATION FOR VISUALLY IMPAIRED PEERS

**BCSF187Z50 - PROJECT WORK PHASE- II**

**REPORT**

##### Batch 2023 (Semester VIII)

SUBMITTED BY

**S DHEERAJ KUMAR** - **11199A063**

**J HIMAVANTH** - **11199A097**

GUIDED BY

DR M SENTHIL KUMARAN

ASSOCIATE PROFESSOR



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**SRI CHANDRASEKHARENDRA SARASWATHI VISWA MAHAVIDYALAYA**

**May-2023**

##### Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya

Enathur, Kanchipuram – 631 561



**BONAFIDE CERTIFICATE**

This is to certify that the PROJECT WORK PHASE-II Report entitled **Real-Time Object** **Identification For Visually Impaired** **Peers** is the bonafide work carried out by Mr. **S Dheeraj Kumar** Reg.No 11199A063 and Mr. **J Himavanth** Reg.No11199A097 during the academic year 2022-2023.

DR.M.SENTHIL KUMARAN DR.M.SENTHIL KUMARAN

Associate Professor Head of the Department,

Department of CSE, Department of CSE,

SCSVMV. SCSVMV.

Submitted for the project work viva-voce examination held on

Place: Kanchipuram.

Date :

***Examiner 1 Examiner2***

**DECLARATION**

It is certified that the PROJECT WORK PHASE-II work titled Real-Time Object Identification For Visually Impaired Peers is originally implemented by us. No ideas, processes, results, or words of others have been presented as our work. Due acknowledgment is given wherever others' work or ideas are utilized.

a. There is no fabrication of data or results which have been compiled /analyzed.

b. There is no falsification by manipulating data or processes or changing or omitting data or results.

We understand that the project is liable to be rejected at any stage (even at a later date) if it is discovered that the project has been plagiarized, or significant code has been copied. We understand that if such malpractices are found, the project will be disqualified, and the Degree awarded itself will become invalid.

Signature of the student with date

|  |  |  |
| --- | --- | --- |
| **CHAPTER NO** | **TITLE** | **PAGE NO** |
|  | **ABSTRACT** | **I** |
|  | **LIST OF FIGURES** | **II** |
| **1** | **Introduction** | **1-3** |
| 1.1 | Objectives | **3** |
| 1.2 | Scope of the Project | **3** |
| **2** | **Literature Survey** | **4-6** |
| **3** | **Problem Statement/Proposed Method / Algorithm /Architecture / Project/ Description** | **7-11** |
| 3.1 | Problem Statement | **8** |
| 3.2 | Proposed method | **8** |
| 3.3 | System Architecture | **9** |
| 3.4 | Algorithm | **10** |
| 3.5 | Methodology | **10** |
| 3.6 | Project Description | **11** |
| **4** | **Implementation Work** | **12-20** |
| 4.1 | Implementation process | **13** |
| 4.2 | Code | **15-16** |
| 4.3 | Result | **17-20** |
| **5** | **Conclusion** | **22** |
|  | **References** | **25** |
|  |  |  |

**TABLE OF CONTENTS**

**ABSTRACT**

Real-Time Object Detection is a challenging aspect, improving day to day in its accuracy. Deep Learning gained a tremendous influence on the world is adapting to Artificial Intelligence. Nowadays blind people can recognize surroundings through sensors and Actuators according to the situation around. Some of the popular algorithms for Object detection are Region-based Convolutional Networks (RCNN), Single Shot Detector(SSD), Faster-RCNN. CNN architecture models have the ability to eliminate highlights and identify objects in any given image. When implemented appropriately, CNN models can address issues like deformity diagnosis, creating educational or instructive applications, etc. The main objective is to provide improved accuracy in identifying and transferring the object detected in real-time through audio as a medium of communication. The parameters like minimum Average precision (mAP), Frames Per Second (FPS) are considered for better accuracy. The YOLO ( You Look Only Once) algorithm is applied for Computation. A Real-Time System is developed to help visually impaired people.

KEYWORDS:

OpenCV, NumPy, Time, Video Capture

LIST OF FIGURES

|  |  |  |
| --- | --- | --- |
| S.no | Title | Pg.no |
| 1 | System Architecture | 09 |
| 2 | Detection of Person | 17 |
| 3 | Detection of Person | 18 |
| 4 | Detection of Motor vehicle | 19 |
| 5 | Detection of Traffic light | 20 |
| 6 | Publication details | 24 |

# CHAPTER 1

## Introduction

**1 Introduction**

Computer Vision in Computer technology explains the break and elucidation of the visual world . In classifying the objects’ accuracy, machines use deep learning models17 and digital images such as cameras and videos. In 1950s, demonstrations have already started in computer vision to identify the edges and align the simpler objects with falling under categories such as circles and squares by the techniques of first neural networks. In 1970s, Optional character recognition came into existence of computer vision, handwritten data on its primary trading tool. The illustrated data mainly used for the blind as a development. Visually impaired people find it hard to recognize the smallest detail with healthy eyes. Those who have the visual acuteness of 6/60 or the horizontal range of the visual field with both eyes open have less than or equal to 20 degrees. These people are regarded as blind. A survey by WHO (World Health Organization) carried out in 2021 estimates that in the world 1billion people includes those with moderate or severe distance vision impairment due to unaddressed refractive error (88.4 million), cataract (94 million), glaucoma (7.7 million), corneal opacities (4.2 million), diabetic retinopathy (3.9 million), and trachoma (2 million), as well as near vision. The main problem with blind people is how to navigate their way to wherever they want to go. Such people need assistance from others with good eyesight. As described by WHO, Blind Assistance System is a vision-based module specifically for BLIND VICTIMS. The system is designed in such a way in which the blind person can take the help of An Application which in turn sends Real-Time Frames to the Laptop-Based Wireless Networked System. It has a core feature of approximate distance calculation and Voice-Based wireless Feedback generation w.r.t the Distance Calculation.

It makes the work of Blind easy, efficient and reliable by sending wireless Voice based feedback whether the particular object is either too close to him or is it at a safer distance.

Therefore, the motivation of this research is given, the fact of building a technological tool supported by computer science (in this case Deep Learning) that allows to overcome some of these barriers, through the creation of a service that recognizes and automatically characterize images taken or provided by a user. The following text presents in the background, a historical tour of the evolution of computer science and, more precisely, the evolution that probabilistic algorithms have had, as well as works that likewise have wanted to identify objects in images using deep learning.

* 1. Objectives

The main purpose of object detection is to identify and locate one or more effective targets in real time. It comprehensively includes a variety of important techniques, such as image processing, pattern recognition, Machine learning .

1.2 Scope of the Project

The scope of the project is to help visually impaired people to identify the objects mainly Motor Vehicles, Traffic Lights, and Persons they face and calculate the distance of the object to avoid accidents and inconvenience on the roads by converting the recognized object to speech through applying various techniques.

# CHAPTER 2

## Literature Survey

Abdul Muhsin M, Farah F. Alkhalid, Bashra Kadhim Oleiwi have proposed their work on “Online Blind Assistive System Using Object Detection” in 2020. In this work, the function of computer vision is to detect indoor objects accurately. The visually impaired people can be assisted by navigating the purposes of the CNN framework.4,5,14 To identify the specific objects first, we need to detect the pixels available in the images. If the lighting conditions are wrong, then it is challenging to capture and identify the objects with high accuracy. To detect the indoor objects, the algorithm needs to extract the image features with a particular class, and it can be done by RetinaNet.25 To enable the network for small object detection by a Region Proposal Int J Cur Res Rev | Vol 12 • Issue 20 • October 2020 160 Mandhala et al.: Object detection using machine learning for visually impaired people Networks (RPN), which involves subsampling to obtain the image information. The Resort with 152 samples achieved an average precision with 83.1%, and Dense Net with 121 samples achieved an average precision with 79.8%. Pooja Maid, omkar Thorat and Sarita Deshpande have proposed their work on “Object Detection for Blind User’s” in 2018. Based on the relation models, this work assigned an equal quantity of work by considering its features. This removes duplication and attains accuracy at specific standards. Since the objects are aligned in the 2D scale ratio, it uses objects rather than words. Further, the model is categorized into two components that fall under geometric and original weights.15 Venkata Naresh Mandhala, Debnath Bhattacharyya, Vamsi B., Thirupathi Rao N. have proposed their work on “Object Detection Using Machine Learning for Visually Impaired People” in 2005. In this work, the main objective is to focus mainly on time complexities and their accuracies depending upon the various test that has been performed by the greedy approach the module which detects the text in the images which can be improved for visually impaired people. The quality of the model can be measured by F.P. and F.N. rates. The decision capability of the algorithm can be done by a set of training images and classifiers. The smart telescopic system will be used for vision problem people. On the micro screen visuals, the image represents itself in a emphasize way leaving certain spots of the image behind. this work mainly in the medical field, specific changes have been made in identifying the disease and recoveries by considering its effectiveness and accuracy. Magnetic resonance imaging (MRI) and computed tomography (C.T). are analysed on image processing algorithms that are highly examined rather than DSA. While DSA is referred to the diagnosis of several neurovascular conditions which is used at the time of surgeries, on these considerations, it could be concluded that the framework is designed based upon the patients diagnosed with ischemic stroke.18 Rui (Forest) Jiang have proposed their work on “Let Blind People See: Real-Time Visual Recognition with Results Connected to 3D Audio” in 2015. In this work, a CNN model produces the best performance for image classification with a single label. Due to complexity, multi labelling is an open challenge for training image layouts. A single image object is taken as an input will be given for hypotheses extraction, and this is shared with CNN to get individual scores by max pooling. The image’s hypotheses are identified with different colours that can be indicated by different clusters.10 The extraction method produces predictive results that are utilized by max pooling. By comparing the I-FT and HCP models, the HCP model improves the system performance by 5.7%. K.S.Manikanta , T.S.S. Phani have proposed their work on “Implementation & Design Of smart Blind Stick For Obstacle Detection And Navigation System” in 2018. In this work, the multi-model is used for visually impaired people to detect the objects with a multi-class strategy in an indoor area. This model takes at a time more than one label. The CVNN and multi-label techniques associate the image with labels that correspond to categories of objects at once.16 The clusters can be made based on multi labelling by ML-CVNN, and the L-CVNN method works by image transformation to classify the problem by ranking solution. The input strategy captures the image by multi-label and multi classes to generate the contexts of realistic and non-realistic of nested and exclusive structures. Dr. K. Sreenivasulu, P. kiran Rao have proposed their work on “A Comparative Review on Object Detection System for Visually Impaired” in 2016. This model is used for detecting the patterns in urban areas such as public streets, raining, restaurants, etc.13 This method characterizes the audio clips, which yields the patterns. The main limitation of this model is to require a trained data set. 6

J. Dharanidharan, R. Puviarasi have proposed their work on “Object Detection System for Blind People” in 2012. In this work, the Camera wearer’s day is a dense storyboard briefing the recommended methods. On the other hand, in traditional essential chunk selection techniques, the final presentation of these techniques mainly examines the vital objects and people who interact using this camera wearer. A few chucks/data packets required for the storyboard are reflected by the vital object-driven circum- 161 Int J Cur Res Rev | Vol 12 • Issue 20 • October 2020 Mandhala etc al.: Object detection using machine learning for visually impaired people stances in this method. Based on the practices 17 hours of self-centred data depending upon the existing techniques, it shows excellence in saliency and summarisation. This has been done in 4 main steps; they are: (a) the image about a famous person or object could be predicted using a novel self-centric salient cycles which it trains a group independent regression model. (b) Separation of each task/event In dividing the video into subcategories of tasks.

# CHAPTER 3

## Problem Statement/Proposed Method / Algorithm / Architecture / Methodology / Project Description

#### Problem Statement

Object Detection task is applied in Security cameras, Face recognition, Vehicle detection in traffic and many more areas. Visually impaired people face a lot of problem. This project aims broadly to help visually impaired people to detect persons, tracking vehicles and it’s distance. The field of computer vision has been developed as years passing by mainly deep learning is specifically conventional neutral nets. Whereas the image processing classifies or labels to image based on this content classifying and finding an unknown number of individual objects was considered an extremely difficult problem. Now, this task can be called as object detection is now adapted by the MNC’S such as IBM, GOOGLE etc.it can be rectify the major problem statement that is required for image recognition.

##### Proposed System

##### 

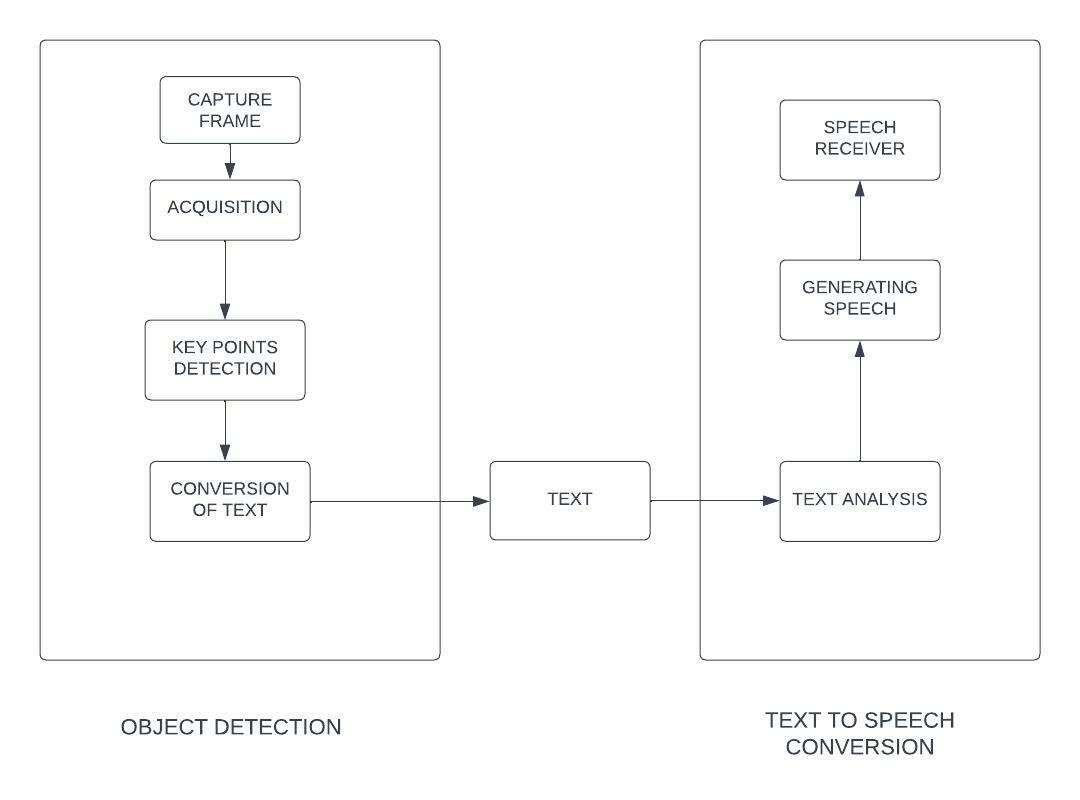
##### The usage of real-time object detection is widespread. Open CV idea that focuses on locating and detecting objects from several classes in the input image. Different approaches and strategies are used to detect the items, together with their distinctive characteristics and background information. Among them, one is building bounding boxes to enable detection and finding of pixels for all characteristics of the objects  in accordance with specific calculations. The suggested system's objective is to recognise items that are displayed in front of it and translate such recognitions into speech. To detect and identify items, object detectors can be created as a portable gadget that can be worn or taken anywhere. The proposed system uses You Only Look Once (YOLO) to detect objects in less time and high speed. The Proposed model is trained to recognize around 80 objects in real time.

###### Advantages:

###### High detection precision, low misdetection rate

###### Real-time detection, simple network structure.

###### Locate objects with rotation bounding box, using rotation anchors.

**Architecture: **

**Fig.1.System architecture**

* 1. **Algorithm**

Step 1: Start

Step 2: Load YOLO

Step 3: Loading Video

Step 4: Detecting Objects

Step 5: Showing information on the screen

Step 6: Stop.

### Methodology

**YOLO (You Only Look Once):**

The YOLO algorithm uses a convolutional neural network (CNN) to detect objects in real time. As its name suggests, this algorithm requires only a single forward propagation through the neural network to detect an object. The YOLO algorithm uses three techniques: Residual Block Bounding Box Regression Union (IOU)Residual Block-First, the image is divided into different grids. Each grid has dimensions S x S. The image below shows how the input image is divided into grids. CNNs are used to predict probabilities and bounding boxes for different classes simultaneously. The YOLO algorithm consists of different variants. The most common are tiny YOLO and YOLOv3.The YOLO algorithm works in three ways: Residual Block-Bounding Box Regression Union (IOU),Residual Block First. the image is divided into different grids. Each raster has dimensions S x S.

The ultimate idea of the YOLO algorithm is to identify a class of an object and the bounding

box pointing to object location. Each bounding box can be divided in descriptors:

1.Center of the box (bx,by)

2.Width (bw)

3.Height (bh)

4.Value c corresponds to the class of an object

Along with that large number of objects are predicted, which is the probability that there is oran object in the bounding box can be calculated.

#### Project Description:

Object Detection and tracking is one of the areas of computer vision that is maturing very rapidly. It allows us to identify and locate in an image or video. With this kind of identification and localization, object detection and tracking can be used to count objects in a particular scene and determine and track their precise locations, all while accurately labeling them.

In this project, we have made use the most popular Python libraries for object detection, OpenCV, numpy, and time.

# 

# 

# CHAPTER 4

## Implementation / Code / Results and Description of Results

#### Implementation

The Flowchart is a communication outline that shows how objects work with each other and in what order. The above Flowchart of the framework clarifies the stream that is first, the user starts and wears the system. Once the source code is implement, it will implement its internal process/code. The code keeps on executing till prescribed output arrive. Initially, we will import all the libraries that are: OpenCV, Pyttsx3, Time, and NumPy and will read the text file containing class names, YOLO weights, and configuration files. After that, the code will initialize the camera connected to it. The camera will capture real-time frames at 1fps (frame per second), then the code will read the input image/frame and get its width and height to an adequate level. Then an object detection algorithm in this case YOLO is applied to this altered frame. Before forward passing this altered image to YOLO weights and YOLO configuration files, a 'BLOB from image' is constructed. To obtain (correct) predictions from deep neural networks such as YOLO, you first need to pre-process your data. In the context of deep learning, feature extraction, and image classification, we have used the OpenCV function blobFromImage. This function performs the following:

1.Mean subtraction – is used to help combat illumination changes in the input images in the dataset.

2.Scaling by some factor – is used to scale the input image space into a particular range

3.Optionally channel swapping.

Then the code performs a forward pass of the YOLO object detector, giving us our bounding boxes, class ids, and associated class probabilities.

Another advantage of YOLO other than being fast is that it provides three methods to improve its performance:

Intersection over Union (IoU) decides which predicted box is giving a good outcome. It calculates the IoU of the actual bounding box and the predicted bounding box.

Non-max suppression suppresses weak, overlapping bounding boxes.

Anchor Boxes detects multiple objects in a single grid.

Further, the frames are divided into a 3×3 grid, which helps in finding the position of objects. The system aims to produce an audio output for the visually impaired. The Detected object labels are converted into speech using the pyttsx3 library.

Lastly, Upon successful recognition of an object and as per grids, the system will provide speech output stating the name of the object along with its grid name, for e.x. Mid left car, Mid right car. Hence helping the visually impaired people in recognizing the objects in the field of view.

Division of Image into grids

###### Required Specifications

Python:

Visual Studio Code

Python 3.10.8 (Above 3 versions)

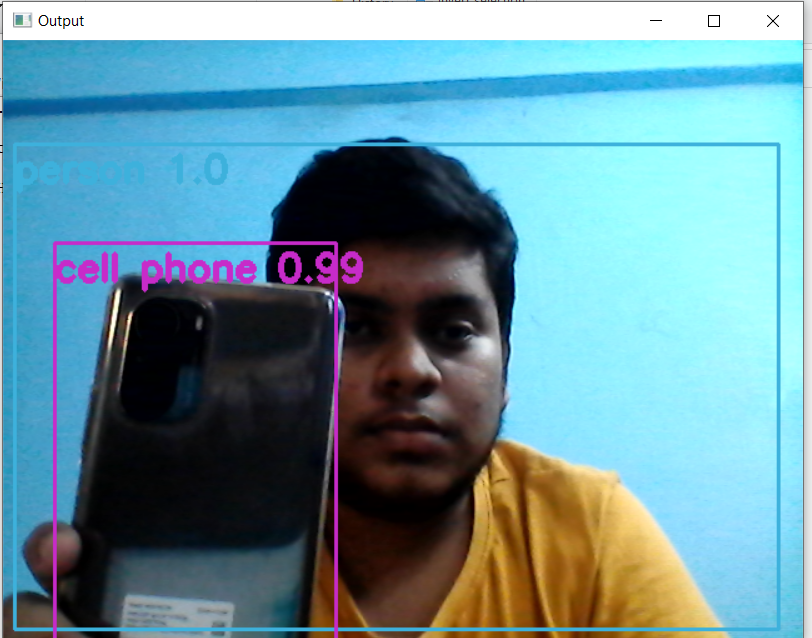
**Requirements:**

**Anaconda** 3.x(version)

Cuda 10**(version)**

**Operating System:** Windows 10,11

* 1. Code:
  2. import cv2
  3. import numpy as np
  4. import time
  5. import pyttsx3
  6. # Initialize text-to-speech engine
  7. engine = pyttsx3.init()
  8. # Load Yolo
  9. net = cv2.dnn.readNet("yolov3.weights", "yolov3.cfg")
  10. classes = []
  11. with open("coco.names", "r") as f:
  12. classes = [line.strip() for line in f.readlines()]
  13. layer\_names = net.getLayerNames()
  14. output\_layers = [layer\_names[i - 1] for i in net.getUnconnectedOutLayers()]
  15. colors = np.random.uniform(0, 255, size=(len(classes), 3))
  16. # Loading video
  17. cap = cv2.VideoCapture(0)
  18. font = cv2.FONT\_HERSHEY\_SIMPLEX
  19. starting\_time = time.time()
  20. frame\_id = 0
  21. while True:
  22. \_, frame = cap.read()
  23. frame\_id += 50
  24. height, width, channels = frame.shape
  25. # Detecting objects
  26. blob = cv2.dnn.blobFromImage(frame, 0.00392, (416, 416), (0, 0, 0), True, crop=False)
  27. net.setInput(blob)
  28. outs = net.forward(output\_layers)
  29. # Showing info on the screen
  30. class\_ids = []
  31. confidences = []
  32. boxes = []
  33. for out in outs:
  34. for detection in out:
  35. scores = detection[5:]
  36. class\_id = np.argmax(scores)
  37. confidence = scores[class\_id]
  38. if confidence > 0.2:
  39. # Object detected
  40. center\_x = int(detection[0] \* width)
  41. center\_y = int(detection[1] \* height)
  42. w = int(detection[2] \* width)
  43. h = int(detection[3] \* height)
  44. # Rectangle Coordinates
  45. x = int(center\_x - w / 2)
  46. y = int(center\_y - h / 2)
  47. boxes.append([x, y, w, h])
  48. confidences.append(float(confidence))
  49. class\_ids.append(class\_id)
  50. indexes = cv2.dnn.NMSBoxes(boxes, confidences, 0.8, 0.3)
  51. for i in range(len(boxes)):
  52. if i in indexes:
  53. x, y, w, h = boxes[i]
  54. label = str(classes[class\_ids[i]])
  55. confidence = confidences[i]
  56. color = colors[class\_ids[i]]
  57. cv2.rectangle(frame, (x, y), (x + w, y + h), color, 2)
  58. cv2.putText(frame, label + " " + str(round(confidence, 2)), (x, y + 30), font, 1, color, 3)
  59. # Convert label and confidence to speech
  60. text = label + " with confidence " + str(round(confidence, 2))
  61. engine.say(text)
  62. engine.runAndWait()
  63. elapsed\_time = time.time() - starting\_time
  64. fps = frame\_id / elapsed\_time
  65. cv2.imshow("Output", frame)
  66. key = cv2.waitKey(1)
  67. while True:
  68. key = cv2.waitKey(1)
  69. if key == 27: # Check if 'esc' key is pressed
  70. break
  71. cap.release()
  72. cv2.destroyAllWindows()
  73. **Results:**

****

* 1. **Fig.2. Detection of Person**

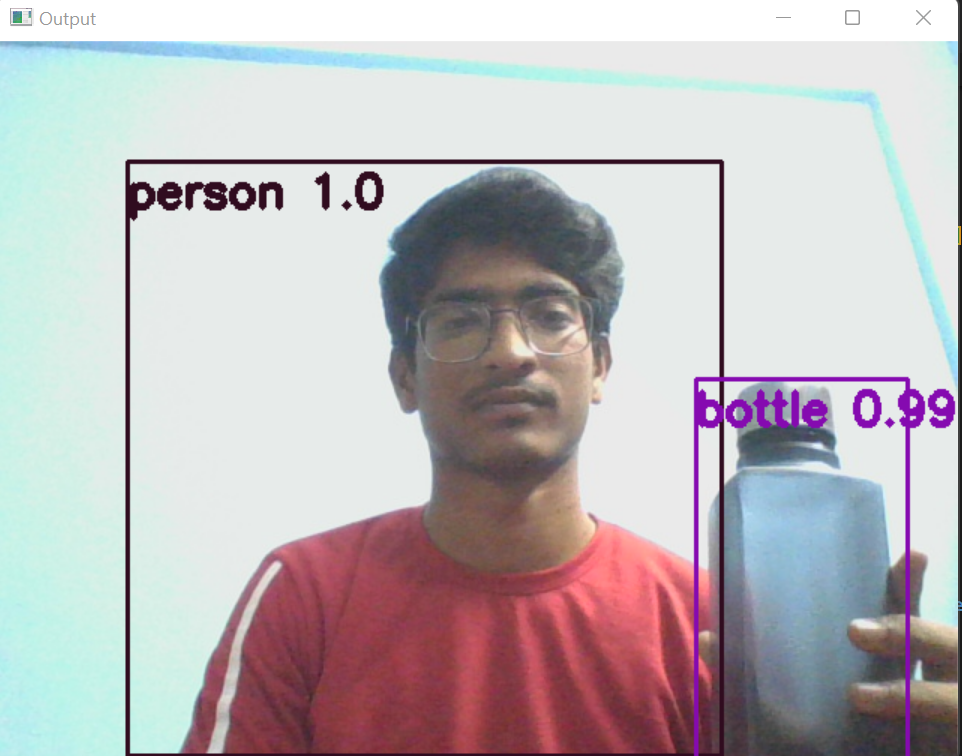
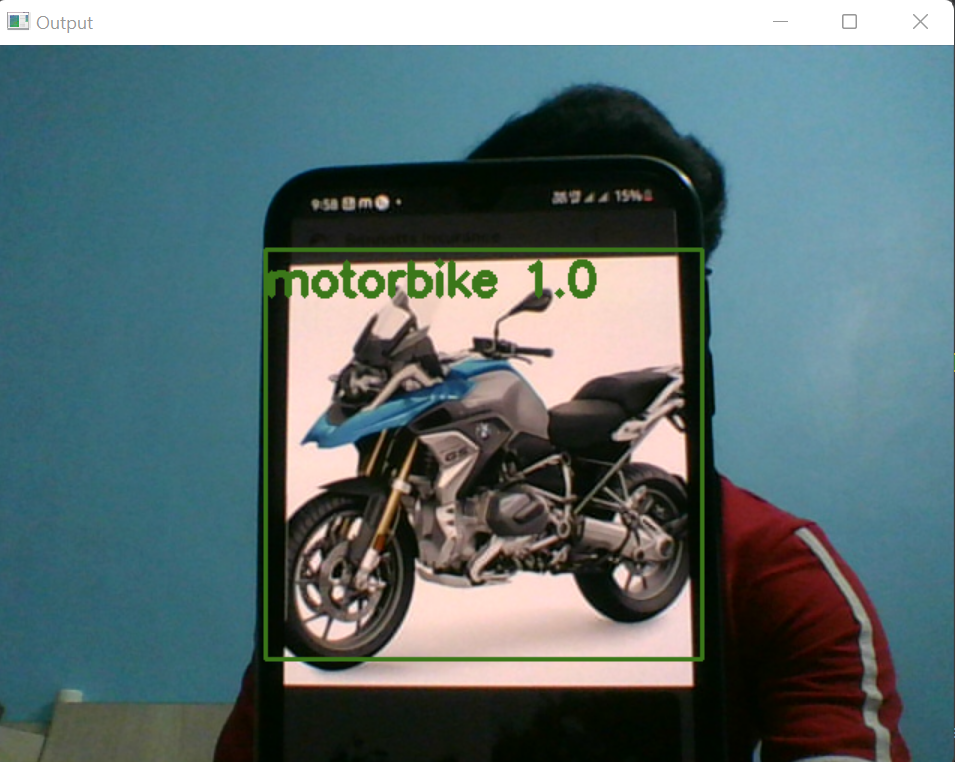
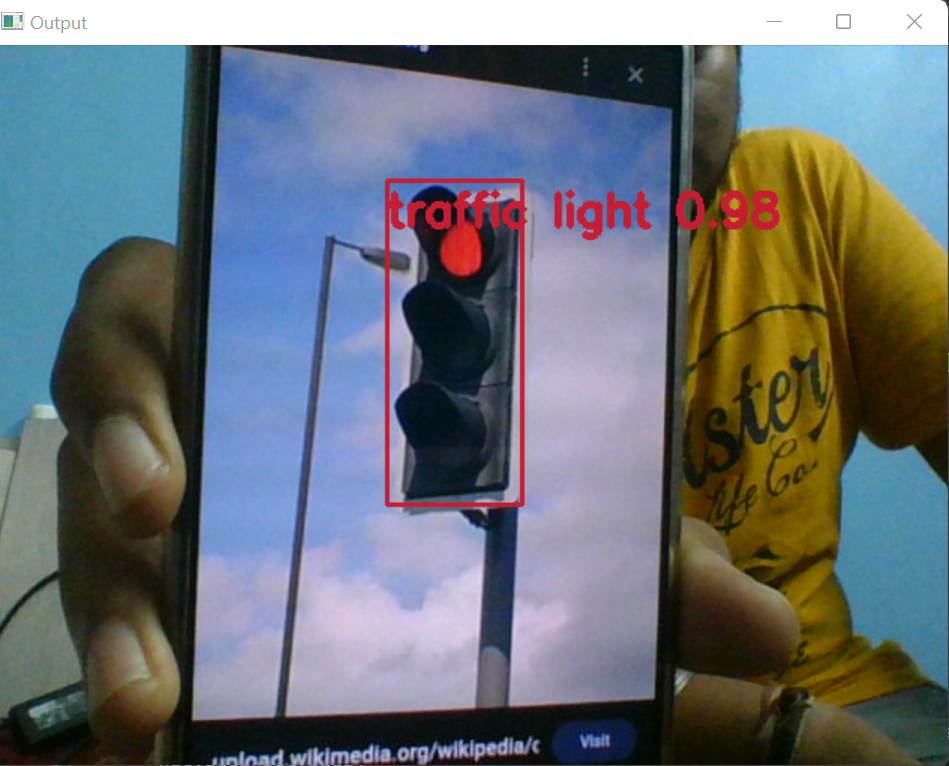


Fig:3 Detection of Person

****

**Fig.4. Detect Motorvehicle**

****

**Fig.5. Detect Traffic light**

# 

# 

# CHAPTER 5

## Conclusion/Future Enhancement/References

#### Conclusion

#### Object recognition technology has experienced decades of development. From traditional computer vision technology to the current popular deep learning recognition technology, the recognition effect is more and more accurate, and the algorithm is more robust. But the better the effect and efficiency of the algorithm means the improvement of the network and hardware requirements, which is a relatively large restriction for the application of the algorithm. And the current recognition algorithm based on deep learning requires a sufficient number of training sets to train the model to achieve a good recognition effect. This is not a priority for objects without a large number of similar object data sets, such as some rare collections in museums. In addition, the current object recognition almost always requires network transmission of images, and then transmission of recognition feedback results through the network, which is much less efficient in environments with relatively poor network conditions, such as the wild and crowded indoor environments. In summary, the future development prospects of object recognition technology should not be limited to the recognition accuracy, recognition efficiency, algorithm robustness, etc., but also specific issues should be considered. For objects with different characteristics and the environment in which the objects are located, designing corresponding recognition algorithms and considering the degree of dependence on network and equipment performance should also be a focus of future research.

#### 5.2 Publication Details

#### 1. Name of the Journal published: International Journal Of Creative Research Thoughts (IJCRT).

#### 2. ISSN Number : 2320-2882.

#### 3. Edition : Volume II Issue 3 March 2023.

#### 4. Published Web-link :www.ijcrt.org.

#### Proofs:

#### 

**Fig.6. publication details**

**References:**

1. Sunit Vaidya, Niti Shah,Naisha Shah, Prof. Radha Shankarmani Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) IEEE Xplore Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2.

2. Online Blind Assistive System using Object Recognition, International Research Journal of Innovations in Engineering and Technology, Vol. 3, Iss. 12, (Dec 2019) 47-51.

3. A Comparative Review on Object Detection System for Visually Impaired, Turkish Journal of Computer and Mathematics Education (TURCOMAT 2021) Vol. 12 No.2.

4. J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” arXiv preprint arXiv:1804.02767, 2018. 2

5. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2818–2826. 1

6. Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition

7. Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng..

Tensorflow: Largescale machine learning on heterogeneous distributed systems, 2016, arXiv:1603.04467. 2

8. S. Ren, K. He, R. B. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” CoRR, vol. abs/1506.01497, 2015. [Online]. Available: http://arxiv.org/abs/1506. 01497. 1

9. Bengio, Y. (2012). “Deep learning of representations for unsupervised and transfer learning,” in

ICML Unsupervised and Transfer Learning, Volume 27 of JMLR Proceedings, eds I. Guyon, G. Dror,

V. Lemaire, G. W. Taylor, and D. L. Silver(Bellevue: JMLR.Org), 17–36.

10. Bourdev, L. D., Maji, S., Brox, T., and Malik, J. (2010). “Detecting people using mutually consistent

Pose let activations,” in Computer Vision – ECCV2010 – 11th European Conference on Computer

Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part VI, Volume 6316 of

Lecture Notes in Computer Science, eds K. Daniilidis, P. Maragos, and N. Paragios

(Heraklion:Springer), 168–181.

11. Bourdev, L. D., and Malik, J. (2009). “Poselets: body part detectors trained using 3dhuman pose

annotations,” in IEEE 12th International Conference on ComputerVision, ICCV 2009, Kyoto, Japan,

September 27 – October 4, 2009 (Kyoto: IEEE),1365–1372

12. Anitha.J, Subalaxmi.A, Vijayalakshmi.G Real Time Object Detection for Visually Challenged Persons International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-8 June, 2019

13. Aniqua Nusrat Zereen, Sonia Corraya Detecting real time object along with the moving direction for visually impaired people 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE).

14. K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition",Proceedings of the IEEE conference on computer vision and pattern recognition*,* pp. 770-778, 2016*.*

15. Jamal S. Zraqou, Wissam M. AlKhadour and Mohammad Z. Siam, "Real-Time Objects Recognition Approach for Assisting Blind People"*,*International Journal of Current Engineering and Technology, 2017.

16. A. Vidyavani, K. Dheeraj, M. Rama Mohan Reddy and KH. Naveen Kumar, "Object Detection Method Based on YOLOv3 using Deep Learning Networks", vol. 9, no. 1, November 2019, ISSN 2278-3075.

17. R. L. Galvez, A. A. Bandala, E. P. Dadios, R. R. P. Vicerra and J. M. Z. Maningo, "Object Detection Using Convolutional Neural Networks", TENCON 2018–2018 IEEE Region 10 Conference*,* 2018.

#### 18. N.A. Othman and I. Aydin, "A new deep learning application based on movidius ncs for embedded object detection and recognition", 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 2018.*

#### 19. Wei Xiang, Dong-Qing Zhang, Heather Yu and Vassilis Athitsos, "Context-AwareSingle-ShotDetector", 2018 IEEE Winter Conference on Applications of Computer Vision*,* pp. 1784-1793, 2018.

#### 20. Abbas Q, Ibrahim MEA, Arfan Jaffar M (2019) A comprehensive review of recent advances on deep vision systems. Artif Intell Rev 52(1):39–76