IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Real-Time Object Identification For Visually Impaired Peers

[1]S Dheeraj Kumar, [2]M Senthil Kumaran, [3]J Himavanth

[1] Student, [2]Associate professor, [3] Student

Department of Computer Science and Engineering,

SCSVMV, Kanchipuram, India.

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.It acts as a basic and most aiding tool for blind people or visually challenged people. Nowadays blind people can recognize surroundings through sensors and Actuators according to the situation around them. Some of the popular algorithms for Object detection are Region-based Convolutional Networks (RCNN), Single Shot Detector(SSD), Faster-RCNN, and You Look Only Once (YOLO). 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:

NumPy, OpenCV, Time, Video Capture, YOLO.

Introduction:

For breaking down and clarifying the visual environment, which explains computer vision in computer technology. Machines utilize deep learning models17 and digital pictures from cameras and videos to accurately label the items. Demonstrations of computer vision to recognize edges and align simpler objects with falling under categories such as circles and squares using the methods of the first neural networks began in the 1950s. Optional character recognition for computer vision, handwritten data on its main trade tool, was created in the 1970s. The illustrated data is mostly utilized as a development for the blind. People with visual impairments find it challenging to see even the slightest detail with healthy eyes. The horizontal range of the visual field with both eyes open is less than or equal to 20 degrees for those with a visual acuity of 6/60. These individuals are said to be blind. According to a WHO (World Health Organization) survey from 2021, there are 1 billion people in the world who have moderate or severe distance vision impairment from untreated refractive error (88.4 million cases), cataract (94 million cases), glaucoma (7.7 million cases), corneal opacities (4.2 million cases), diabetic retinopathy (3.9 million cases), and trachoma (2 million cases), as well as near vision. The major challenge for blind persons is finding their way to their desired destination. Such folks require the support of others with good vision. According to WHO, the Blind Assistance System is a vision-based module designed especially for victims who are blind. The system is set up such that an application that delivers real-time frames to a laptop-based wireless networked system can be used by a blind person. Its primary features include voice-based wireless feedback generation and approximative distance estimation.

It sends wireless voice-based feedback on whether a certain object is either too close to him or is it at a safer distance, making the task of the blind simple, effective, and trustworthy.

The goal of this research is to develop a technical solution supported by computer science, namely Deep Learning, that enables overcoming some of these limitations by developing a service that automatically detects and describes photographs shot or submitted by a user. The following article provides a historical overview of the development of computer science and, specifically, the growth of probabilistic algorithms. It also discusses related efforts that aimed to use deep learning to recognize objects in photos.

Literature Survey:

Sunit Vaidya, Niti Shah, Naisha Shah, and Prof. Radha Shankarmani have Proposed their work on "Real-Time Object detection For Visually Challenged People" in 2020. In this work, The objective of the proposed work is to change the visual world into an audio world by notifying the blind people about the objects in their path. This will help visually impaired people to navigate independently without any external assistance just by using the real-time object detection system. The application uses image processing and machine learning techniques to determine real-time objects through the camera and inform blind people about the object and its location through the audio output.

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.

Pooja Maid, omkar Thorat, Sarita Deshpande have proposed their work on "Object Detection for Blind Users" 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.

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, rain, restaurants, etc.

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.

Rui (Forest) Jiang 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-labeling is an open challenge for training image layouts. A single image object is taken as 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%.

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 problems. This project aims broadly to help visually impaired people to detect persons, and track vehicles and their distance. The field of computer vision has been developed as years pass by mainly deep learning is specifically conventional neural nets. Whereas image processing classifies or labels to images 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 object detection is now adapted by MNC's such as IBM, GOOGLE, etc. It can 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 recognize 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.

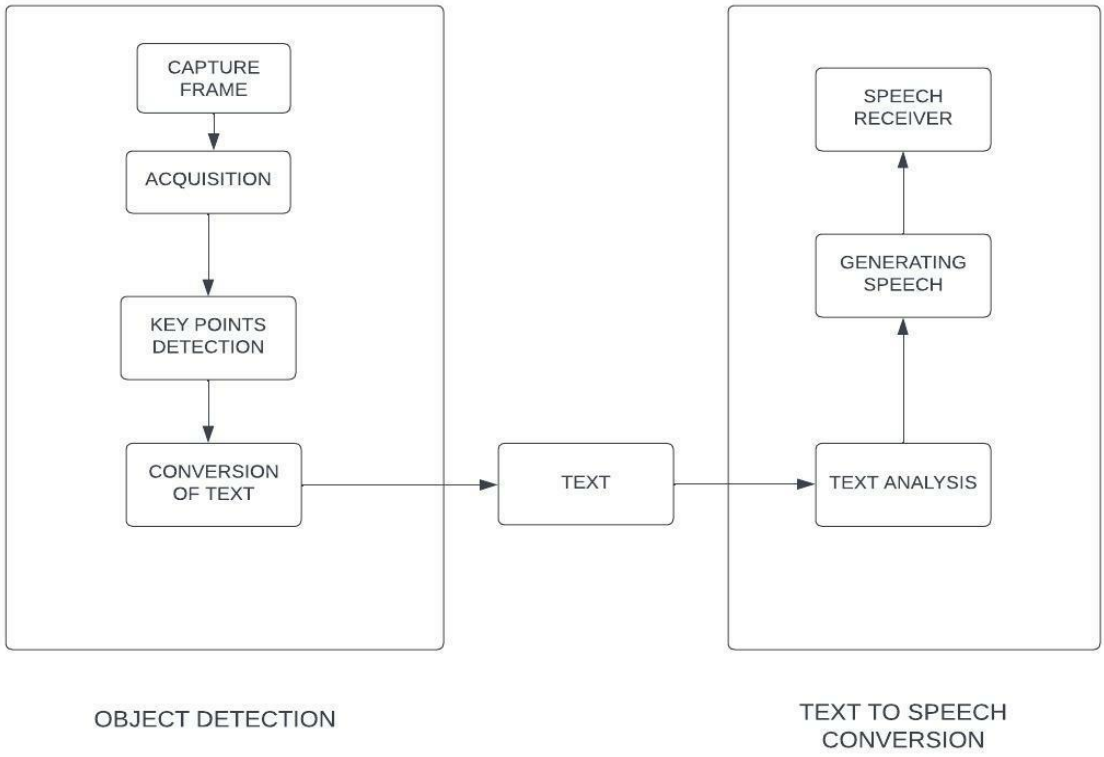

Fig1:System Architecture

Methodology:

Convolutional Neural Networks (CNN) are used by the YOLO method to recognize items instantly. This technique, as its name suggests, only needs one forward propagation through the neural network to identify an item. Three strategies are used by the YOLO algorithm: Bounding box regression with residual block union (IOU) Remaining Block- The image is first separated into several grids. Each grid has S x S dimensions. The grids created from the input image are displayed in the image below. CNNs are used to simultaneously forecast the probability and bounding boxes for various classes. There are various iterations of the YOLO algorithm. Little YOLO and YOLOv3 are the most popular. Residual Block-Bounding Box Regression Union (IOU), Residual Block, and the YOLO algorithm all function in three different ways. The image is first separated into several grids. Each raster has S x S dimensions.

The YOLO algorithm's main goal is to identify an object's class and the bounding box that points to the object's position. Our descriptors allow us to categorize each bounding box as follows:

1. In the box's centre (bx, by)
2. Width (bw)
3. Height (bh)
4. Value c corresponds to the class of an item. In addition, we forecast a real number of objects, making it possible to determine the likelihood that there is an object inside the bounding box.

Experimental Analysis:

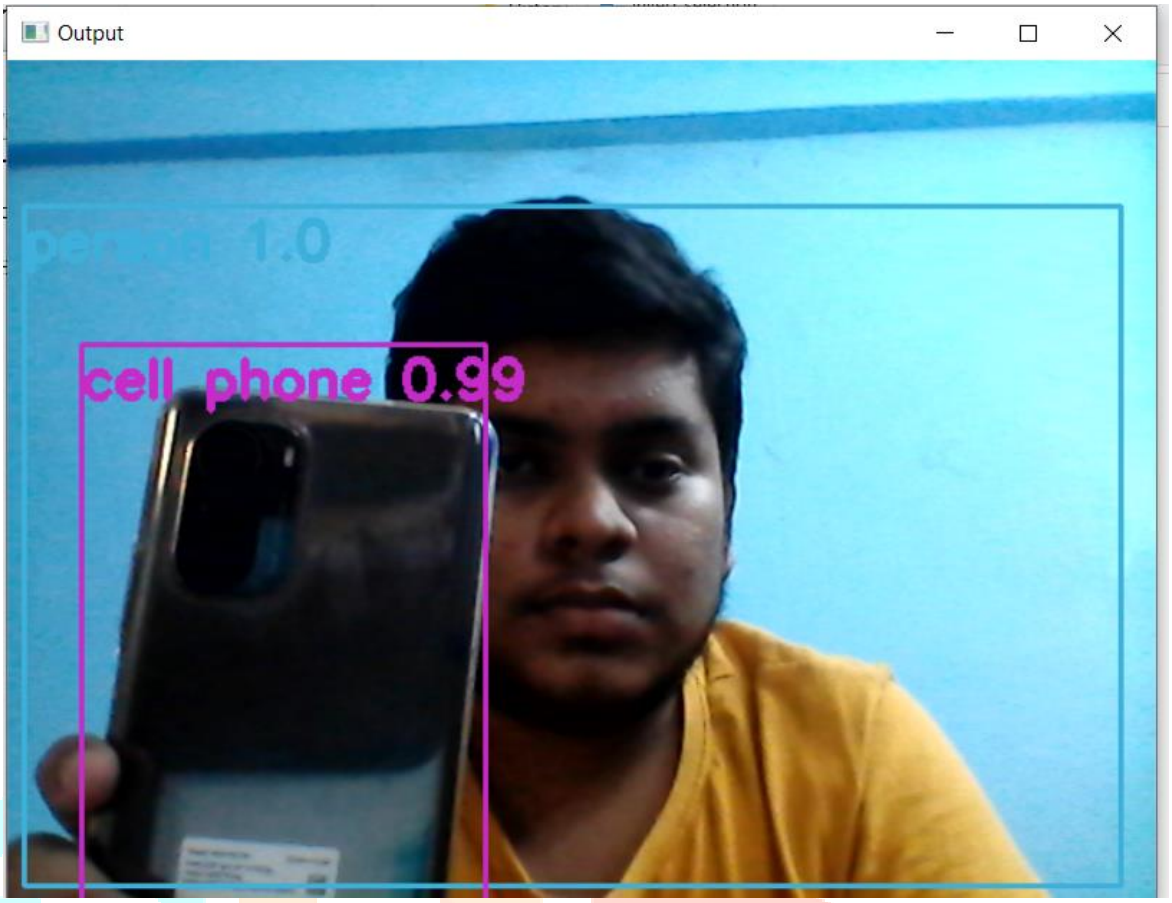

Fig.2. Detection of Person

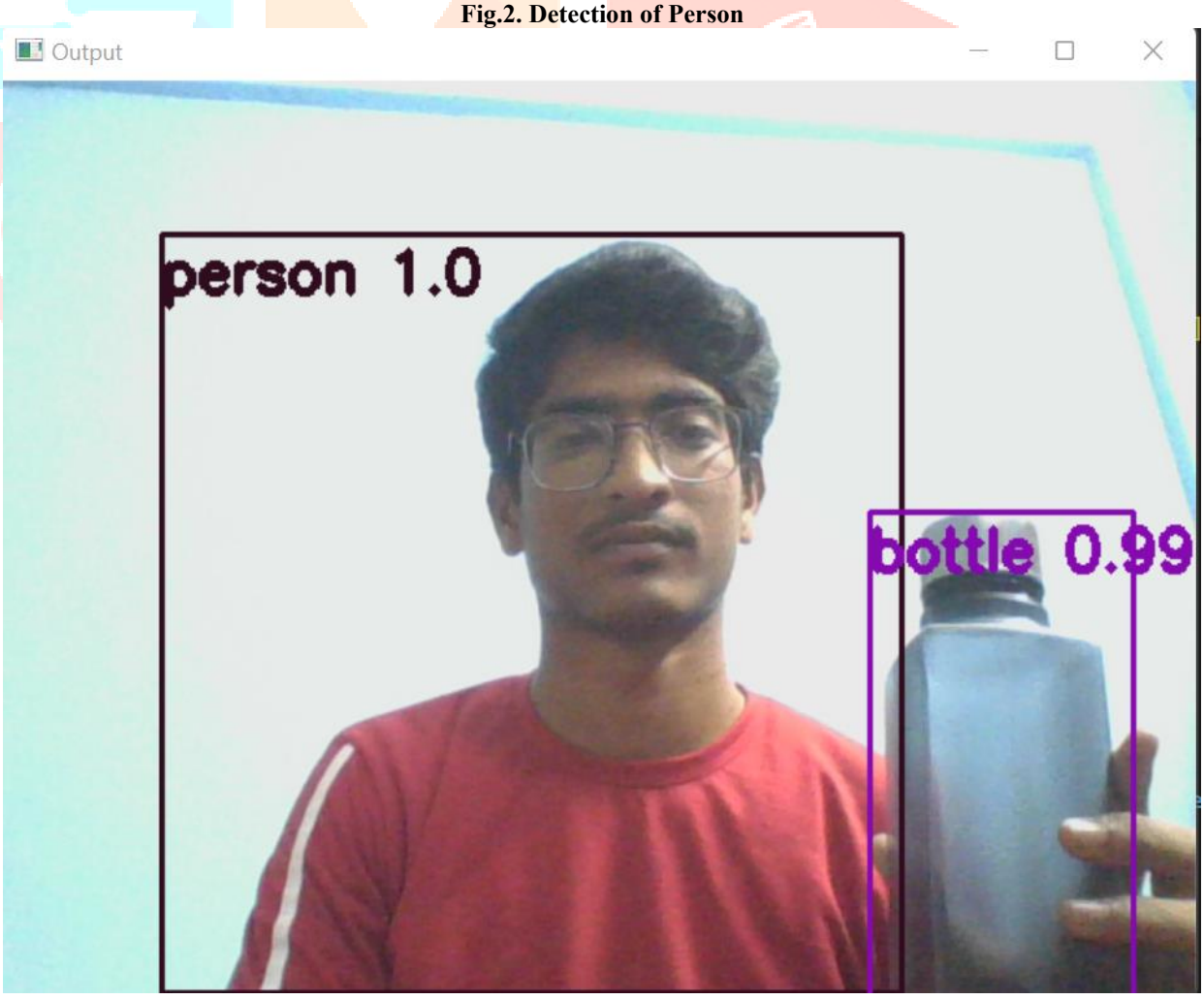

Fig:3 Detection of Person

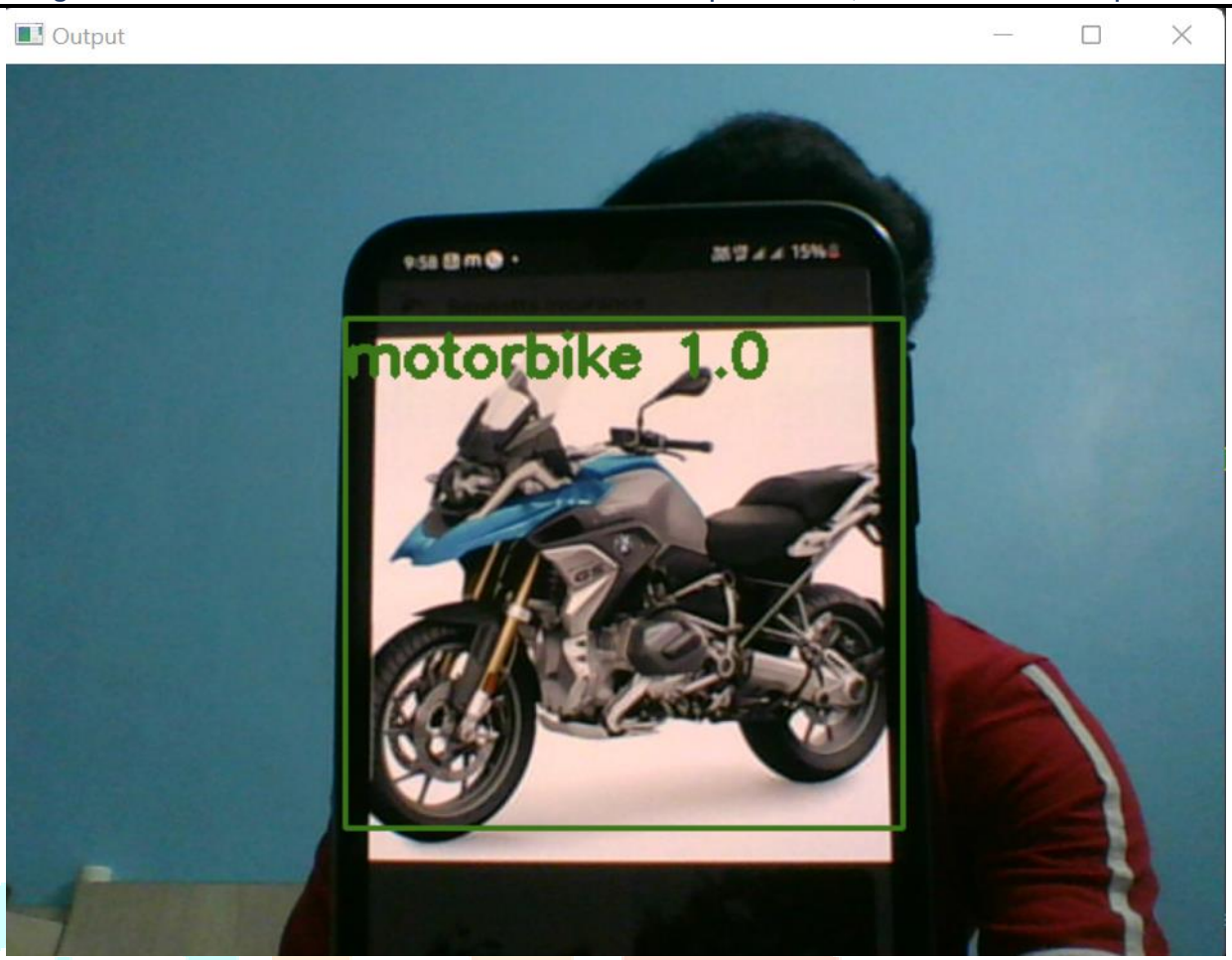

Fig.4. Detect Motor vehicle

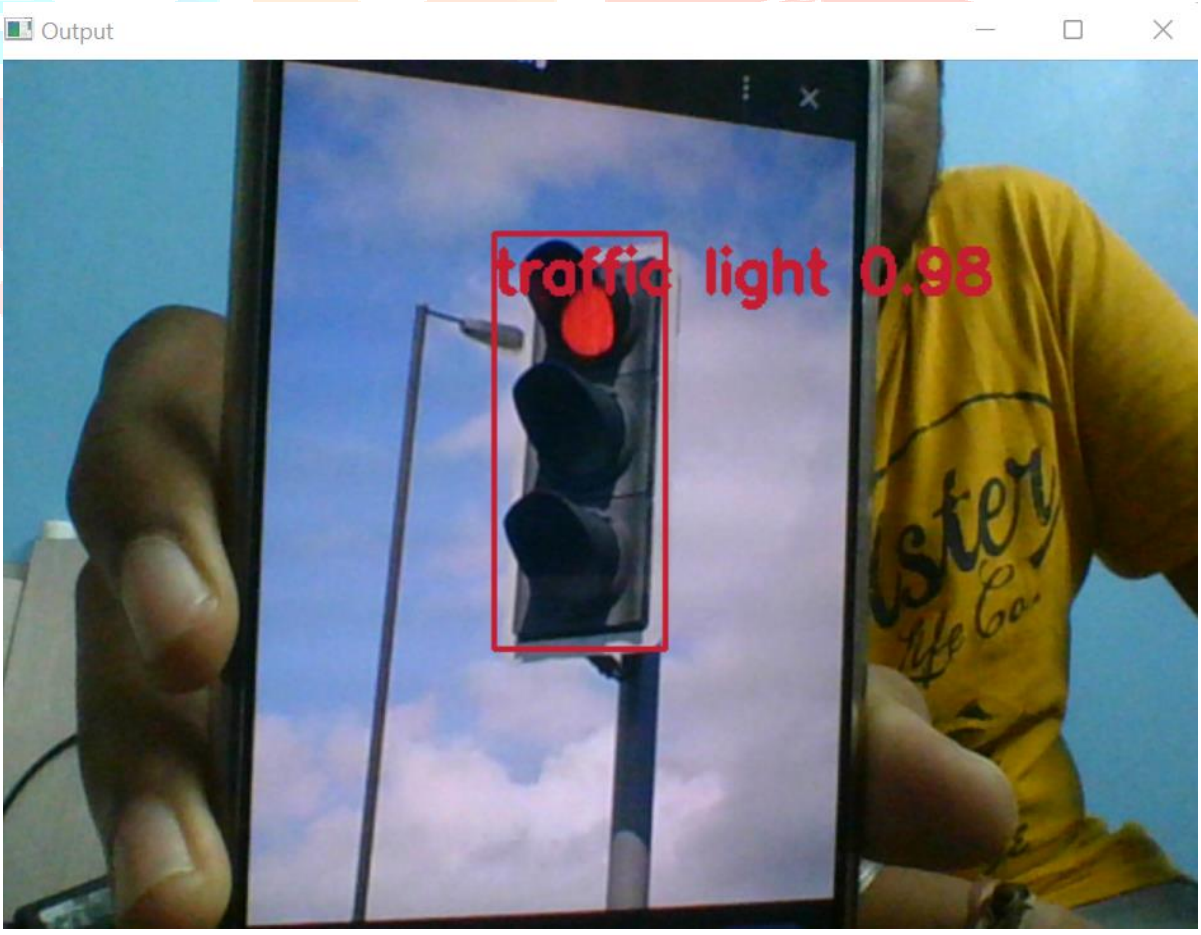

Fig.5. Detect Traffic light

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).