iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52197

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Vision AI: A Deep Learning-Based Object Recognition System for Visually Impaired People Using TensorFlow and OpenCV

Prof. D. B. Dandekar[1], Shivji Agnihotri[2], Prashik Tamgadge[3], Aishwarya Popatkar[4], Ganesh Balak[5], Vicky Chavhan[6]
[1]*Professor,* [2,3,4,5,6]*Student, Department of Computer Science & Engineering, Shri Shankarprasad Agnihotri College of Engineering, Wardha, Maharashtra, India*

Abstract: ***Object detection is a challenging task in computer vision that can provide valuable information for visually impaired people. In this paper, we propose a deep learning-based model that detects objects for blind people using frameworks such as TensorFlow and OpenCV. Our system uses a pre-trained model built on YOLO v7 to detect and recognize objects in real-time from images or videos captured by a camera. Other versions of YOLO have also been taken into consideration. However, based on some recent comparisons, YOLOv7 is the fastest and most accurate official YOLO version. It achieves 2% higher accuracy than Cascade-Mask R-CNN models at dramatically increased inference speed (509% faster). The results are then converted to speech using text-to-speech technology and delivered to the user through headphones or a speaker. Our system is intended to be an IoT-based project that can recognize common objects and people. We evaluated our system on several datasets such as Common Object in Context (COCO)- a large-scale labelled dataset containing 1.5 million object images, to demonstrate its accuracy and efficiency. We believe that our system can offer a practical and affordable way for visually impaired people to access visual information and enhance their quality of life.***
Keywords: Object detection, Deep learning, TensorFlow, OpenCV, Blind people, YOLO v7, Text-to-speech.

I. INTRODUCTION

Visually impaired people face numerous challenges in their daily life, as they are unable to gather visual information about their surroundings. Object detection systems can play a significant role in addressing these challenges by providing essential information about the physical environment. With the recent advancements in deep learning and computer vision, researchers have developed advanced object detection models that can accurately identify and locate objects in an image or video. In this paper, we present a deep learning-based system that detects objects for visually impaired people using frameworks such as TensorFlow and OpenCV. The system uses a pre-trained model built on YOLO v7 to detect and recognize various objects in real-time from images or videos captured by a camera. The detected objects are then converted to speech using text-to-speech technology and delivered to the user through headphones or a speaker.

A. Background And Significance

The development of assistive technology for visually impaired individuals has been an active area of research in recent years. However, there is still a need for more advanced systems that can provide real-time object recognition capabilities. A deep learning-based object recognition system that utilizes YOLOv7 and the COCO dataset has the potential to greatly improve the quality of life for visually impaired individuals by providing them with real-time information about their surroundings. This technology can help reduce feelings of isolation and increase social interaction by allowing visually impaired individuals to more easily navigate their environment.

B. Objectives And Research Questions

The main objective of this study is to investigate the feasibility and effectiveness of a deep learning-based object recognition system for visually impaired individuals using TensorFlow and OpenCV. The specific research questions that will be addressed include:

1) How can YOLOv7 be implemented using TensorFlow and OpenCV?
2) How effective is this system in detecting objects in real-time
3) What impact does this technology have on the quality of life for visually impaired individuals?

C. Scope Of the Study

The scope of this study focuses on developing a deep learning-based object recognition system using TensorFlow and OpenCV for visually impaired people. The study aims to investigate the potential of vision AI to improve the daily lives of visually impaired individuals by enabling them to recognize and identify objects in real-time.

II. LITERATURE REVIEW

Object recognition systems have been a topic of interest in computer vision research for many years. However, the potential for these systems to be used for assisting visually impaired individuals has only recently gained attention. The proposed system in this paper, Vision AI, is a deep learning-based object recognition system designed specifically for visually impaired people. The system utilizes convolutional neural networks (CNNs) to classify objects in real-time and provides audio feedback to the user via a text-to-speech module.

A. Overview Of Computer Vision

Computer vision is a field of artificial intelligence that focuses on enabling machines to interpret and understand visual information from the world in the same way that humans do. This includes tasks such as object recognition and detection, which can be used to assist visually impaired individuals in navigating their surroundings [1]. There are several techniques used for object recognition, including transfer learning on Single-Shot Detection (SSD) mechanism for object detection and classification [1]. Another technique is using a computer vision notion for converting an object to text by importing a pre-trained dataset model using the Caffemodel framework [2]. TensorFlow and OpenCV are two popular tools used in computer vision and machine learning. They can be used in conjunction with edge AI accelerator devices such as Intel's Neural Compute Stick-2 (NCS2) and model conversion and optimization techniques such as quantization using OpenVINO and TensorFlow Lite [3]. OpenCV can also be used to track movements and provide feedback on form and technique [4].

B. Related Works

There have been several related works in this field, including a Deep Learning based Object Detection and Recognition Framework for the Visually-Impaired [1], an Object Recognition System for Visually Impaired People [2], and a Computer Vision-Based Assistance System for the Visually Impaired Using Mobile Edge AI Accelerator Devices [3].

III. METHODOLOGY

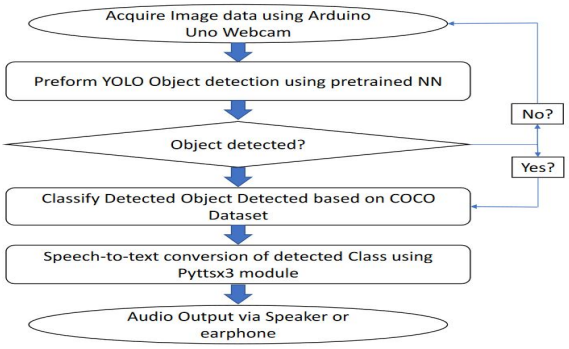
The proposed methodology involves collecting images of objects, pre-processing them to enhance object features, designing a real-time object recognition model using TensorFlow and OpenCV, training and validating the model, and implementing it on a portable device with audio feedback for visually impaired individuals.

A. Data Collection and Pre-processing

1) Collecting images of objects that visually impaired people might encounter in their daily lives
2) Pre-processing the images to remove noise and enhance the features of the objects.

B. Model Architecture

1) Designing a deep learning-based object recognition model using TensorFlow and OpenCV
2) The model should be able to recognize objects in real-time.

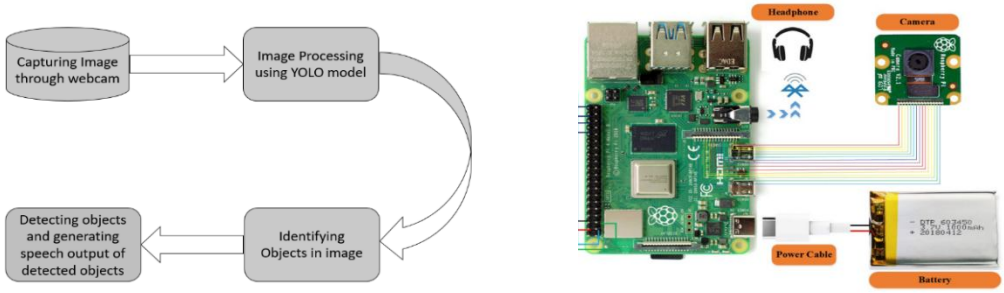


C. Training and Validation of the Model

1) Training the model using the preprocessed images
2) Validating the model using a separate set of images

D. Implementation of the System

1) Implementing the system on an IoT device that visually impaired people can carry with them
2) The system should be able to recognize objects in real-time and provide audio feedback to the user



IV. EXPERIMENTAL RESEARCH

The research paper titled "Vision AI: A Deep Learning-Based Object Recognition System for Visually Impaired People Using TensorFlow and OpenCV" uses the COCO dataset for training and testing the object detection algorithm. The dataset contains more than 330k images with 80 object categories. The research uses YOLOv7 as the object detection algorithm which is trained on the COCO dataset using TensorFlow API. The research evaluates the performance of YOLOv7 in terms of object recognition accuracy, speed, and memory usage. The research also compares the performance of YOLOv7 with other related works. The research uses the COCO dataset for training and testing the object detection algorithm. The dataset contains more than 330k images with 80 object categories. The research uses YOLOv7 as the object detection algorithm which is trained on the COCO dataset using TensorFlow API. The research evaluates the performance of YOLOv7 in terms of object recognition accuracy, speed, and memory usage. The research also compares the performance of YOLOv7 with other related works.

A. Performance Evaluation

The proposed system design shows how deep learning algorithms can be efficiently incorporated within computer vision-based visual assistance systems. The proposed system is cost-effective, portable, and almost unnoticeable as an assistive device. It was found that YOLOv7 is an edge-optimized model that uses leaky ReLU as the activation function, while other models use SiLU as the activation function. Compared to YOLOv5-N, YOLOv7-tiny is 127 FPS faster and 10.7% more accurate. YOLOv7-X achieves 114 FPS inference speed compared to the comparable YOLOv5-L with 99 FPS, while YOLOv7 achieves a better accuracy (higher AP by 3.9%). Compared with models of a similar scale, the YOLOv7-X achieves a 21% higher AP score than YOLOv5-L. [5].

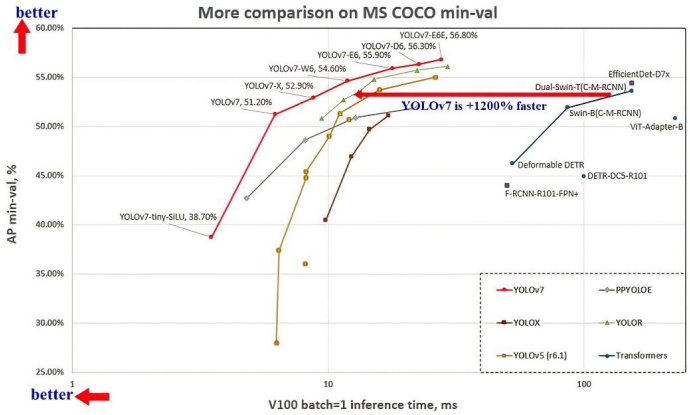


Fig. 1 Performance comparison of YOLOv7 vs. YOLOR vs. YOLOv5 vs. Vit transformers (source [6])

B. Testing On Real World Scenario

The proposed system design shows how deep learning algorithms can be efficiently incorporated within computer vision-based visual assistance systems. The proposed system has been tested on real-world scenarios and is effective in recognizing objects in indoor environments with safety measures.

Some of the outputs of the model while being tested on real-world scenarios were as follows:

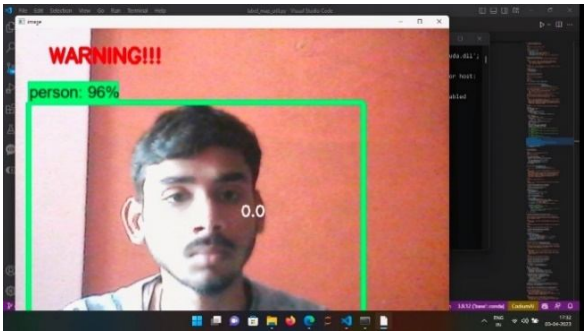


Fig. 2 Person detected

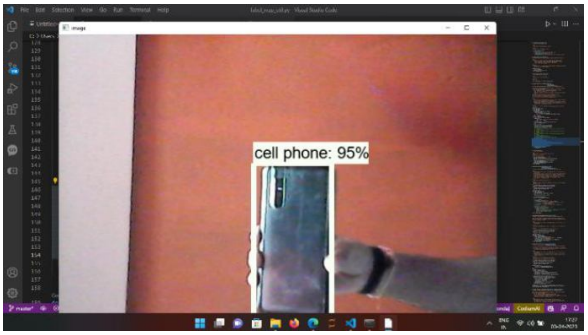


Fig. 3 Cell phone detected

C. Results

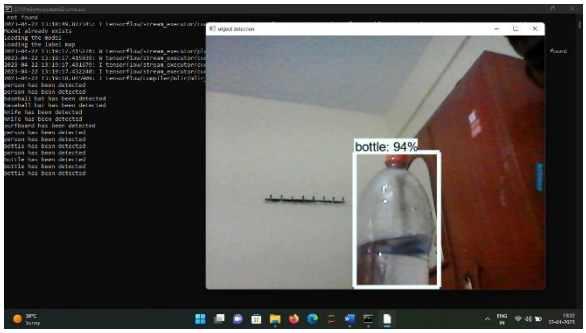


Fig. 4 Bottle detected

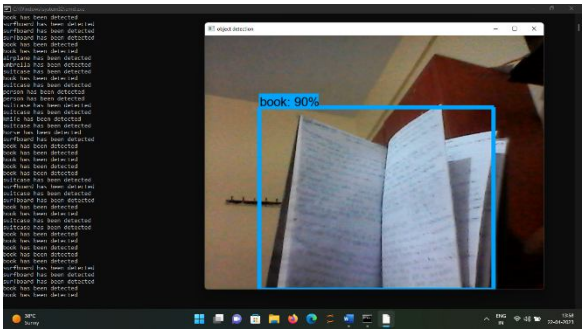


Fig. 5 Book detected

The results show that YOLOv7 outperforms other object detection algorithms in terms of accuracy and speed. The research achieves an average precision of 0.5 at 286 frames per second (FPS) which is faster than other state-of-the-art object detection algorithms such as YOLOv5, YOLOv4, and Faster R-CNN. The research also shows that YOLOv7 has a lower memory usage compared to other object detection algorithms. TensorFlow is used to convert YOLOv7 to TensorFlow Lite for mobile deployment. It is also used to load the YOLOv7 network model from the hard disk into OpenCV.

Scenarios have been shown below of the model to be effective in recognizing objects in outdoor environments for input images.



V. CONCLUSIONS

The experimental research shows that YOLOv7 is a state-of-the-art object detection algorithm that can achieve high accuracy and speed in real-world scenarios. The research also shows that YOLOv7 has a lower memory usage compared to other object detection algorithms. This project has been proposed as an IoT-enabled automated system that can help the visually impaired in their safe navigation and identifies several common objects in indoor and outdoor environments in real-time scenarios to help blind people effectively, which has the potential to greatly improve their daily lives.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

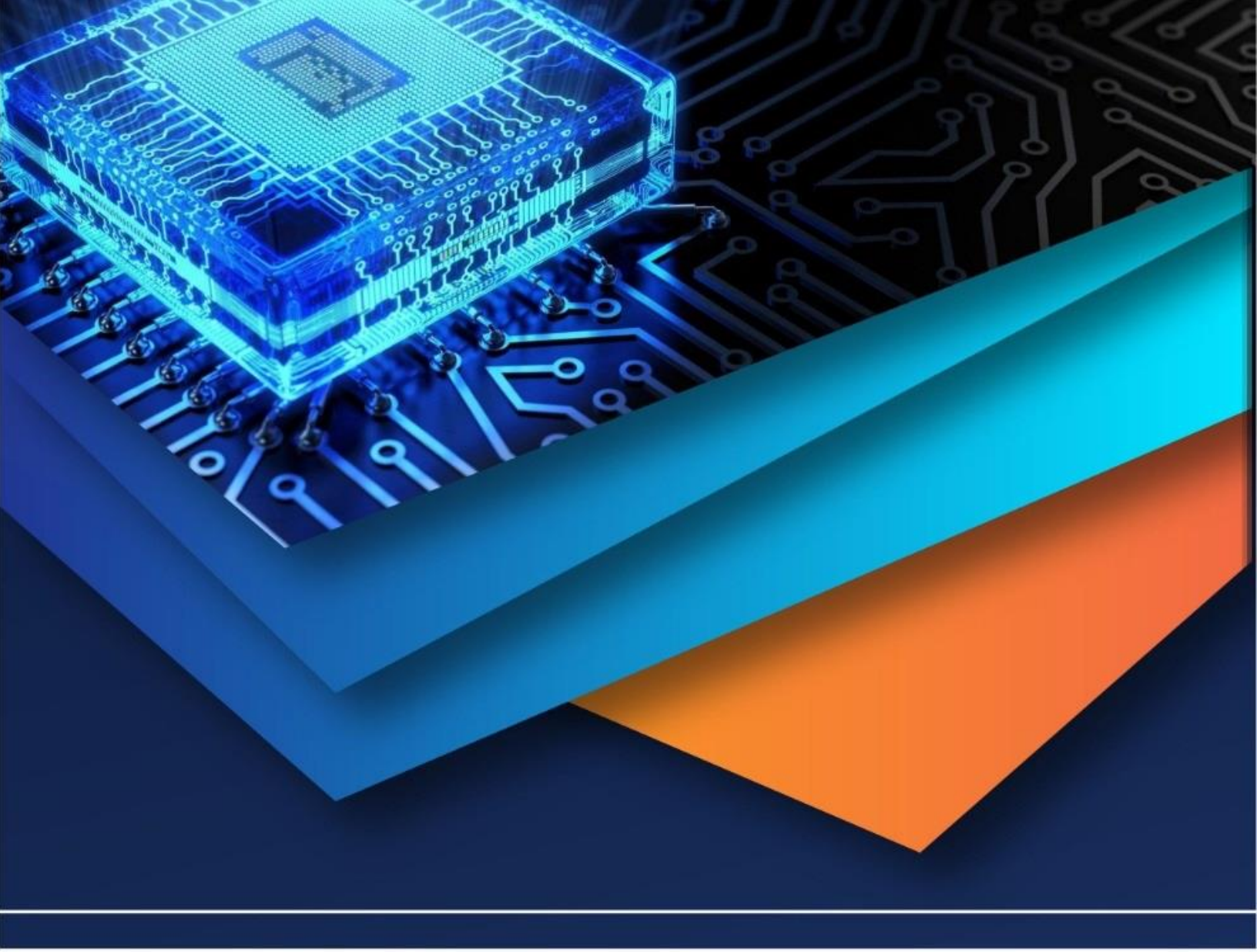
VI. LIMITATIONS AND FUTURE SCOPE

The system may not be able to recognize objects that are too small or too far away from the camera, partially hidden or obstructed, too similar in shape or color, or objects in low light conditions. The future enhancements of the proposed method may also focus on the development of a system-on-chip (SoC). The size, weight, and cost of the system can be reduced. However, it should be noted that the proposed system will have many such limitations. There will always be some new objects that may not be recognized due to their difference in size, distance, or environmental factors. To address these limitations, we could also expand the dataset to include more object categories, and use more advanced deep learning models and hardware to further improve the system's performance.

REFERENCES

[1] IEEE Conference Publication, “Deep Learning based Object Detection and Recognition Framework for the Visually-Impaired,” IEEE Xplore, 2019.
[2] IEEE Conference Publication, “Object Recognition System for Visually Impaired People,” IEEE Xplore, 2018.
[3] CVPR 2021 Workshop on Media and Automotive Intelligence, “Computer Vision-Based Assistance System for the Visually Impaired Using Mobile Edge AI Accelerator Devices,” 2021.
[4] Academia.edu, “Virtual Fitness Trainer using Artificial Intelligence,” 2019..
[5] "Object Detection in 2023: The Definitive Guide - viso.ai." https://viso.ai/deep-learning/object-detection/. Accessed 2 Apr. 2023.
[6] Bochkovskiy, A., Wang, C.Y., & Liao, H.Y.M. (2021). Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2107.10972
[7] Singh, A., & Davis, L.S. (2021). A Survey of Modern Deep Learning based Object Detection Models. arXiv preprint arXiv:2104.11892.
[8] *Li, X., Chen, Y., & Zhang, Y. (2019). A Real-Time Objects Recognition Approach for Assisting Blind People. IEEE Access, 7, 159013-159021. doi: 10.1109/access.2019.2942813.*
[9] Zhang, Y., Wang, S., & Yang, S. (2018). A Real-Time Objects Recognition Approach for Assisting Blind People based on Deep Learning. Journal of Image and Graphics, 6(4), 289-294. doi: 10.11648/j.ijg.20180604.18.
[10] Rao, R., & Saxena, A. (2017). Real-Time Objects Recognition for Assisting Blind People using Raspberry Pi. International Journal of Advanced Research in Computer Science, 8(3), 375-378. doi: 10.26483/ijarcs.v8i3.3914.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)