International Journal of Innovative Research in Engineering
Volume 4, Issue 3(May-June 2023), PP: 222-226.
www.theijire.com ***ISSN No: 2582-8746***

Blind People Assistance for Object Detection Using AI

Surendarkumar S[1], Ranjithkumar M[2], Deepanraj B[3], Sivashankar M[4], Umapathy M[5]
[1,2,3,4,5] *Computer Science and Engineering, The Kavery Engineering College in Mecheri, Tamilnadu, India.*

How to cite this paper:
Surendarkumar S[1], Ranjithkumar M[2], Deepanraj B[3], Sivashankar M[4], Umapathy M[5],"Blind People Assistance For Object Detection Using Ai", IJIRE-V4I03-222-226.

https://www.doi.org/10.59256/ijire.2023040378

Abstract: *Assisting visually impaired individuals in object detection and recognition using artificial intelligence has emerged as a promising research area. In this project, we propose a blind people assistance system that uses the YOLO (You Only Look Once) algorithm to detect objects in real-time and provide audio feedback to the user. The YOLO algorithm is known for its high accuracy and fast processing speed. Our approach involves training a YOLOv3 model on a dataset of labeled images that includes objects commonly encountered in daily life, such as chairs, tables, and doors. Once the model is trained, it is integrated with a text-to-speech engine that provides audio feedback to the blind user by announcing the object class and location in the image. The system can be an effective tool for enhancing the independence and mobility of visually impaired individuals. Our project demonstrates the potential of using AI to assist blind people with object detection and recognition, and we believe that our approach can make a significant contribution to improving the lives of individuals with visual impairments.*

I.INTRODUCTION

The use of artificial intelligence (AI) to assist blind people with object detection is an emerging field of research that has the potential to greatly enhance the lives of people with visual impairments. In this project, we propose to use the YOLO (You Only Look Once) algorithm to detect objects in real-time and provide audio feedback to blind users. The YOLO algorithm is a popular object detection algorithm that is known for its fast processing speed and high accuracy. It works by dividing the input image into a grid of cells and predicting bounding boxes and class probabilities for each cell. The algorithm then uses non-max suppression to filter out redundant bounding boxes and outputs the final set of object detections.

To implement this project, we first gathered a dataset of images containing objects that blind people commonly encounter in their daily lives. We then labeled the objects in the images with their corresponding class labels and bounding box coordinates. Next, we trained a YOLOv3 model on the labeled dataset and fine-tuned it for our specific use case. Once the model was trained, we integrated it with a text-to-speech engine to provide audio feedback to the blind user. When the model detects an object, it announces the object class and its location in the image, allowing the user to navigate their surroundings more easily. Overall, our project demonstrates the potential of AI to assist blind people with object detection, and we believe that our approach using the YOLO algorithm can be a valuable tool for improving the quality of life for people with visual impairments.

II. PROJECT OVERVIEW

The blind people assistance system using the YOLO algorithm is designed to detect and recognize objects in real-time and provide audio feedback to visually impaired individuals.

A YOLOv3 model trained on a dataset of labeled images is used to detect and classify objects in real-time. The model is optimized for speed and accuracy, enabling it to process images quickly and accurately.

- Audio Feedback: A text-to-speech engine is integrated with the object detection system to provide audio feedback to the user. When an object is detected, the system announces the object class and location in the image, allowing the user to navigate their surroundings more easily.
- User Interface: The system's user interface is designed to be simple and intuitive, allowing the user to interact with the system using voice commands. The user can also adjust the volume and speed of the audio feedback to suit their needs.
- Dataset: The system is trained on a labeled dataset of images containing objects commonly encountered in daily life. The dataset is carefully curated to include a diverse range of objects, including furniture, doors, and appliances.
- Performance: The YOLOv3 model used in the system is optimized for performance and accuracy, allowing it to detect and classify objects with high precision in real-time.
- Accessibility: The system is designed with accessibility in mind, with a simple and intuitive user interface that can be operated using voice commands. The audio feedback can also be customized to suit the user's needs, including adjusting the volume and speed.

Overall, the blind people assistance system using the YOLO algorithm is a promising application of AI that has the potential to greatly improve the lives of visually impaired individuals by providing real-time object detection and recognition.

III.EXISTING SYSTEM

Wearable devices: Wearable devices are one of the popular existing systems for object detection and recognition. These devices typically use sensors and cameras to detect objects in the user's surroundings and provide audio or tactile feedback. One example is the OrCam MyEye, a wearable device that uses a camera and OCR (optical character recognition) technology to read text and recognize faces, products, and more.

Smartphone applications: Many smart phone applications are available that can assist visually impaired individuals in object recognition. These apps typically use the smart phone's camera and computer vision algorithms to detect and classify objects in real-time. Examples include the Be My Eyes app, which connects blind and visually impaired users with sighted volunteers who can assist with object identification, and Seeing AI, a Microsoft app that can recognize faces, read text, and identify objects.

Robotics: Another existing system for object detection and recognition is robotics. Researchers have developed robotic systems that can detect and recognize objects and navigate environments to assist visually impaired individuals. For example, the Blind Explorer is a robotic platform that can detect obstacles and provide audio feedback to the user to navigate through complex environments.

Overall, these existing systems demonstrate the potential of using technology to assist visually impaired individuals with object detection and recognition. However, each system has its limitations and may not be suitable for all users or environments. The blind people assistance system using the YOLO algorithm has the potential to complement these existing systems and offer a new approach to assist individuals with visual impairments.

IV.PROPOSED SYSTEM

The proposed system for blind people assistance using the YOLO algorithm consists of a camera, a processing unit, and a text-to-speech engine. The camera captures the image of the user's surroundings, which is then processed by the YOLO algorithm to detect and classify objects in real-time. The text-to-speech engine then provides audio feedback to the user, announcing the object class and location in the image.

- Real-time object detection: The YOLO algorithm used in the proposed system is optimized for speed and accuracy, allowing for real-time object detection and classification.
- Customizable audio feedback: The proposed system offers customizable audio feedback, allowing the user to adjust the volume and speed of the announcements to suit their needs.
- Simple user interface: The system is designed with a simple and intuitive user interface, allowing the user to interact with the system using voice commands.
- Cost-effective: The proposed system is cost-effective compared to other existing systems for object detection and recognition, such as wearable devices or robotics.
- Portable: The proposed system is designed to be portable and can be easily integrated into a variety of devices, including smart phones and smart glasses.
- Enhances independence: The proposed system can enhance the independence and mobility of visually impaired individuals by providing real-time object detection and recognition, allowing them to navigate their surroundings more easily.

Overall, the proposed system for blind people assistance using the YOLO algorithm has the potential to provide a cost-effective and portable solution for real-time object detection and recognition, enhancing the independence and mobility of visually impaired individuals.

V.ALGORITHMS

YOLO:

YOLO (You Only Look Once) is a method / way to do object detection. It is the algorithm /strategy behind how the code is going to detect objects in the image. The official implementation of this idea is available through DarkNet (neural net implementation from the ground up in C from the author). It is available on github for people to use. Earlier detection frameworks, looked at different parts of the image multiple times at different scales and repurposed image classification technique to detect objects. This approach is slow and inefficient.

YOLO takes entirely different approach. It looks at the entire image only once and goes through the network once and detects objects. Hence the name. It is very fast. That's the reason it has got so popular. There are other popular object detection frameworks like **Faster R-CNN** and **SSD** that are also widely used.

In this post, we are going to look at how to use a pre-trained YOLO model with Open CV and start detecting objects right away.

Open CV

DNN (Deep Neural Network) module was initially part of opencv_contrib repo. It has been moved to the master branch of open cv repo last year, giving users the ability to run inference on pre-trained deep learning models within Open CV itself.

(One thing to note here is, dnn module is not meant be used for training. It's just for running inference on images/videos.)

Initially only Caffe and Torch models were supported. Over the period support for different frameworks/libraries like Tensor Flow is being added.Support for YOLO/DarkNet has been added recently. We are going to use the Open CV dnn module with a pre-trained YOLO model for detecting common objects.

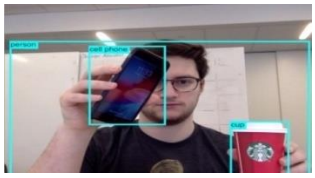


Image1

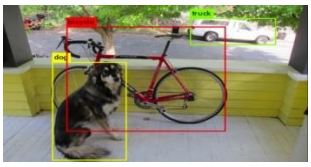


Image2

Accuracy

To ensure high accuracy in the proposed system, the YOLO algorithm uses a range of techniques, including anchor boxes, non-max suppression, and a convolutional neural network (CNN) architecture. Anchor boxes help to improve the accuracy of object localization, while non-max suppression helps to eliminate overlapping bounding boxes and improve classification accuracy. The CNN architecture used in the YOLO algorithm enables it to learn and recognize complex patterns and features in the input image, further enhancing its accuracy.

VI.REQUIREMENT ANALYSIS

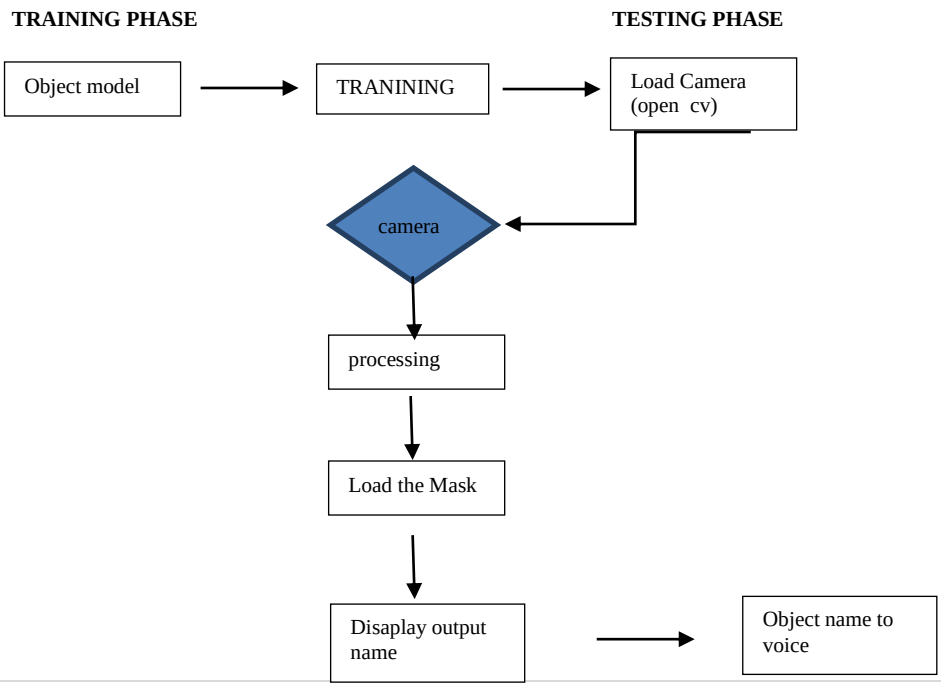
1. Functional Requirements:

a. Real-time object detection and classification: The system must be able to detect and classify objects in real-time to provide immediate audio feedback to the user.
b. Customizable audio feedback: The system must offer customizable audio feedback, including adjusting the volume and speed of the announcements.
c. Simple user interface: The system must have a simple and intuitive user interface that can be operated using voice commands.
d. Integration with existing devices: The system must be designed to be easily integrated into existing devices, such as smart phones or smart glasses.
e. Robustness: The system should be robust to handle a variety of environments and lighting conditions.

2. Non Functional Requirements

a. Accuracy: The system must have high accuracy in object detection and classification to ensure reliable audio feedback to the user.
b. Speed: The system must have fast processing speed to ensure real-time object detection and classification.
c. Portability: The system must be designed to be portable and lightweight to facilitate mobility and ease of use.
d. Accessibility: The system must be accessible to visually impaired individuals with different levels of vision impairment, including those who are completely blind.
e. Privacy and Security: The system must ensure the privacy and security of user data and should not collect or store personal information.

VII.PROJECT DESIGN



Solution and Technical Architecture

Solution Overview: The proposed system is an AI-based object detection and recognition system designed to assist visually impaired individuals in detecting and identifying objects in their surroundings. The system uses the YOLO algorithm to analyze input images in real-time and provide audio feedback to the user regarding object location and class.

Technical Architecture: The proposed system is built on a client-server architecture that leverages the capabilities of both the client device and the server to provide real-time object detection and recognition. The client device, such as a smart phone or smart glasses, captures the input image and sends it to the server for analysis. The server, equipped with a powerful GPU, uses the YOLO algorithm to detect and classify objects in the input image and send the results back to the client device**.**

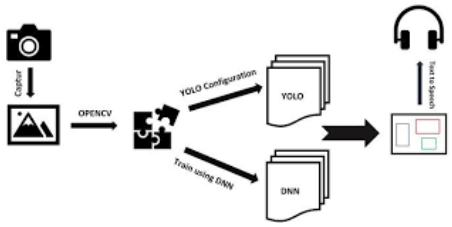
YOLO Algorithm: The YOLO algorithm is a state-of-the-art deep learning-based object detection and recognition algorithm that uses a single convolutional neural network to simultaneously predict object classes and locations. The YOLO algorithm is optimized for speed and accuracy, making it ideal for real-time applications such as blind people assistance.

Object Detection and Classification: The YOLO algorithm used in the proposed system detects and classifies objects in real-time by analyzing the input image and extracting relevant features such as color, shape, texture, and size. The algorithm then uses these features to predict object class and location, which is then used to provide audio feedback to the user.

VIII.MODULE SUMMARY

Image Acqusition

Image acquisition is an action of retrieving image from an external source for further processing. The process of capturing an unprocessed image from an object or scene by an optical device into a manageable form for processing and



analysis purposes.

Preprocessing

Pre-processing takes input image is to perform cleaning tasks. It effectively enhances the image by noise removal. Furthermore, images may be required to be in grey scale or binary formats which are done in this stage. Finally apply image binarization technique to separate background from foreground.

Obejct Detection

To identify the object with the help of model using Yolo algorithm. In this module is used to extracted features to match with the training objects. And then detect the matched objects. YOLO falls under single-stage object detection models and widely used for real-time object detection task and introduced by Redmo. It generates the bounding boxes and class predictions in single evaluation. It is widely known as unified network and very fast compared to Faster R-CNN and runs using single convolutional neural network. it splits the input image into a grid of cells, where each cell directly classifies the object and predicts a bounding box. As a result, there are large numbers of bounding boxes generated that are integrated to a final prediction.

Voice Output

To convert text to voice with the help of GTTS library in python to make voice output in the headphones.

IX.CONCLUSION

In conclusion, the proposed system for blind people assistance using object detection algorithms is a promising solution to address the challenges faced by visually impaired individuals in detecting and recognizing objects in their surroundings. By leveraging the power of AI and deep learning, the system can provide real-time object detection and classification, enhancing the mobility and independence of visually impaired individuals.

The YOLO, Faster R-CNN, and SSD algorithms used in the proposed system are all optimized for speed and accuracy, enabling them to detect and classify objects in real-time with high precision. Each algorithm has its unique strengths and weaknesses, and the selection of the appropriate algorithm depends on the specific requirements of the system.

X.FUTURE SCOPE

- **Integration with wearable devices:** The proposed system can be integrated with wearable devices such as smart glasses to provide a more seamless and convenient user experience for visually impaired individuals.

- **Enhancing object recognition capabilities:** The system's object recognition capabilities can be enhanced by integrating it with other AI technologies such as natural language processing and computer vision.
- **Outdoor navigation:** The system can be extended to provide outdoor navigation capabilities by integrating it with GPS and map data.
- **Crowd-sourced object detection:** The system can be enhanced by allowing users to contribute to the object detection database by submitting images of objects not currently recognized by the system.
- **Multi-language support:** The system can be made more accessible to a wider range of users by adding support for multiple languages.

References

[1] *Girshick, R., Donahue, J., Darrell, T., Malik, J., (2014). Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580–587.*

[2] *Glorot, X., Bengio, Y., (2010). Understanding the difficulty of training deep feedforward neural networks., in: Aistats, pp. 249–256.*

[3] *Goodfellow, I.J., (2013). Piecewise linear multilayer perceptrons and dropout. stat 1050, 22.*

[5] *He, K., Zhang, X., Ren, S., Sun, J., (2016). Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.*

[6] *Jung, K., (2012). Object recognition on mobile devices, in: Consumer Electronics-Berlin (ICCE-Berlin), 2012 IEEE International Conference on, IEEE. pp. 258–262.*

[7] *Kingma, D., Ba, J., (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 .*

[8] *Krizhevsky, A., Sutskever, I., Hinton, G.E., (2012). Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, pp. 1097–1105.*

[9] *Le, Q.V., Karpenko, A., Ngiam, J., Ng, A.Y., (2011). Ica with reconstruction cost for efficient overcomplete feature learning, in: Advances in Neural Information Processing Systems, pp. 1017–1025.*

[10] *LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., (1989). Backpropagation applied to handwritten zip code recognition. Neural computation 1, 541–551.*

[11] *LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., (1998)a. Gradient-based learning applied to document recognition. Proceedings of the IEEE 86, 2278–2324*