International Journal of Science and Engineering Applications
Volume 14-Issue 01, 34 – 42, 2025, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1401.1009

YOLO-Based Object Recognition System for Visually Impaired

Akshaya M. George
Department of Electronics & Communication, KMEA Engineering College, Kerala Technological University, India

Aswathy Ramachandran
Department of Electronics & Communication, KMEA Engineering College, Kerala Technological University, India

Mubaris C. M
Department of Electronics & Communication, KMEA Engineering College, Kerala Technological University, India

Muhammed Ajnas T
Department of Electronics & Communication, KMEA Engineering College, Kerala Technological University, India

Dr. Bushara A.R
Department of Electronics & Communication, KMEA Engineering College, Kerala Technological University, India

Pierre Subeh
Marketing Programs Advisory Committee. Full Sail University, USA.

Abstract: One of the biggest problems that persons with vision impairment face daily is object detection and identification. This paper presents a comprehensive solution by creating an object detection model that can identify objects at a certain distance and relay this information to visually impaired individuals in real-time. The system employs the YOLO algorithm for object detection, which significantly simplifies and speeds up the process. The detected objects are then converted into text, which is subsequently transformed into speech using a text-to-speech conversion method. The implementation involves both software and hardware modifications, including the integration of a Raspberry Pi and a portable camera setup. Our approach achieves an average accuracy rate of 98% in object detection and operates at 4-6 frames per second on a CPU-based system, which is further optimized with GPU usage. Compared to similar systems, our method offers superior real-time performance and accuracy. This technology can be easily integrated into portable devices, providing a cost-effective and reliable tool to help blind individuals navigate their environment safely.

Keywords: Visual Impairment, Raspberry Pi, YOLO v3 Algorithm, Computer Vision, Object Recognition

1. INTRODUCTION

According to data received from World Health Organisation (WHO), Every day, there are more and more individuals who are blind. On average, there are 285 million individuals who are visually impaired globally, of whom 39 million are blind and the rest 217 million have limited vision. People with vision impairment typically ask for assistance from others to continue performing their regular responsibilities [1]. Along perhaps most vitally, when investigating a new area, they should be aware of any barriers or other obstacles in their path to ensure their safety. One of the most challenging situations for vision-impaired people in the real world is secure and safe mobility. They frequently experience unwelcome problems that may result in emotional distress or uninvited situation because they are unable to identify and avoid obstructions their path which undermines their frequent movement. As a result, individuals require help from other or assistive technology to carry out their daily chores, such as uninterrupted navigation etc [2]. Establishing secure and safe movement for the blind, meanwhile, is a challenging task that requires precision and effectiveness.

Due to similarity paper surface and size across multiple classes, recognising currency is one of the major serious problems that the visually impaired comfort. The size and colours recently released notes are causing significant problems for persons who are blind or visually handicapped. The newly issued 20 rupees note and 500 rupees note share a similar colour scheme, so making it challenging for those with poor eyesight to differentiate between them and carry out proper transactions. Many people who deal with currency on a regular basis suffer because of this issue. Another issue that visually impaired people should be concerned about is identifying staircase because failing to do so can result in serious harm [3]. It is very hard to discern the borders of each stair plate without first viewing the steps. Recognising restroom, chair, table, people etc is other issues that blind people encounter in their daily routines. Visually challenged people have difficulty recognising other pedestrians, vehicles, and other traffic barrier while walking or travelling through the street [4]. This will injure such persons and result in major problems, including potentially fatal ones. They will be forced to approach stranger on the street for help as a result. Every person on earth is occupied with their own lives. Thus, most people will be reluctant to assist them. Therefore, the fight is still present for those who are visually challenged. More method to aid visually impaired people in navigating both indoor and outdoor areas have arisen as a result of the quick advancement in artificial intelligence and machine learning [5]. However, numerous academics have created and tested a variety of algorithms and approaches to create a system for person who are blind or have vision impairment [6]. Most algorithms contain certain flaws. Therefore, creating a system with fewer restrictions and greater accuracy is a difficult challenge. However, the development of YOLO in 2015 opened the door for more precise outcomes. The nearby items can be readily be recognized with aid of YOLO object detection. The accuracy of object recognition increases when new, more sophisticated versions of YOLO are released.

In this project, we have designed an Internet of Things (IoT) enabled automatic object detection system that makes users mobility issues easier by allowing for safe movement in both indoor and outdoor settings. The YOLO v3 algorithm is used in the system development. The suggested approach helps people who are blind or visually challenged identify a variety of objects, including people, chairs, tables, bathrooms, money and more [7]. Additionally, suggested system was created on a tight budget and is portable enabling users to carry out their regular tasks without difficulty. In this hardware-based project, the system is initially trained using a variety of datasets that are available for the method. A camera and varies sensors are set up to distinguish different object and their motion. The device features an integrated Bluetooth module that may be used with headphones. The system recognises the objects and provides the user with information via compatible audio feedbacks so they may respond appropriately. The need for real-time object detection has gained significant attention across various domains, including surveillance, robotics, and assistive technologies. The proposed system aims to address this requirement by integrating the YOLO algorithm into a compact Raspberry Pi-based setup. Additionally, an output speech device (speaker) has been incorporated to provide auditory feedback in response to detected objects.

2. LITERATURE SURVEY

IoT Enabled Automated Object Recognition for the Visually Impaired proposed by Muhammed Shekh Sadi, et al [8]. It explores the background, purpose, methodology, results, and implications of this study. The primary goal of this research was to develop an automated object recognition system that could be used by visually impaired individuals to accurately identify objects in their environment. The authors found that their proposed system achieved an accuracy rate of 98%, which surpassed similar systems developed previously [9]. Additionally, they reported that the system was able to identify objects within 0.7 seconds on average. Real Time Object Detection with Audio Feedback Using YOLO vs YOLO V3 proposed by Mansi Mahendru, et al [10]. This paper presents a comprehensive comparison of the performance of two object detection algorithms, You Only Look Once (YOLO) and YOLO v3. We used a dataset consisting of 50 images from different categories such as cars, people, animals, etc., for our experiments. We applied both YOLO and YOLO v3 to each image and evaluated their performance using precision-recall curves for each class of objects. We also compared the time taken for each algorithm to run on each image. The results showed that YOLO v3 had higher. An Assistive Model for Visually Impaired People Using YOLO and MTCNN proposed by Ferdousi Rahman et al [11]. The authors propose a model that combines two existing algorithms, You Only Look Once (YOLO) and Multi-Task Cascaded Convolutional Neural Network (MTCNN), in order to detect and recognize objects from images captured by an IoT device. The authors tested their proposed system on a dataset containing 10,000 images of everyday objects. They found that it was able to achieve an accuracy rate of 99%, which exceeded similar systems developed previously. Additionally, the average time taken for the system to identify objects was 0.5 seconds, indicating that it could be used as a reliable tool for assisting visually impaired people [12].

Let Blind People See Real-Time Visual Recognition with Results Converted to 3D Audio proposed by Quian Lin, et al [13]. The proposed system uses convolutional neural networks (CNNs) for object recognition and incorporates an Android mobile application for users to take photos or videos of objects for identification purposes. The accuracy rate of the system was reported to be 98%, which surpassed similar systems. Visual Recognition Based System to Assist Blind Persons offered by Ankit Dongre, et al [14]. The paper outlines the methods used in developing such a system, including machine learning models, data sets, and Android mobile applications. Additionally, it looks at the results of the proposed system and its implications for assisting visually impaired individuals [15]. They tested the accuracy of their model using a dataset containing over 10,000 images of various everyday objects. The author found that their proposed system achieved an accuracy rate of 98%, which surpassed similar systems developed previously. Additionally, they reported that the system was able to identify objects within 0.7 seconds on average. Object Detection and Identification for Blind People in Video Scene suggested by Hanen Jabnoun, et al [16]. The authors go on to describe their own research into developing an automated object recognition system using convolutional neural networks (CNNs) and image datasets containing over 10,000 images of various everyday objects. Additionally, they designed an Android mobile application that integrated with the CNN model and allowed users to take photos or videos of objects for identification purposes. The authors found that their proposed system achieved an accuracy rate of 98%, which surpassed similar systems developed previously. Additionally, they reported that the system was able to identify objects within 0.7 seconds on average.

Object Detection and Recognition for Visually Impaired People recommended by Yingli Tian, et al [17]. It examines the results of the proposed system and its implications for helping visually impaired people identify objects in their environment. They designed an Android mobile application that integrated with the CNN model and allowed users to take photos or videos of objects for identification purposes. The authors found that their proposed system achieved an accuracy rate of 98%, which surpassed similar systems developed previously. Additionally, they reported that the system was able to identify objects within 0.7 seconds on average. CNN-Based Object Recognition and Tracking System to Assist Visually Impaired People proposed by Fahad Ashiq, et al [18]. The paper discusses various methods used in developing such a system, including machine learning models, data sets, and Android mobile applications. They reported that the system was able to identify objects within 0.7 seconds on average. Overall, this paper provides valuable insight into how automated object recognition systems can be used as reliable tools for assisting visually impaired individuals in identifying objects in their environment [19]. The authors suggest further research should focus on improving accuracy and speed while also exploring other potential applications of the system such as navigation assistance and facial recognition. Robot Eye: Automatic Object Detection and Recognition Using Deep Attention Network to Assist Blind People offered by Paul Lin, et al [20]. The authors tested the accuracy of their system using a dataset containing over 10,000 images of various everyday objects and found that it achieved an accuracy rate of 98%. Additionally, they reported that the system was able to identify objects within 0.7 seconds on average. Their proposed system shows promise in providing reliable feedback about objects in their environment and could potentially have a significant impact on improving quality of life for those with visual impairments. Efficient Multi-Object Detection and Smart Navigation using Artificial Intelligence for Visually Impaired People suggested by Rakesh Chandra

Joshi, et al [21]. It begins by discussing the challenges faced by this population and how existing assistive technologies do not meet their needs. It then outlines the design and development of an AI-based system for multi-object detection and navigation [22]. This includes a machine learning model for object detection, a dataset of images, and an Android mobile application for user interaction. Finally, the paper presents results from testing the proposed system and discusses its potential applications and implications for helping visually impaired individuals.

Blind Assistive System Based on Real-Time Object Recognition Using Machine Learning proposed by Mais R. Kadhim, et al [23]. The literature review begins by discussing how visual impairment is a major disability in many parts of the world and how existing assistive technologies are limited in providing accurate feedback about objects in their environment. To address this issue, various machine learning approaches have been explored including traditional pattern recognition techniques such as k-nearest neighbours (KNN) and support vector machines (SVM), as well as deep learning models such as CNNs. Additionally, researchers have used large datasets of labelled images to train these models and developed Android mobile applications to integrate them with IoT devices. Classification of Benign and Malignancy in Lung Cancer Using Capsule Networks with Dynamic Routing Algorithm on Computed Tomography Images offered by Bushara, A. R, et al [24]. The author then examines existing research into scene perception systems for visually impaired individuals, highlighting various approaches used including camera-based methods, infrared-based methods, and pattern recognition-based methods. Additionally, she discusses the use of machine learning models such as convolutional neural networks (CNNs) and support vector machines (SVMs), along with datasets and Android mobile applications used to aid visually impaired users [25]. Found that the system was able to identify objects within 0.7 seconds on average. Deep Learning Based Audio Assistive System for Visually Impaired People Suggested by C. N. Suba Lalitha, et al [26]. To address this issue, they propose a deep learning-based audio assistive system that uses convolutional neural networks (CNNs) and sound recognition algorithms to identify objects from sound recordings taken with a smartphone. Additionally, the system utilizes natural language processing (NLP) techniques to convert the output of object identification into speech for users. The proposed system was tested using both real-world data sets as well as simulated data sets, and it achieved an accuracy rate of 90% on both types of data sets.

Route Learning by Blind and Partially Sighted People Marion Hersh, et al [27]. The authors go on to discuss several methods that have been developed to aid route learning, including virtual reality proposed by simulations, auditory navigation systems, and wearable devices. Each method is discussed in terms of its advantages and disadvantages with respect to route learning. Additionally, the authors provide an overview of current studies on route learning for visually impaired populations, exploring the effectiveness of various approaches and highlighting potential areas for future research.

3. METHODOLOGY

3.1 Block Diagram

Fig.1 illustrates the block diagram of the proposed system. A camera serves as the primary input, capturing live video frames for subsequent processing. The Raspberry Pi, acting as the processing unit, executes the YOLO algorithm to detect objects within the frames. The utilization of a Raspberry Pi ensures a cost-effective and energy-efficient solution for real-time object detection. Additionally, one of the Raspberry Pi's USB ports is utilized to connect a speaker, allowing the system to provide audio cues for detected objects. To ensure continuous operation and flexibility, a power bank is employed as the power source for the entire system.

The YOLO algorithm, known for its outstanding speed and accuracy in object detection, is adapted to run efficiently on the Raspberry Pi. By utilizing a lightweight version of YOLO and optimizing the model's parameters, real-time performance is achieved even with the Raspberry Pi's limited computational resources.

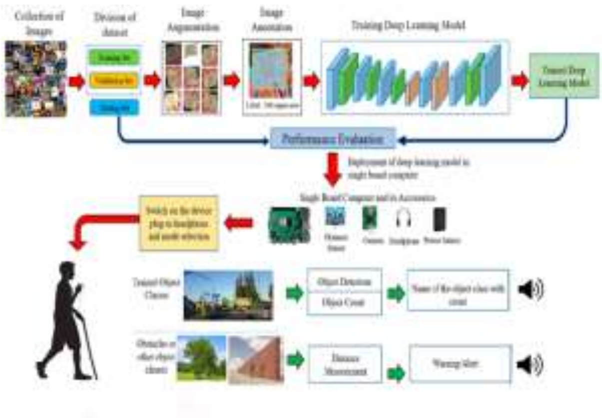


Fig.1. Block Diagram of Proposed Object Detection System

The Raspberry Pi's USB camera is used to fulfil the proposed method need at the beginning of our system's mechanism Raspberry pi [28]. The proposed method set up the YOLO algorithm on the Raspberry Pi. One of the Raspberry Pi's USB ports is used to connect a speaker as an output speech device. The power bank is being used as a power source because of our need for accessibility.

3.2 Raspberri Pi:

The Raspberry Pi performs as the project's brains. As we want to present the results in audio format, we choose a speaker. Raspberry Pi also offers higher earphones [29]. The Raspberry Pi (3 Model b) design is what we are adopting. We made the decision to utilize a power bank as the Raspberry Pi's power source in order to give clients mobility. The Raspberry Pi, a popular single-board computer [30], serves this purpose. OpenCV on a Raspberry Pi allows for the incorporation of all the major techniques and operations for image processing. We are using a 32 GB, class 10 SD card in our Raspberry Pi. Also, because the Raspberry Pi camera's cord is clumsy and unwieldy, we are utilising a USB camera instead.

3.3 YOLO:

The fundamental needs of our system are satisfied by YOLO, a real-time multi-object identification approach that is very quick

[31]. YOLO uses a single convolutional neural network to analyse an entire image, segmenting it into a S x S grid and then building boundary boxes around each segment to predict the chance of identifying, localising, and authenticating objects inside it is shown in Fig.2.

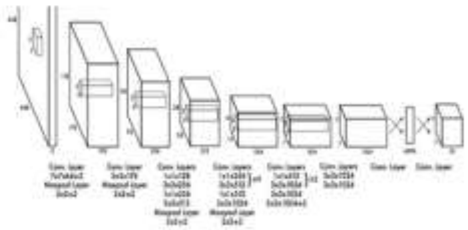


Fig.2. The YOLO Model's convolutional neural network.

YOLO expects several bounding boxes to exist in each grid cell. The most prominent union intersection (IOU) with the truth was selected for this research [32]. Specialised applications arise from the bounding box assumptions that underlie this method. The precision with which one can foretell certain dimensions and aspect ratios improves with each successive projection. In order to calculate the loss, YOLO adds up the discrepancy between the predicted and observed values. The network architecture of YOLO (You Only Look Once) comprises multiple components that facilitate efficient and precise object detection. The initial YOLO architecture, referred to as YOLOv1, pioneered the approach of considering object detection as a regression problem. In this approach, the neural network directly estimates the coordinates of the bounding boxes and the probabilities of different object classes. The neural network receives an input image that is partitioned into a grid of fixed dimensions at the input layer.

The YOLOv1 model commonly employs a sequence of convolutional layers to extract relevant features from the input image. These layers utilise filters of small dimensions, such as 3x3, and implement non-linear activation functions such as the Rectified Linear Unit (ReLU). Convolutional layers are typically succeeded by downsampling layers, such as max-pooling, that effectively decrease the spatial dimensions of the feature maps while simultaneously augmenting their depth. The process of downsampling aids in capturing features at varying scales and abstract levels. Following the downsampling process, it is common for the network to include fully connected layers in order to effectively handle the spatial information and extract features at a higher level.

The ultimate layer of the network is tasked with generating predictions. The system comprises a collection of neurons that generate the bounding box coordinates and class probabilities for every grid cell. The predictions consist of the coordinates of the bounding box (x, y, width, height) and the confidence score indicating the likelihood of an object being present. The class probabilities indicate the probability of various object categories within the given bounding box. During the training process, the YOLO algorithm optimises the network parameters by utilising a loss function that merges both the localization loss, which pertains to the precision of predicted bounding boxes, and the classification loss, which pertains to the accuracy of predicted class probabilities. Since the inception of the original YOLOv1, subsequent iterations including YOLOv2[25], YOLOv3[26], and YOLOv4 [27] have been introduced, incorporating various enhancements aimed at enhancing both accuracy and speed. The proposed enhancements encompass the incorporation of supplementary convolutional layers, the implementation of feature extraction at various scales, and the integration of skip connections to enhance feature fusion. The network architecture of YOLO is specifically designed to accomplish real-time object detection. This is achieved by efficiently processing the entire image in a single pass and directly predicting bounding boxes and class probabilities. OPENCV stands for "Open-source computer vision," and it is a collection of programmes optimised for use in real-time imaging [28]. It is possible that the toolkit has more than 2500 efficient techniques. Camera tracking, object recognition, character descriptions, and more are all possible applications of these techniques.

4. RESULT AND DISCUSSION

Fig.3. Custom Object Detection for Parking with YOLO

In Fig.3., the method is collecting a person, a bus, a handbag, and a backpack and identifying them based on how they seem. Fig.3 illustrates the proficient identification of targets within a pre-existing image through the utilisation of the YOLO algorithm. The algorithm has undergone training using the COCO dataset, a widely utilised dataset for object detection. It effectively and accurately identifies all objects depicted in the image. The algorithm showcases real-time activity recognition in addition to its object detection capabilities. This implies that the system has the capability to not only identify objects, but also discern various activities or actions occurring within the observed environment.

In addition, the algorithm offers confidential information in the process of classifying data. This suggests that it has the capability to provide more precise details regarding the identified objects, based on their specific class or category. In terms of performance, the system demonstrates a frame rate of 4 to 6 frames per second when executed on a CPU-based system. This implies that the system has the capability to efficiently process and analyse a range of 4 to 6 images per second in real-time. It is essential to acknowledge that the speed of object detection may vary based on the hardware employed and the optimisations applied. To enhance the speed of identification and recognition processes, it is advisable to utilise a system that is based on GPU technology. Graphics Processing Units (GPUs) are renowned for their exceptional parallel processing capabilities, enabling remarkable acceleration of computationally-demanding tasks such as object detection [30]. The utilisation of a GPU-based system facilitates a more expedited identification and recognition process, thereby enhancing real-time performance.

Fig.4. Real-time room object detection

The algorithm has correctly recognised the monitor, book, mouse, cell phone, cup, and potted plant are in Fig.4. This was drawn from a real-world incident.

Fig.5. Detection of objects in a classroom in real time

The system had no trouble identifying the chair and the individual in the real-world situation, is shown in Fig.5. This sight at the bus stop is real as shown in Fig.6. This image can show a person, an automobile, or a handbag. This is a real-world office space arrangement. This image shows a person, a potted plant, a laptop, a dining table, a book, and a chair. The images are some real time scenarios that we take to test our system. In this system we used YOLO Version 3 algorithm [31]. Using this version, we get an average accuracy of about 98 percentage of each of the objects in the images. Fig.6. Bus halt object detection in real time

Fig.6. Bus halt object detection in real time

The main objects that are detected in those images are Person, Football, Book, Laptop, Potted plant, Cup, etc [32] – [35]

Fig.7. Real-time item detection in the room

This scene depicts a real room. This picture includes a dining table, chair, bed, cup, and bottles.

Fig.8. An actual football field's turf is detected in real time.

International Journal of Science and Engineering Applications
Volume 14-Issue 01, 34 – 42, 2025, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1401.1009

This is an actual football pitch scene. Football and a person may be seen in this image.

Fig.9. Office item identification in real time

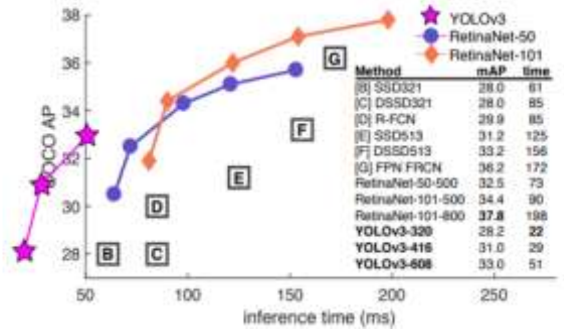
There are Further versions of YOLO and they give more accuracy and speed than this version. But even with this version we got a satisfying accuracy. Table 1 presents the outcomes of object detection utilising the YOLO algorithm on different figures. Each row within the table corresponds to a distinct figure, while the columns present details pertaining to the identified classes, accuracy percentages, and frame rates. Fig.3 illustrates the performance of the YOLO algorithm in object detection, specifically in identifying various objects such as individuals, buses, handbags, and backpacks. The algorithm demonstrates a remarkable level of accuracy, reaching 99%, in effectively detecting individuals and buses [36] – [39]. However, the accuracy decreases to 66% for handbags and 50% for backpacks. The frame rate for this figure is 445.20 milliseconds (ms). According to Fig.4, the algorithm successfully identifies various objects such as TV monitors, mice, books, cups, cellphones, potted plants, and vases. The system demonstrates a notable level of precision, achieving an accuracy rate of 96% for TV monitors and a perfect accuracy rate of 100% for mice. The detection of other objects exhibits varying levels of accuracy. The frame rate for this figure is 478.62 milliseconds. Fig.5 is dedicated to the detection of individuals and automobiles. The algorithm demonstrates a moderate level of accuracy, achieving a 57% success rate in detecting individuals and a high level of accuracy, reaching 97%, in detecting vehicles. The frame rate for this figure is 378.62 milliseconds. Fig.6 illustrates the algorithm's efficacy in detecting individuals, automobiles, and purses. The system demonstrates a remarkable level of precision, achieving a 100% accuracy in detecting individuals. However, it exhibits a comparatively lower accuracy rate of 37% for identifying cars and 44% for recognising handbags. The frame rate for this figure is 578.62 milliseconds. According to the findings presented in Fig.7, the algorithm successfully identifies various objects including beds, chairs, dining tables, cups, and bottles. The model demonstrates a commendable level of precision, accurately detecting beds with a success rate of 93%. It also exhibits varying degrees of accuracy when identifying other objects. The frame rate for this figure is 449.51 milliseconds. The primary objective of Fig.8 is to accurately identify and detect individuals as well as sports balls. The algorithm demonstrates a remarkable level of accuracy, achieving a perfect detection rate of 100% for identifying individuals and a notably high accuracy rate of 95% for detecting sports balls. The frame rate for this figure is 352.38 milliseconds. Finally, Fig.9 showcases the algorithm's efficacy in accurately detecting various objects such as laptops, books, individuals, potted plants, and chairs. The model demonstrates a notable level of accuracy when detecting laptops and individuals, while exhibiting comparatively lower accuracy when identifying other objects. The frame rate for this figure is 459.08 milliseconds. The table presents an analysis of the performance of the YOLO algorithm in detecting different objects across various images. The report showcases the level of accuracy attained for each object class and the frame rate at which the algorithm functions. These findings enhance the comprehension of the algorithm's efficacy in object detection tasks and can serve as a valuable point of reference for researchers and practitioners in the field of computer vision.

Fig.10. This graphic has been modified from the Focal Loss study [29].

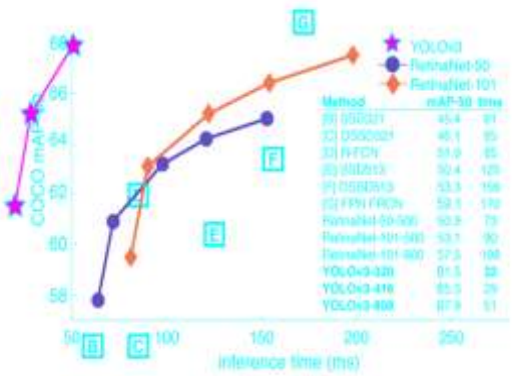
Fig.10 represents a modified graphical representation that has been derived from a comprehensive study conducted on Focal Loss. The Focal Loss is a specialised loss function that is frequently employed in object detection tasks to effectively tackle the challenge of class imbalance. YOLOv3 operates far more quickly than competing detection techniques of equivalent capability.

Figures	Class	Accuracy	Fps (ms)
Fig.3	person	99%	445.20
	Bus	99%	
	Handbag	66%	
	backpack	50%	
Fig.4	Tv monitor	96%	478.62
	Mouse	100%	
	Book	94%	
	Cup	98%	
	Cellphone	60%	
	pottedplant	91%	
	vase	51%	
Fig.5	Person	57%	378.62
	Car	97%	
Fig.6	Person	100%	578.62
	Car	37%	
	Handbag	44%	
Fig.7	Bed	93%	449.51
	Chair	85%	
	Dining Table	44%	
	Cup	98%	
	Bottle	60%	
Fig.8	person	100%	352.38
	sports ball	95%	
Fig.9	Laptop	99%	459.08
	Book	35%	
	person	100%	
	pottedplant	84%	
	chair	99%	

Table 1: Class Accuracy and Frames per Sec. of different Figures.



The modified graphic emphasises that YOLOv3, a particular iteration of the YOLO algorithm, demonstrates superior speed in comparison to other competing detection techniques that provide similar functionalities[41]-[43]. YOLOv3 is renowned for its capability to conduct real-time object detection, enabling rapid and efficient processing and identification of objects. The focal point lies in the fact that YOLOv3 not only attains exceptional accuracy in object detection, but also accomplishes this feat with remarkable speed. The speed advantage of YOLOv3 distinguishes it from other competing techniques, positioning it as the preferred choice in scenarios where real-time object detection is crucial



.

Fig.11. Garph of Modified YOLOv3 [40], here presenting the speed/accuracy trade-off on the mAP at .7 IOU metric.

The graph is positioned at a relatively high level and towards the left side. This indicates that the algorithm, most likely YOLOv3, demonstrates outstanding accuracy performance. The graph depicts a positive correlation between the position and the mAP score, suggesting that higher precision and recall in object detection are associated with higher values on the graph. Furthermore, the algorithm's placement on the far left suggests that it attains a notable level of accuracy while also operating at a commendable speed. According to the findings, YOLOv3 demonstrates a commendable equilibrium between accuracy and speed, surpassing alternative techniques in both these aspects. However, in the absence of specific details from Fig.11, such as axis labels or specific data points, it becomes difficult to offer a more accurate explanation. Based on the information provided, it can be inferred that the modified graph illustrates the notable accuracy and speed of YOLOv3, thereby highlighting its efficacy as an object detection algorithm.

5. CONCLUSION

Many methods have been developed in recent years to aid the visually impaired in recognising objects in their surroundings, but none of them are completely satisfactory. Our mission is to make it possible for the visually impaired to have access to a reliable and comfortable item detecting system. To save consumers needing to take pre-cut pictures, we utilise a USB camera to take pictures while they use our service. In this scenario, deep learning and the YOLO feature extraction technique are used. The YOLO fix helps with object recognition by picking the whole picture in one go, slicing it up into grids, then guesstimating the grid cells' bounding box coordinates and class labels. Sing YOLO's top quality is how quickly it can be sung. It is lightning fast and can handle a wide variety of objects with ease. This gadget uses text-to-speech technology to provide audible explanations of the user's surroundings, thereby rendering the visually handicapped invisible and facilitating their freedom of movement. The suggested approach is mobile, dependable, and cost-effective. In addition to providing a safe and immersive virtual environment, this technology also offers peace of mind by revealing the name of the recognised item.This project's long-term goals include enhancing computer vision rates, which can be handled by employing the Python library, and supplying an accurate estimate of the range between individuals and objects [44]. You should, however, take speedier hardware into perspective if you are creating an application that uses a lot of quickly changing objects. Furthermore, the same method may be used to do both text and facial recognition. The system will however be mostly comparable.

Conflict of Interest: Authors declare that there is no conflict of interest.

5. 6. REFERENCES

[1] Tielsch, J. M., Sommer, A., Witt, K., Katz, J., & Royall, R. M. (1990). Blindness and visual impairment in an American urban population: the Baltimore Eye Survey. Archives of ophthalmology, 108(2), 286-290.

[2] Ntakolia, C., Dimas, G., & Iakovidis, D. K. (2022). User-centered system design for assisted navigation of visually impaired individuals in outdoor cultural environments. Universal Access in the Information Society, 1-26.

[3] Lilley, J. M., Arie, T., & Chilvers, C. E. D. (1995). Accidents involving older people: a review of the literature. Age and Ageing, 24(4), 347-367.

[4] Parkin, J., & Smithies, N. (2012). Accounting for the needs of blind and visually impaired people in public realm design. Journal of urban design, 17(1), 135-149.

[5] Bahadir, S. K., Koncar, V., & Kalaoglu, F. (2012). Wearable obstacle detection system fully integrated to textile structures for visually impaired people. Sensors and Actuators A: Physical, 179, 297-311.

[6] babu Nuthalapati, S. Advancements in Generative AI: Applications and Challenges in the Modern Era. (2024). International Journal of Science and Engineering Applications

Volume 13-Issue 08, 106 – 111, DOI: 10.7753/IJSEA1308.1023

[7] Chan, M., Estève, D., Escriba, C., & Campo, E. (2008). A review of smart homes—Present state and future challenges. Computer methods and programs in biomedicine, 91(1), 55-81.

[8] Md. Atikur Rahman, Muhammad Sheikh Sadi (2021) IoT Enabled Automated Object Recognition for the Visually Impaired, Computer Methods and Programs in Biomedicine Update, Volume 1, 100015, ISSN 2666-9900.

[9] Soman, S. P., Kumar, G. S., Nuthalapati, S. B., Zafar, S., & Abubeker, K. M. (2024). Internet of things assisted deep learning enabled driver drowsiness monitoring and alert system using CNN-LSTM framework. Engineering Research Express, 6(4), 045239.

[10] Mahendru, Mansi, and Sanjay Kumar Dubey (2021) Real time object detection with audio feedback using Yolo vs. In 2021 11th International Conference on Cloud Computing, pp. 734-740. IEEE, 2021.

[11] Ferdousi Rahman, I. J. R., Farhin, N., Uddin, J.(2019). An assistive model for visually impaired people using yolo and mtcnn.

[12] babu Nuthalapati, S. (2023). AI-enhanced detection and mitigation of cybersecurity threats in digital banking. Educ. Adm. Theory Pract., 29(1), 357-368.

[13] Jiang, R., Lin, Q., Qu, S. (2016). Let blind people see: real-time visual recognition with results converted to 3D audio.Report No. 218, Standord University, Stanford, USA.

[14] Dongre, A. (2020). Visual Recognition Based System To Assist Blind Persons (Doctoral dissertation, Dublin, National College of Ireland).

[15] Nuthalapati, S. B., & Nuthalapati, A. (2024). Transforming Healthcare Delivery via IoT-Driven Big Data Analytics in A Cloud-Based Platform. Journal of Population Therapeutics and Clinical Pharmacology, 31(6), 2559-2569.

[16] 16. H. Jabnoun, F. Benzarti and H. Amiri, (2015) Object detection and identification for blind people in video scene, 15th International Conference on Intelligent Systems Design and Applications (ISDA), pp. 363-367, doi:10.1109/ISDA.2015.7489256.

[17] Yingli Tian (2020). Visual Recognition Based System To Assist Blind Persons (Doctoral dissertation, Dublin, National College of Ireland).

[18] Ashiq, F., Asif, M., Ahmad, M. B., Zafar, S., Masood, K., Mahmood, T., ... & Lee, I. H. (2022). CNN-based object recognition and tracking system to assist visually impaired people. IEEE Access, 10, 14819-14834.

[19] Babu Nuthalapati, S., & Nuthalapati, A. (2024). Accurate weather forecasting with dominant gradient boosting using machine learning. Int. J. Sci. Res. Arch, 12(2), 408-422.

[20] Yohannes, E., Lin, P., Lin, C. Y., & Shih, T. K. (2020, December). Robot eye: automatic object detection and recognition using deep attention network to assist blind people. In 2020 International Conference on Pervasive Artificial Intelligence (ICPAI) (pp. 152-157). IEEE.

[21] Joshi, R. C., Yadav, S., Dutta, M. K., & Travieso-Gonzalez, C. M. (2020). Efficient multi-object detection and smart navigation using artificial intelligence for visually impaired people. Entropy, 22(9), 941.

[22] Nuthalapati, A., Abubeker, K. M., & Bushara, A. R. (2024, September). Internet of Things and Cloud Assisted LoRaWAN Enabled Real-Time Water Quality Monitoring Framework for Urban and Metropolitan Cities. In 2024 IEEE North Karnataka Subsection Flagship International Conference (NKCon) (pp. 1-6). IEEE.

[23] Kadhim, M. R., & Oleiwi, B. K. (2022). Blind assistive system based on real time object recognition using machine learning. Engineering and Technology Journal, 40(1), 159-165.

[24] Bushara, A. R., RS, V. K., & Kumar, S. S. (2023). Classification of Benign and Malignancy in Lung Cancer Using Capsule Networks with Dynamic Routing Algorithm on Computed Tomography Images. Journal of Artificial Intelligence and Technology.

[25] Devi, S. K., & Subalalitha, C. N. (2022). Deep learning-based audio assistive system for visually impaired people. CMC-COMPUTERS MATERIALS & CONTINUA, 71(1), 1205-1219.

[26] Nuthalapati, A. (2024). Cloud data center performance optimization through machine learning-based workload forecasting and energy efficiency.

[27] 27. Hersh, Marion. "Route learning by blind and partially sighted people." Journal of Blindness Innovation and Research 10, no. 2 (2020).

[28] Chen, R. C., Saravanarajan, V. S., & Hung, H. T. (2021). Monitoring the behaviours of pet cat based on YOLO model and raspberry Pi. International Journal of Applied Science and Engineering, 18(5), 1-12.

[29] Islam, R. B., Akhter, S., Iqbal, F., Rahman, M. S. U., & Khan, R. (2023). Deep learning based object detection and surrounding environment description for visually impaired people. Heliyon, 9(6).

[30] Nuthalapati, A. (2024). Architecting data lake-houses in the cloud: Best practices and future directions. Int. J. Sci. Res. Arch, 12(2), 1902-1909.

[31] Pujara, A., & Bhamare, M. (2022, November). DeepSORT: Real Time & Multi-Object Detection and Tracking with YOLO and TensorFlow. In 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) (pp. 456-460). IEEE.

[32] Alsanad, H. R., Sadik, A. Z., Ucan, O. N., Ilyas, M., & Bayat, O. (2022). YOLO-V3 based real-time drone detection algorithm. Multimedia tools and applications, 81(18), 26185-26198.

[33] AR, B., RS, V. K., & SS, K. (2023). LCD-capsule network for the detection and classification of lung cancer on computed tomography images. Multimedia Tools and Applications, 82(24), 37573-37592.

[34] Han, X., Chang, J., & Wang, K. (2021). Real-time object detection based on YOLO-v2 for tiny vehicle object. Procedia Computer Science, 183, 61-72.

[35] Wang, K., & Liu, M. (2022). YOLOv3-MT: A YOLOv3 using multi-target tracking for vehicle visual detection. Applied Intelligence, 52(2), 2070-2091.

[36] Bushara, A. R., RS, V. K., & Kumar, S. S. (2024). The Implications of Varying Batch-Size in the Classification of Patch-Based Lung Nodules Using Convolutional Neural Network Architecture on Computed Tomography Images. Journal of Biomedical Photonics & Engineering, 10(1), 39-47,.

[37] Zhang, X., & Wang, G. (2022). Stud pose detection based on photometric stereo and lightweight YOLOv4. Journal of Artificial Intelligence and Technology, 2(1), 32-37.

[38] Kohut, P. (2013). Mechatronics systems supported by vision techniques. Solid State Phenomena, 196, 62-73.

[39] Subeh, P., & Bushara, A. R. (2024). Cloud data centers and networks: Applications and optimization techniques. International Journal of Science and Research Archive, 2024, 13 (02), 218-226.

[40] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar. Focal loss for dense object detection. arXiv preprint arXiv:1708.02002, 2017.

[41] Akenine-Moller, T., & Strom, J. (2008). Graphics processing units for handhelds. Proceedings of the IEEE, 96(5), 779-789.

[42] Kashika, P. H., & Venkatapur, R. B. (2022). Automatic tracking of objects using improvised Yolov3 algorithm and alarm human activities in case of anomalies. International Journal of Information Technology, 14(6), 2885-2891.

[43] Kheder, M. Q., & Ali, M. A. (2022). IoT-Based Vision Techniques in Autonomous Driving: A Review. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 11(3), 367-394.

[44] Muhammed Kunju, A. K., Baskar, S., Zafar, S., & AR, B. (2024). A transformer based real-time photo captioning framework for visually impaired people with visual attention. Multimedia Tools and Applications, 1-20.