Object Detection for Blind People with ML Algorithm

[1]Rohan Mule, [2]Shubham Chavan, [3]Sanihussain Patel, [4] Pooja Dehankar
[1-3]Bachelor of Technology, [4] Professor
[1]Department of Computer Engineering,
[1]Ajeenkya DY Patil University, Pune-412105, India

Abstract - Object detection has numerous applications in today's world, directly or indirectly affecting people's lives and helping people in many ways. Given this application and its usage, accurately detecting real-time objects is very efficient. This article describes the concept of real-time object detection and its conversion to audio format.

Index Terms - Image Processing, Machine Learning, Visually Impaired, Object Detection, YOLO, Key Point Extraction, Voice Output.

I. INTRODUCTION

In this research, we propose a method to automatically and accurately determine the size of an object using a stereo camera system and structured light. Four steps are performed before calculating the size: preprocessing, object detection, key point extraction, and depth interpolation. The first step in preprocessing is to align depth and RGB frames. We then use depth thresholding to identify objects and extract key points using the proposed key point extraction approach in combination with the ShiTomasi corner detector. A depth interpolation algorithm is created to correct for inaccurate depths on object edges. Finally, the Euclidean distance is used to derive the object dimensions from the 3D coordinates of the points of interest.

II. LITERATURE SURVEY

Paper 1: Real-Time Object Detection for Visually Challenged People
Authors: Sunit Vaidya, Naisha Shah, Niti Shah, Radha Shankarmani
Abstract: One of the most important senses for survival is sight. Visual impairment affects millions of people in this world. These people struggle to move independently and safely and have problems accessing information and communication. The aim of the proposed work is to transform the visual world into the auditory world by informing blind people about objects in their path. This allows visually impaired people to move around on their own without assistance, simply by using a real-time object recognition system. The application uses image processing and machine learning techniques to identify objects in real-time via cameras and inform visually impaired people of objects and their locations via audio output. The inability to distinguish between objects has led to many limitations of existing approaches, including reduced accuracy and poor performance. The main goals of the proposed work are to provide good accuracy, best performance results, and viable options for blind people to make the world a better place.

Paper 2: Real Time Object Detection using YOLO Algorithm
Authors: I.V.Sai Lakshmi Haritha, M.Harshini, Shruti Patil, Jeethu Philip
Abstract: The purpose of this research is object recognition using the You Look Only Once (YOLO) method. This method is much more efficient than existing models in terms of speed and performance. Some algorithms do not scan all regions in a single forward propagation, but YOLO's algorithm uses a convolutional neural network and class probabilities to predict binding boxes, which allows the entire image to be Analyze the YOLO is faster than other algorithms.

Paper 3: YOLO-Green: A Real-Time Classification and Object Detection Model Optimized for Waste Management
Authors: Wesley Lin
Abstract: Deep neural networks (DNNs) play an important role in our daily lives, from helping with simple tasks to solving global problems like detecting cancer cells. However, little research has been done to solve the global garbage crisis by using DNNs and deep learning models to classify and detect garbage. This is because current DNNs struggle to be efficient and accurate in detecting obscure objects such as debris. To address this issue, this paper focuses on YOLO-Green, a new real-time object detection model designed specifically for garbage detection. The model was trained on a dataset collected from real-world trash and classified into the 7 most common types of solid trash. With just 100 training epochs, YOLO-Green achieves an impressive 78.04 mAP and 2.72 FPS while maintaining a model size of only 117MB. Based on YOLOv4's unique object detection, YOLO-Green outperforms his YOLOv4 and other popular deep learning models in both accuracy and efficiency while maintaining a relatively small model size. Ultimately, this have a look at sheds a superb mild at the ability of the usage of deep getting-to-know fashions as an opportunity to guide waste management.

Paper 4: YOLO-compact: An Efficient YOLO Network for Single Category Real-time Object Detection
Authors: Yonghui Lu1, Langwen Zhang2, Wei Xie3
Abstract: In real applications, the number of categories for object detection is always simple. In this article, we propose an efficient YOLO compact network designed for real-time object detection within categories. In this article, we first explored a number of methods for transforming large and deep networks into compact and efficient networks through a series of ablation experiments. These methods were then incorporated into his YOLOv3 network, which maintains his YOLO-compact infrastructure. A network structure design approach is proposed that separates the down sampling layer from all network modules and facilitates the modular design of the network. The connection structure inside the RFB module has been changed to a direct connection structure, adopting a 1x3+3x1+3x3 convolution structure instead of a 5x5 convolution, realizing an efficient RFB-c module. The remaining bottleneck block was improved by removing the last his 1x1 fold layer and using his depth-separable 3x3 folds. Since pedestrians are the most representative objects in real-world applications, this paper uses results from the people category of the VOC2007 test set to demonstrate network performance. YOLO-compact's model size is only 9MB, 3.7 times smaller than tiny-yolov3, 6.7 times smaller than tiny-yolov2, and 1/26 smaller than YOLOv3. AP result for YOLO-compact is 86.85tiny-yolov3.32.

Paper 5: Object Detection and Distance Estimation Tool for Blind People Using Convolutional Methods with Stereovision
Authors: Rais Bastomi, Firza Putra Ariatama, Lucke Yuansyah Arif Tryas Putri, Septian Wahyu Saputra, Mohammad Rizki Maulana , Mat Syai'in, Ii Munadhif, Agus Khumaidi, Mohammad Basuki Rahmat, Annas Singgih Setiyoko, Budi Herijono, E.A. Zuliari, Mardlijah
Abstract: This research creates a tool that can provide information about objects in the environment. The tool can also estimate the distance of detected objects in combination with a camera and glasses, making it easier for blind people to use. This tool helps you identify objects around you and improve your skills and abilities. The tool uses a camera as its main sensor, which works like the human eye, providing real-time video as visual data. The visual RGB data is processed with a 2-fold 176 x 132 pixel Convolutional Neural Network. A small pixel of size 41x33 pixels is generated, so the classification weights are obtained using backpropagation and a specific dataset. Once we have the detection results, the next step is to find the centroid value and measure the distance between the object and the camera in stereo vision. The results are converted to audio and plugged into headphones so that blind people can hear the information. Test results show that the tool can recognize specific objects: people, tables, chairs, cars, bicycles, and motorcycles with an average accuracy of 93.33%. There is an error of about 6.1% when measuring distances between 50cm and 300cm.

Paper 6: The development of image-based distance measurement system
Authors: Lawrence Y. Deng
Abstract: This system uses two parallel laser beams as a scale to quickly calculate the actual horizontal or vertical distance. The basic idea of using ambient light to remove image noise worked perfectly in many conditions, but it had some pitfalls in low light and was a challenge when developing Night Sight. It turns out there is.
1) Auto White Balance (AWB) doesn't work in low light.
2) Tone mapping a scene that is too dark to see.

III.Problem Statement

Through Python language and artificial intelligence image processing technology, accurately detect objects, get the true distance, pronounce the name and approximate distance of the detected object.

IV.Proposed System

Proposed Method In this document, a proposed object detection and recognition system is presented. The proposed system uses You Only Look Once (YOLOv3) based on user-defined datasets to locate and recognize objects in indoor environments such as offices and rooms. The OpenCV library is used for image acquisition and processing. YOLOv3 detects and recognizes objects in each frame. Additionally, it uses the playsound library to play sounds and tell the user the position of the object (center, left, or right) within each frame. The proposed system is implemented on a personal computer using the Python programming language.

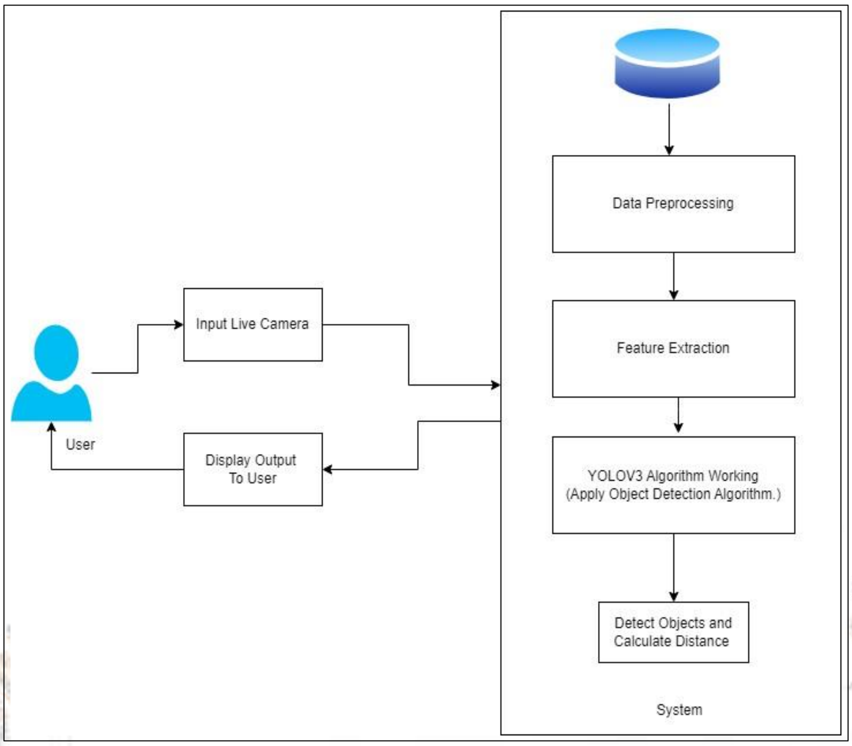

Fig. System Architecture

The following steps are used to capture and recognize objects using the proposed system.
Step 1: Used the OpenCV library to power a webcam and capture a frame (image).

Step 2: Used OpenCV to resize each frame (image).

Step 3: If there are objects in the frame, YOLOv3 will detect these objects and recognize them from the weight file. However, if the frame contains no objects, the next frame is selected. A weight file is generated from the training process.

Step 4: Once YOLOv3 detects and recognizes objects, it calculates their distance.

Step 5: Play a sound from the sound dataset using the playsound library to let the user know the object found in the frame and its distance.

V. CONCLUSIONS

In the proposed system, an effective real-time object measurement approach is suggested for industrial systems in the proposed system. Computer vision is employed in the system that is being applied to find, gauge, and measure items. In a real-time, the system is able to recognize and measure items. The distance of each object is determined after it has been identified using the YOLOv3 Algorithm for Object Detection. Object detection deals with detecting objects inside a certain image or video. The navigation systems is costly which may not be affordable for the common blind people. Thus, main aim of this project is to the develop a research level product which can be further enhanced to build a simple, refined, cost-effective product for blind people to detect nearby objects/obstacles.

VI.Future Scope

The extension of this system identifies any kind of entity with faster frame rate. The text to speech module has also been developed according to the futuristic pace. Instead of using the pre-trained models, self-trained models can be used. The model can be trained to recognize objects which are frequently encountered by the user. Thus, it can be customized for the specific needs of the user and ensures safer navigation.

VII. REFERENCES

[1] **"Real-Time Object Detection for Visually Challenged People",** Sunit Vaidya, Naisha Shah, Niti Shah, Radha Shankarmani, Information Technology Department Sardar Patel Institute of Technology Mumbai, India, 2022 6th International Conference on Electronics, Communication and Aerospace Technology, 2022.

[2] **"Real Time Object Detection using YOLO Algorithm",** I.V.Sai Lakshmi Haritha, M.Harshini, Shruti Patil, Jeethu Philip, 2022

[3] **"YOLO-Green: A Real-Time Classification and Object Detection Model Optimized for Waste Management",** Wesley Lin, 2021

[4] **"YOLO-compact: An Efficient YOLO Network for Single Category Real-time Object Detection",** Yonghui Lu1, Langwen Zhang2, Wei Xie3, 2020 Chinese Control And Decision Conference (CCDC), 2020.

[5] **"Object Detection and Distance Estimation Tool for Blind People Using Convolutional Methods with Stereovision",** Rais Bastomi, Firza Putra Ariatama, Lucke Yuansyah Arif Tryas Putri, Septian Wahyu Saputra, Mohammad Rizki Maulana , Mat Syai'in, Ii Munadhif, Agus Khumaidi, Mohammad Basuki Rahmat, Annas Singgih Setiyoko, Budi Herijono, E.A. Zuliari, Mardlijah, 2019 International Symposium on Electronics and Smart Devices (ISESD), 2019.

[6] **"The Development of Image-based Distance Measurement System",** Lawrence Y. Deng, Nick C. Tang, Timothy K. Shih, Dong-Liang Lee, Yu-Hsin Cheng, Kuo-Yen Lo, *Journal of Internet Technology*, vol. 10, no. 1 , pp. 65-71, Jan. 2009.