International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 06 | June 2021 www.irjet.net p-ISSN: 2395-0072

Real-Time Object Detection Accomplished by Assistance

Hima Patel[1], Ravi Patel[2], Parth Shah[3], Honey Patel[4]

[1-3]UG students, Department of Computer Science and Engineering, Institute of Technology and Management Universe, Gujarat, India
[4]Professor, Department of Computer Science and Engineering, Institute of Technology and Management Universe, Gujarat, India

---

Abstract - *Many people suffer Form temporary or permanent disabilities. Technology can be in many ways to help visually impaired people, but the problem of identifying objects is still a challenging task. Even though there exist many detection methods, the accuracy, Rapidity, and efficiency of detection are not good enough. So this paper demonstrates real-time object detection using the YOLO v3 algorithm by deep learning technique, aim is to design an efficient and cost-effective assistive device for people who are suffering from vision loss or impairment. It uses a smartphone to capture real-time objects at which the user is looking in its surroundings and gives voice feedback to the user. This project was designed and developed for the android device as an application.*

Key Words: **Deep learning, Computer vision, Object detection**

1. INTRODUCTION

In a cumulative study till 2015, there were around 940 million people around the world who were suffering from a certain level of vision loss. Out of these 940 million people, 240 million had very low vision and 39 million were completely blind. They are a crucial part of our society. Visually impaired people face complex issues while performing tasks that may seem very trivial to normal people, such as looking for something, perceiving and identifying objects in the surrounding, navigating through a route indoors and outdoors as well. In general, avoiding obstacles and perceiving the elements in the surrounding is a challenge for them. This project aims to create an assistive device using concepts of deep learning and computer vision to help blind people overcome hurdles in their way. The device should be capable of classifying different objects and identify them in real-time and notify the object class to the user through voice feedback. All the design and layout are kept by considering the visually impaired people in the mind. The opinions and views of the people are taken into consideration because they are the ones who are going to use this app.

1.1 Deep Learning

Deep learning is a subcategory of machine learning in Artificial Intelligence that utilizes multiple layers of neural networks to imitate human brain functions in processing data and creating patterns for critical decision making. Deep learning is a function of AI which consists of networks capable of unsupervised learning from the dataset which is unstructured. Deep learning has evolved in the digital era exponentially which has brought upon an explosion of data in all forms and from every region in the entire world. This data, known as big data, can be obtained from online sources like social media, e-commerce platforms, internet search engines, and online cinemas among other sources. This enormous amount of data is readily available and can be shared through applications like cloud computing. However, the dataset used can be unstructured and in such an enormous size that it could take decades for a human to comprehend it and extract relevant information and trends. Deep learning is an unsupervised learning algorithm that does not require data with desired variables and features. Instead, it utilizes an iterative way to get the desired result. Deep learning usually works fine with big data. Deep learning utilizes neural networks to scan the massive dataset to find patterns and fine correlations automatically. After the model is trained using the dataset, the learned associations are used to interpret new information.

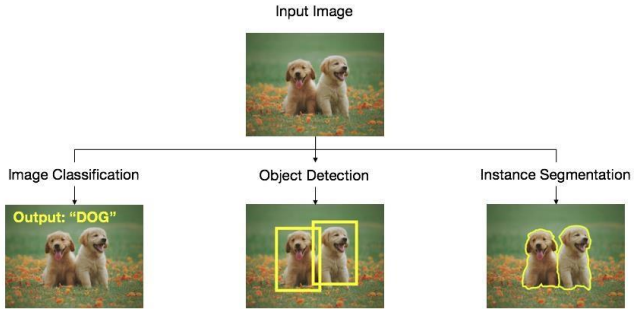
1.2 Computer Vision

Computer vision is the technology that is capable of using machines to process, understand and analyze imagery (both photos and videos). Although algorithms similar to computer vision have been around since the 1960s, due to recent advancements in machine learning, as well as leaps forward in data-storage technologies, much-improved computing capabilities, and cheap high-quality input devices, there is a major improvement in how well the software can explore this kind of data. Computer vision is a broader term for any computation involving visual content, such as images, videos, icons, or pixels in general. But within computer vision, there are a few specific tasks that are the core building blocks. In object classification, a model is trained on a dataset of specific objects. The trained model classifies the object belonging to one or more of the trained classes. In object identification, since the model is trained on a dataset

of an object, it should be able to recognize a specific instance of the object.

1.3 Object Detection

Object detection is a computer vision technique that allows us to identify and locate objects in images or video. With this kind of identification and localization, object detection can be used to count objects in a scene and determine and track their precise location, all while accurately labeling them. imagine, for example, an image contains two chairs, one bottle, and one cup. Object Detection allows us to at once classify the types of things found while also locating instances of them within the image. object detection draws bounding boxes around detected objects. Object detection is inextricably linked to other computer vision techniques like image recognition and image segmentation, in that it helps us understand and analyze scenes in images or video.



2. LITERATURE SURVEY

[1] Dr. Hemant I. Joshi and Dr. Balkrishna B. Soneji

The present novel invention Object Detection Based Smart Blind Stick Based on Object Detection, Navigation, Emergency Alert, Obstacle Detection, water detection, and Sound & Vibration Alert is capable of detecting objects and also capable of converting the name of the object into speech using Camera module (10), microprocessor, program, and Speaker (12) to provide ease to blind people for the object recognition from some distance without touching it

[2] Dr. Ravichandran Krishamoorthy and Dr. S. Chandra Sekaran

The present disclosure is related to the object Identification System which will be helpful for Blind People. The eye is one of the most wanted human sense organs, and it plays a critical part in human understanding of the surroundings. As a result, thousands of articles have been written on these topics, proposing a wide range of machine vision products and programs for the creation of modern technological aids for the blind. This invention aims to present a suggested system for restoring a core feature. The invention proposes a visual replacement method for blind people dependent on object detection in a video scene in this innovation. For target recognition, this invention employs SIFTS keypoint extraction and feature matching. The experimental results show that the proposed invention works well in identifying the objects in the video.

[3] Suraj Rameshsinh Pardeshi and Pravin Laxmikant Yannawar

This Invention "Optimized Deep Convolution Neural Networks Based Object Detection for Blind and Visually Impaired People" is a method and system for classifying the captured images in an indoor environment to identify the object. We recommend a machine-based object recognition and navigation tool for visually disabled people. There are three stages in the proposed system, viz. image acquisition, object's position, and class rating and subsequent score merging to create the latest article classes. The contours of the captured images are recognized with a remote webcam and by independently applying two incredible deep learning calculations, SSD and YOLO. After producing the scores for each class, they are ideally blended using the Syntax Conflict Resolution standard to create a tiered set. In considering these score estimates, the classification of the elements with the class, the most significant probability score is selected. The separate audio document indicating the name of the item is selected and the acoustic signal is sent to the blind customer via the consumable headset provided. The main objective to invent the system is, Walking firmly and without any human assistance with city conditions or darkness is a worrying mistake for amazed individuals. Narcotics face some problems in their daily life, one of these most important problems is the Identification of disincentives during the walk. When they move from one place to another, they need help from another place individuals around. The walk loses its independence. Sticks can be used but these are not reliable and not all have them. A person with a disability needs external authentication to help them solve problems along the way due to their inability. The activity is basically focused on providing a kind of visual guide to people with visual disabilities.

[4] Dr. Chetan Sharma, Dr. Anoop Kumar Shukla and Dr. Shray Pathak

In this invention Smart Spex: Detect the object and speak object name using machine learning is a integrating electronic components into an electro-active frame for driving electro-active focusing lenses and also this is accomplished in a cosmetically pleasing manner that allows a platform of frame systems to be built from a single electronic module. The invention is to controlling an electro-active lens in a deliberate, hands free manner that gives the user control of the electro-active lens and an approach that detects objects crossing a virtual boundary line is provided.

The invented technology is also including an object detection tool that provides this capability and the object detection tool comprises a boundary component configured to define a virtual boundary line in a video region of interest and establish a set of ground patch regions surrounding the virtual boundary line. The invented technology is also the object detection tool further comprises an extraction component configured to extract a set of attributes from each of the set of ground patch regions, and update a ground patch history model with the set of attributes from each of the set of ground patch regions. The invention is also an analysis component is configured to analyze the ground patch history model to detect whether an object captured in at least one of the set of ground patch regions is crossing the virtual boundary line in the video region of interest. The invented technology is also including a method for communicating navigation information on a physical environment to a user and the method comprises receiving, by one or more computer processors at least one digital image of a physical environment of a user that is captured by one or more digital video devices and converting the at least one digital image into a three-dimensional image. The method includes analyzing the three-dimensional image using object analysis to generate output data where the output data corresponds to the physical environment and the method includes determining a device associated with the user and formatting the output data for use with the device.

3. OBJECTIVE

The system makes use of object detection and Along with it, the application will send an alert signal to the user if it detects any obstacles which may come in the way.

3.1 Object detection

The application makes use of the YOLOv3 model for object detection. It uses a single neural network for the whole input image. The network then divides the input image into several different regions and predicting the bounding regions in form of boxes with their probability score.

3.2 Dataset

In this project, Common Object in Context (COCO) dataset is used for training the YOLO model and text reader which recognizes 80 different categories.

4. PROPOSED SYSTEM

The system is implemented in an android application that detects various objects in real-time.

4.1 System Overview

The system uses a smartphone to capture real-time input data. The app's camera is automatically accessed and it starts capturing the surrounding objects and App notifies them about it. the system speaks out about every action to make them understand and navigate through their system.

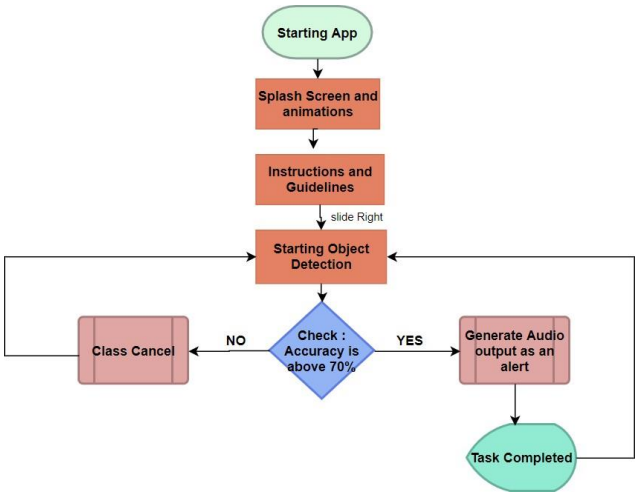


Fig -1: system diagram

4.2 Implementation

The system is implemented by combing various technological stacks which are discussed below.

Android studio is used for the development of the application as it is the official integrated development environment (IDE) designed specifically for android application development.

TensorFlow Lite is the lightweight version that is specifically designed for the mobile platform and embedded devices. It provides a machine learning solution to mobile with low latency and small binary size. TensorFlow supports a set of core operators which have been tuned for mobile platforms. It also supports custom operations in models.

YOLO (You Only Look Once) is used to process the video in real-time. It uses a single convolutional network which is faster than the R-CNN (Regional Convolutional Neural Network) and predicts the detected object in the form of bounding boxes.

5. EXPERIMENTAL RESULTS AND ANALYSIS:

The application can detect and recognize various categories of obstacles that may be faced while walking, objects of daily use, and inform the user through Assistant. A user should have a smart android phone and this application will able to

help them in every way covering corner things also of indoor environment and in noncrowded area.

Fig -2: Results

6. CONCLUSION

Thus, the proposed system would be very helpful and will improve the user's experience by providing them guidance with their android device. The auto assistance will play a crucial and extensive role with the user. Assistance can also be Consider as the third eye of the user. Our main focus is to create such an application that allows the visually impaired person to feel their surroundings just by hearing the assistant. Therefore, it will help to prevent possible accidents. The mobile devices can be carried easily and the camera of the device can be used to detect an object from the surroundings and give output in audio format.

REFERENCES

[1] Liyan Yu, Xianqiao Sansan Zhou, "Research of Image Main Object Detection Algorithm Based on Deep leaning", IEEE, 2018.

[2] Wang Yang, Zheng Jiachun," Real time face Detection Based on YOLO", IEEE, 2018.

[3] Aditya Raj, Manish Kannaujiya, Ajeet Bharti, Rahul Prasad, Namrata Singh, Ishan Bhardwaj, "Model for Object Detection using Computer Vision and Machine Learning for Decision Making", International Journal of Computer Applications, March 2019.

[4] Microsoft "COCO Dataset", http://cocodataset.org/

[5] Android Developer, https://developer.android.com/

[6] YOLO weights, https://pjreddie.com/darknet/yolo/