

Assisting Blind People Using Machine Learning Algorithms

Prof. Ephzibah E.P

Shivam Gupta(19mca0026)

Site, Vit

Vellore, Tamil Nadu

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: In this project, Machine learning techniques are applied in order to assist blind people by detecting real time objects and the distance between the objects and the camera. The application will be prompting the user about the object which is very close to them for avoiding collision. In this project objects are detected with the help of tensorflow lite and result is passed to the user in voice mode. In this Application user can also perform certain actions to know their current location, time etc. User will be opening the application using the "OK GOOGLE" command, which opens the application, the camera gets open and starts detecting the objects and telling the user regarding the closest object to them with the relative distance using TTS (i.e voice command). In between User can also perform different operations like knowing time, locations etc just by performing different gestures on the phone.

Keywords- AndroidStudio, TensorflowLite, TTS, Speech Recognizer.

I. Introduction

There is always a challenge when developing a guiding system for blind people. There are many people who are suffering from visually impaired or blindness, these people face a lot of difficulties in their day to day activities. The most difficult task for them is moving in an unfamiliar world of which they have no idea what's around them. Generally people rely on their visions to familiarize with nearby surroundings around them. In this people who are visually impaired or blind can move freely around the real time environment without the collisions as they will be getting all the information about the about the objects in their surroundings and the distance of object from them through voice mode. This will help users to get a rough idea about where they are what's around them to avoid possible collisions. With the help of this project they can recognise different objects just by pointing the camera around themselves to know the object and their size and relative distance from them. Even for opening the application they can just use "OK GOOGLE" service provided by google and after that the application will help them to move freely.

The main focus in developing this application is to make an application which allows them to fit with other people as most of the system makes them feel like they are different from other people. But with this system, whenever a user will use this application they will look like normal people as most of the people these days carry mobile phones in their day to day life.

- **Tensorflow Lite-** TensorFlow Lite is a deep learning system for on-device inference that is open source. This toolkit is used for running ML inference on devices. It has TFL converter which will convert the ML model in such a form that will work really well on embedded or mobile devices and then the TFL interpreter will run the model and run inference for you on devices. Main advantage of using Tensorflow Lite is that it does model pruning and quantization. Application of Tensorflow lite is the predictive Capacity for emoji on pixel devices.
- **Android Studio-** Android Studio offers a seamless environment where Android phones, tablets, Android Wear, Android TV and Android Auto apps can be created. Structured code modules allow you to divide your project into functionality units that you can create, evaluate, and debug independently.
- **TTS-** TTS stands for Text To Speech. This application helps to convert the text into speech format. It can voice out in different languages and at different speech rate. Also with this application you can set the pitch of the speech.
- **SpeechRecognizer-** This is a service which basically recognizes voice and converts it into text. It basically has many functions with it like postspeaking, BufferSpeaking etc.

Programming Languages/ Tools / DBMS Used for Implementation: JAVA, Python, Tensorflow Lite, Machine Learning, Android Studio, XML

Literature Survey

In paper[1],author has interviewed 4 blind people to know what features or functionalities the ideal navigation system should have.This system was based on Haptic information Tactile simulation .In this there is a device wearable on hand.This device contains motor which rotate at certain frequency to help the blind people.so for example if user need to move left so the device will pinch the user on left side etc.Their whole concept was based on simulation.

In paper[2],author has developed an application which will help them to assist to detect different objects ,currency recognition,ebook ,chatbot etc.This application will assist the user using voice commands.it can detect the text in the hard copy document.According to user,blind people have to click an image to detect the object.They have used NLP and vision API in this application.

In paper[3],author has developed two devices one is in the hand of the blind people and one is in the bus.Both the devices has Zigbee Transceiver.The bus arrival information is passed from the bus device to the blind people device using Zigbee.The user tells the destination through the microphone in the device which passed that information to the device in the bus.The buzzer and led in connected in the bus to alert the bus driver when the device of blind people starts communicating.This not only helps the blind people but also to the elder people.

In paper[4], author has provided different techniques with their pros and cons to help the upcoming developer to choose the best approach to assist blind people .Techniques shown are: indoor positioning, computation offloading, or distributed sensing, to the analysis of spatial related perceptual and cognitive processes of BVI people.In this paper author has taken many approaches that are done in past and figured out why they are not the best or why they are good ,their limitations etc.

In paper [5] ,the author has made a project which will assist blind people to move in an indoor environment freely without any collision.The system detects dynamic obstacles and adjusts path planning in real-time to improve navigation safety.In this paper, they introduce a SmartCane prototype ISANA mobile computing to assist blind people with indoor navigation. ISANA features include the development of indoor semantic charts, navigation and wayfinding, obstacle avoidance, and a multi-modal user interface. With this user, they can be aware of their environment and travel accordingly.

In Paper [6], the author has made a project which will assist blind people to navigate and track their location if they are lost.They have used two languages to guide people that are Bengalli and English.According to author the voice commands are passed to user which will tell the object position.Also in this application user can perform a call without touching their phone just by pressing the headset button.Navigation is performed using GPS.Technology involved:Arduino;Android,RFID,Ultrasonic,GSM,GPS,Serv-ver,Microcontroller, Pulse,Analog and Modulation.

In paper [7]the author has made a project which is to recognize potholes and uneven surfaces that can improve the mobility of blind people even at night. In this paper, computer vision is achieved. The method involves projecting laser patterns, documenting the patterns through a monocular video, analyzing the patterns to extract characteristics and then providing the blind user with path indications.

In paper[8] the author has suggested a prototype blind-assistant object seeking device by camera-based network and matching-based recognition in their paper. In this, they have gathered a dataset of daily necessities and add Speed-Up Robust Features and Scale Invariant Feature Transform feature descriptors to perform object recognition. Their key aim is to help blind users identify their personal objects in everyday life.

In paper[9] the author has made a project which is used to detect static and dynamic objects from the real time environment and transform them into acoustic signals.The device is based on stereo-vision technology, which combines accurate static and moving obstacle and free path tracking technology in real-time. The device is designed to provide three-dimensional ambient

information and to transmit it to the user via the propagation of acoustical signals. It can support blind people with items that range from 5m to 15m.

TABLE 1
Recent System Vs Proposed System

| Properties | Recent System | Proposed System |
|------------|--|---|
| Reliable | Most of the systems are giving results after some time. | giving results at a very fast rate. |
| Portable | Most of the systems involve different sensors, camera, sonar, helmets etc. | only involve camera which is easy to carry. |
| Cost | very expensive | depend on the price of the phone. |
| Misfit | These systems will make user feel that they are different from others. | This will help them to fit with all people as camera is normal to all people |
| Technology | Works on Technology that is old. | works on new Technology as it Uses android and Machine Learning, Tensorflow Lite. |

II. The Proposed Method

The proposed system works on any mobile phone that has a camera in it. In this system there are no other objects taken to detect objects or to get the result. This system is basically an Android application developed using android studio. In this object detection is done using Tensorflow lite which is a deep learning system for on-device inference that is open source. This toolkit is used for running ML inference on devices.

Tensorflow Lite vs Tensorflow

- **Description:** Tensorflow lite is a lightweight solution of Tensorflow for mobile and embedded devices, whereas Tensorflow model TensorFlow is used to create large-scale neural networks with many layers. Also Tensorflow can be used for both networking and inference whereas Tensorflow Lite is developed only for running inference on devices with limited compute.
- **Size:** Tensorflow Lite creates a small binary size as compared to Tensorflow mobile .Tensorflow Lite achieve this thing in two ways:
 - **Quantization:** This reduces the precision of numbers which are used to represent different parts of the Tensorflow model like weights and activation Output . So with this weight with 32-bit floating points gets converted to 16-bit or 8-bit floating point which reduces the size of the whole model.
 - **Weight Pruning:** This basically trims parts of the model which are of very less impact and does very less difference to the performance of the model.

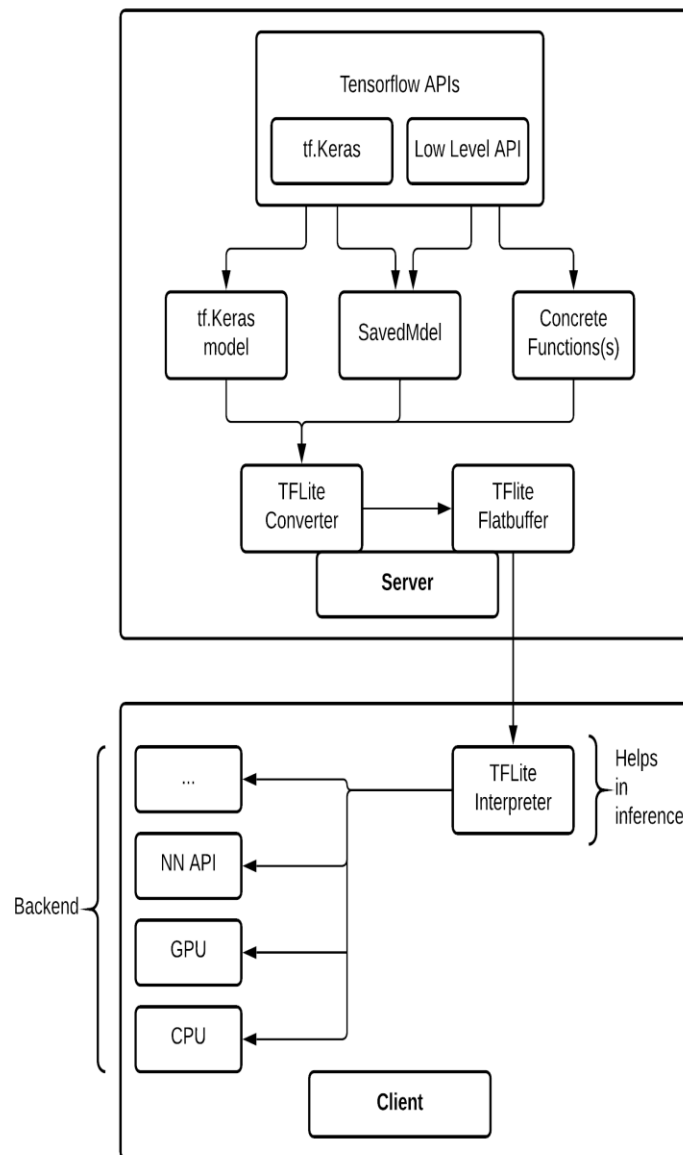


Fig 1.0 Working of TensorFlow Lite

- **Latency:** Tensorflow Lite uses Delegates to increase the efficiency of the model running on mobile devices.it basically hand over parts of graph execution to another hardware acceleration.where as Latency is Tensorflow mobile is higher as compared to Tflite model due to size and as it does not have delegates to increase the efficiency.Also as the inferences are made on the edge devices. so the data transfer from devices to server is eliminated,making inference faster.
- **Secure:** Tensorflow Lite model is deployed on edged devices and inference is also done on edged devices so no data is transferred from edge devices to network so data privacy is there.

In proposed system objects detection and classification is done using Tensorflow Lite Android Example which detects the object and classifies them into different categories. after Classification the distance of the object from the camera is found. For finding the distance between object and the camera different methods are compared :

- In the first method distance is found using width,focal length and pixel of camera. this formula failed to give the correct result as it doesn't take the height of the object detected into the account which gave the wrong output.

$$D = (W * F) / P$$

where D ,W,F and P stands for distance calculated,Width,Focal length and Pixel respectively.

- In the second method ,the distance between object and camera is calculated but in the formula the distance is found using the height of the object without camera .But knowing every object length will not make the project general.
- In the third method which is used in this project,the distance is calculated by using a simple approach which says that when an object is close to the camera the area of the object maximises and vice versa . so by using this simple approach ,distance is calculated. So in this system 5-6 data entries were taken with their relative distance from the camera.then adding those entries and making a graph for them. After that a polynomial equation is formed using Microsoft Excel software and that equation is used to find the distance between detected object and the camera.

$$distance = (int) (((((Math.pow(10.0d, -10.0d) * 3.0d) * Math.pow((double) area, 2.0d)) - ((Math.pow(10.0d, -5.0d) * 2.0d) * ((double) area)))) + 1.6889d) * 39.97d)z$$

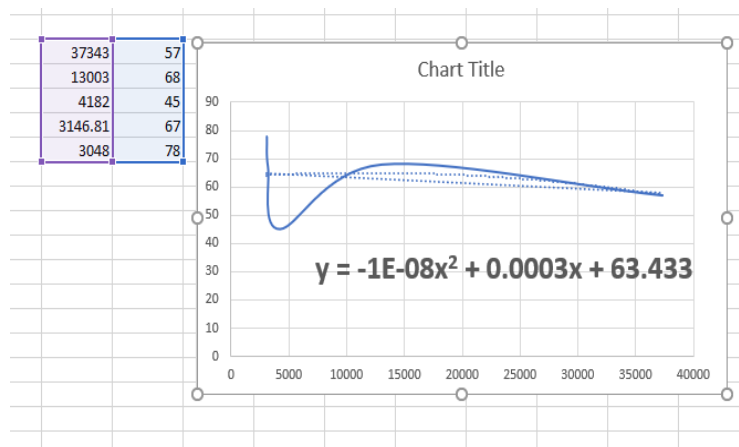


Fig 1.1 Distance formula for calculating distance from detected object to camera

After calculating the minimum distance from that is taken because at one time many objects are recognized and it will create chaos for the user as the user will be hearing about the object and the distance using voice command using the TextToSpeech module. So the minimum distance object is taken into consideration and that object's category and the distance from the object to the camera is conveyed to the user.TTS module is used to convert the result into speech. The pitch and the rate can

be modified and different languages can be changed using the TTS module.

Blind people can also get their current location by tapping on the screen at any place, for knowing date and time they can double tap on the screen. All that information is passed to them using voice commands by the TTS module.

Modules Description

- **Object detection:** With the help of Tensorflow Lite objects are detected from real time environment.
- **Object Description:** Different properties like size, name and their relative distance from the user is gathered.
- **Voice commands:** All the functionalities are communicated to the user through voice commands.
- **Warning Measure:** If the object is too close to the user, User will get a warning voice command to alert the user for possible collision.
- **Important Details:** Users will be notified about some important details like time, current Location and about weather after some intervals. They can get those information just by performing different operations like tapping, double tapping and swiping etc.

Flow Diagram

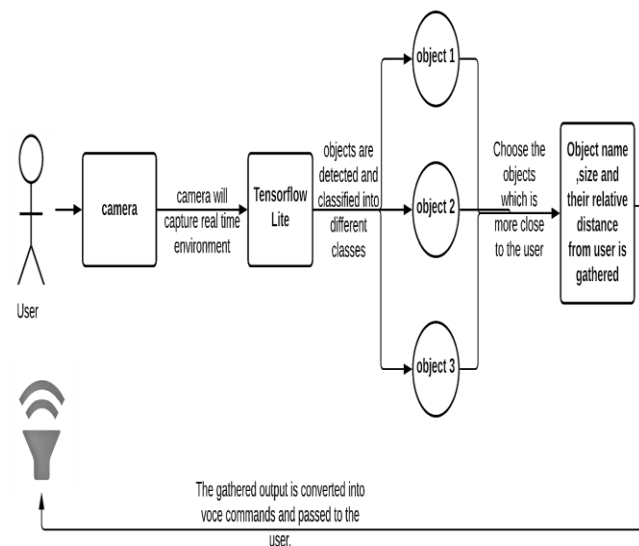


Fig 1.2 Work flow of the system

In fig 1.1, The workflow is started with camera and user as user will opening the Application named as “NOW YOU CAN SEE” by calling “Ok GOOGLE” and then saying open *application name*. This will open the application, as soon as application is opened the camera will get started and started detecting the objects in real time environment and categorising them in different categories with the help of tensorflow Lite. The Minimum_confidence taken in this application is 60% as there should be no faulty categorizing of objects because if minimum confidence is decreased the probability of faulty detection and classification increases which can be dangerous for blind people.

After different objects are classified, distance of objects are calculated one by one and the object with smallest is taken. As it is working in real time environment so objects are detected continuously so as soon as new object is detected and classified distance of new object is compared with the previous object. If its bigger than only smaller distance object is chosen and vice versa. After the object is chosen properties of object like name and distance is passed to the TextToSpeech module which will

transform text into voice commands . There is a sleep time for 3sec after the voice command is passed to the user so that they won't feel irritated or headache by continuous passing of voice commands.

In between users can perform different functions like knowing their current location ,time etc.by performing different gestures on screen.

The detection and classification module will not stop as it is in synchronized function and will be running in background throughout.

This Application will be taking some permission from the user like:-

- `<uses-permission android:name="android.permission.ACCESS_COARSE_LOCATION"/>`
- `<uses-permission android:name="android.permission.CAMERA"/>`
- `<uses-permission android:name="android.permission.RECORD_AUDIO"/>`
- `<uses-permission android:name="android.permission.INTERNET"/>`
- `<uses-permission android:name="android.permission.ACCESS_FINE_LOCATION"/>`
- `<uses-feature android:name="android.hardware.camera" />`
- `<uses-feature android:name="android.hardware.camera.autofocus" />`

IV. Working Images of Application

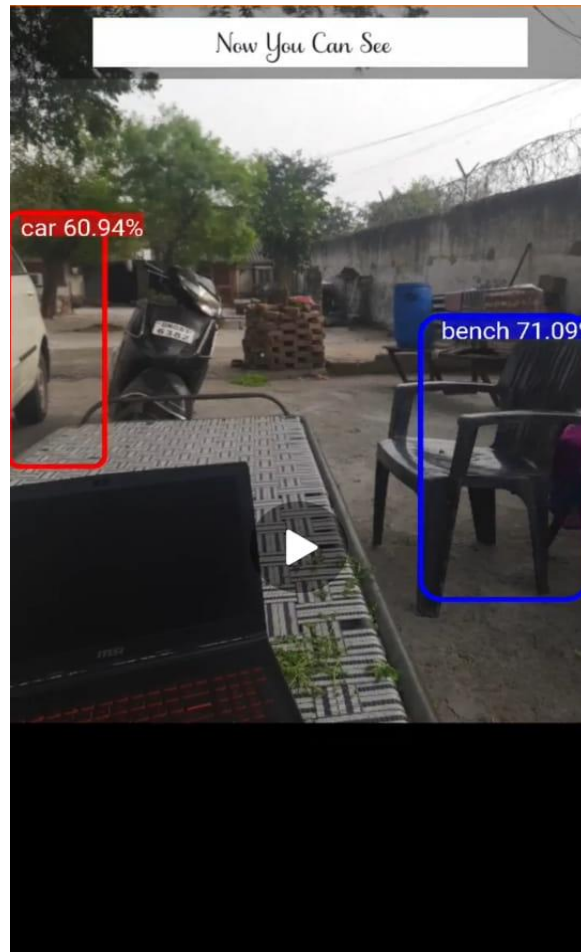


Fig 1.2 Detecting objects



Fig 1.2 Detecting Nearest object

V. Limitations

This system is working fine and it is covering all the modules but it has some limitations too. The Battery consumption as camera is turned on throughout this application which drains the battery very fast. Also this Application is running Machine learning inference on devices which involves lots of processing and this leads to the heating issues of edged devices. The limitations discussed till now are related to the device which is used to run this application but there are some other limitations too which are regarding the applications. The object detection is little slow. There comes little latency because the voice command is passed after all those detection, classification, distance calculation and finding the object which is more close to the user. so this shows little latency. This application does not work at night because of visibility issues of the object. These all issues will work out as the technology is evolving at a very fast rate. so a system with a good processor will cover all these limitations.

VI. Conclusion

In this research, There are many systems which are already developed which can give better results but they are very hard to carry out in normal day to day life. Even they are not feasible and more expensive. The proposed system discussed here will be easy for visually impaired or blind people to carry around. This system will assist the user to know about their surroundings and assist them in knowing different objects. This system will help them to navigate and avoid possible collisions. Using this application blind people can navigate from one place to another without collision and can know their locations and time in between. They can also use this app to recognize day to day life objects at their home.

VII. Future Scope

For future scope, This system will work very efficiently if they are two camera in an edge device as one of them will focus on one object which is close to the user and tell them regarding the object and one camera will focus on other objects and if other object is more close to the user then the focused object so they both will be swapped so this will make the system more optimized and give efficient result than system right now. Also with the advancement of the technology processors are going to be very fast which will be delivering the result very fast.

VIII. References

- [1] Barontini F, Bettelani GC, Leporini B, Averta G, Bianchi M. "A User-Centered Approach to Artificial Sensory Substitution for Blind People Assistance". (2020).
- [2] Swathi, Ms, and Mrs Mimitha Shetty. "Assistance System for Visually Impaired using AI." *International Journal of Engineering Research & Technology (IJERT)* ISSN: 2278-0181 (2019).
- [3] Prasad, V. Siddeswara, and R. Sagar. "An implementation of a device to assist the visually Impaired/blind people for easy navigation through bus." In *IOP Conference Series: Materials Science and Engineering*, vol. 1065, no. 1, p. 012048. IOP Publishing, 2021.
- [4] Real, Santiago, and Alvaro Araujo. "Navigation systems for the blind and visually impaired: Past work, challenges, and open problems." *Sensors* 19, no. 15 (2019): 3404.
- [5] Li, Bing, Juan Pablo Munoz, Xuejian Rong, Qingtian Chen, Jizhong Xiao, Yingli Tian, Aries Arditi, and Mohammed Yousuf. "Vision-based mobile indoor assistive navigation aid for blind people." *IEEE transactions on mobile computing* 18, no. 3 (2018): 702-714.
- [6] Tanveer, Md Siddiqur Rahman, M. M. A. Hashem, and Md Kowsar Hossain. "Android assistant EyeMate for blind and blind tracker." In *2015 18th international conference on computer and information technology (ICCIT)*, pp. 266-271. IEEE, 2015.
- [7] Rao, Aravinda S., Jayavardhana Gubbi, Marimuthu Palaniswami, and Elaine Wong. "A vision-based system to detect potholes and uneven surfaces for assisting blind people." In *2016 IEEE International Conference on Communications (ICC)*, pp. 1-6. IEEE, 2016.
- [8] Yi, Chucai, Roberto W. Flores, Ricardo Chinchá, and YingLi Tian. "Finding objects for assisting blind people." *Network Modeling Analysis in Health Informatics and Bioinformatics* 2, no. 2 (2013): 71-79.
- [9] Dunai, Larisa, Guillermo Peris Fajarnes, Victor Santiago Praderas, Beatriz Defez Garcia, and Ismael Lengua Lengua. "Real-time assistance prototype—a new navigation aid for blind people." In *IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society*, pp. 1173-1178. IEEE, 2010.
- [10] Elrefaei, Lamiaa A., Mona Omar Al-musawa, and Norah Abdullah Al-gohany. "Development of an android application for object detection based on color, shape, or local features." *arXiv preprint arXiv:1703.03848* (2017).
- [11] Alsing, Oscar. "Mobile object detection using tensorflow lite and transfer learning." (2018).
- [12] Xiao, Jizhong, Samleo L. Joseph, Xiaochen Zhang, Bing Li, Xiaohai Li, and Jianwei Zhang. "An assistive navigation framework for the visually impaired." *IEEE transactions on human-machine systems* 45, no. 5 (2015): 635-640.