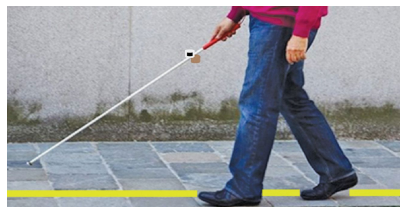




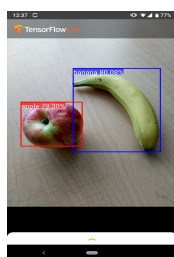
Tell us what your idea is.

The idea is to help visually impaired people to listen to what is in front of them or what they can't see. Instead of the stick they carry to guess where any object is,

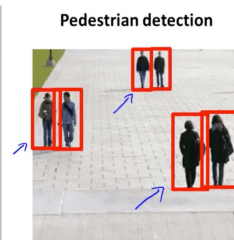


Source- <https://electronicsforu.com/wp-content/uploads/2019/04/3-3.jpg>

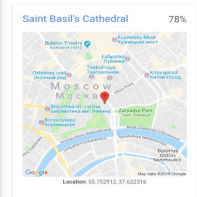
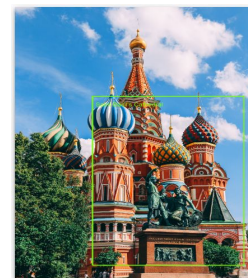
holding an android app with camera access to detect objects in front of them and describe the object through a voice. Using **machine learning's various aspects like real time object detection technique, pose estimation, text recognition, landmark recognition**, I want to create an android app that would detect the surroundings of the visually impaired person and describe them through voice output - text to speech engine.



+



+



=



(Images source - Google Images)

This app will be also capable to solve their day to day problems like :

- Looking for lost objects. Example - wallet, cap, goggles.
- Will use optical character recognition techniques or text recognition to read newspaper or some form's details.

### How machine learning will be used?

This app will be completely based on the **real time object detection and tracking** along with **text and place/landmark recognition** to identify the surrounding objects. **Pose estimation** techniques to identify the behavior of the people in the surrounding to avoid collisions or to ask for help and also for security purposes. Machine learning models will be trained on a regular basis using the the places visited and the tasks done by the user.

### Thinking about how a visually impaired person will open the app to use it?

Google assistant has an answer to every problem. :)

The user will be able to open the app and use it by asking google assistant to open the app and the rest the app will manage itself taking care of everything without much user interactions.

### Some example dialog flow

#### **\*\*While Looking for objects\*\***

*"There is an apple in front of you".*

*"A hat is on slight left".*

*"A bag is on slight right".*

#### **\*\*While walking\*\***

*"There is a building in front".*

*"The board on slight left reads 'Construction work ahead' ".*

*"Two people are present in front of you".*

*"Oh there we got a dog".*

*"There is a bench ahead".*

### Tell us how you plan on bringing it to life.

---

[Github Repository](#)

My project is at the concept phase for now. I am currently planning on what all and how to implement, also which technologies or machine learning tools will be best for this app.

### **Features to be implemented**

- Real time object detection and position tracker.
- Text recognition.
- Landmark recognition and pose estimator.
- Voice outputs - Text to speech engine.
- Linked with Google assistant using deep links.

### **Tools to be used**

- Detect and track objects.
- Tensorflow lite's pose estimation.
- Text recognition.
- Place/Landmarks recognition.

A big thanks to **Google's ML Kit** and **TensorFlow lite** to provide all the required tools with proper optimisation.

### **Timeline**

#### **1. Initial phase - Object detector and voice output (15th February, 2020)**

- a. Set up real time object detector using Google's ML kit and to receive output as voice or media.
- b. Managing connections to earphones, bluetooth headsets etc.
- c. Make the app usable by the user without much user interactions. ( Deep linking with Google assistant)

#### **2. Second phase - Advancement(31st March, 2020)**

- a. Implementing tensorflow lite's pose estimation tool to depict or estimate the behaviour of the person if the detected object is a person.
- b. Implementing optical character recognition or text recognition to read the contents of a notice board or a restaurant name if the detected object is some text.
- c. Landmark recognition to describe the place to the user.
- d. Optimizing dialog flow of the voice output.

#### **3. Testing, Model training and app optimization phase.(18th April, 2020)**

- a. Collecting data and managing a dataset of the user to train our machine learning model to improve itself.
- b. Optimizing and making app more user friendly.
- c. Implementing various tests.

#### 4. Released :D (25th April, 2020)

#### Tell us about you.

---

I am Aryan Ganotra, a Computer Science engineering undergraduate from Delhi Technological University, New Delhi, India. I have been developing android applications and working on various Android concepts from past one year.

[Resume](#)

[Github](#)

#### Here are my some project links

1. Blog application and notes downloading android application for Economics Department of Jesus and Mary College, New Delhi, India.
2. Android Blog application and stock prices viewer for Finance and Investment Cell of Shri Ram College Of Commerce, New Delhi, India

[Source code](#)

[Play Store](#)

[Letter of recommendation](#)

[Source Code](#)

3. Mathomon - The flappy math - A minimalistic educational game based on various arithmetic operations.

[Play Store](#)

[Source Code](#)

#### Conclusion

I am aware of the various drawbacks like latency and continues detection of unnecessary objects but we can always find a solution to a problem if worked on them.

I have tried my best to explain the idea but if you are having a hard time getting it or have some doubts regarding it, I am ready to get back and answer any query.

Thanking You,  
Aryan Ganotra