

NETRA: Android Application for Visually Challenged People to Detect and Recognize Object, Person and Text

Yashashree Bapat¹, Akash Bhardwaj², Akash Prajapati³, Prof. Deepti Vijay Chandran⁴

^{1,2,3,4}Smt. Indira Gandhi College of Engineering (Mumbai University)

Abstract: The inability to recognize known individuals and surrounding objects in the absence of audio or haptic cues severely limits the visually impaired in their social interactions and puts them at risk from a security perspective. In recent years, several prototype systems have been developed to aid this population with the face recognition task as well as object detection. The object and face recognition module helps visually impaired user to recognize various objects around his/her surroundings and also the person to whom they are speaking with.

Key Words: User, Application, OS, Face recognition, Object recognition, Text recognition, Dilib Library, Tenserflow Liter.

I. INTRODUCTION

The blind and the visually impaired face diverse kinds of life challenges that normally sighted people take for granted. As far as out-door activities are concerned the blind indicate difficulties in safe and independent mobility depriving them of normal professional and social life. Visual impairment afflicts approximately 285 million people worldwide according to recent estimates World Health Organization (WHO) and, without additional interventions, these numbers restricted to increase gradually. Many challenges faced by this population some of them are ability to recognize the faces and objects they encounter them in their daily lives. One consequence of this is that whenever a visually impaired individual arrives in a social setting Then the issues dealing with communication and access to information are pointed out. Here a significant help is offered by software applications for computers and touch-screen devices equipped with speech synthesizers that enable browsing the internet and access to text documents.

I.I. PROBLEM STATEMENT

In day-to-day ongoing life everything is getting dependent on computer based technologies. Computing environment is getting more closer towards Human Computer Interaction designs. As far as, out door happenings are considered the blind face difficulties in safe and autonomous mobility .As we know that the visually impaired user is always anxious to know what is going on in his/her surrounding, so a significant help is provided by the android application. This is an innovative Application for the visually impaired people and acts as voice assistance for them. This Application is used to make them aware about what is around them and to whom the person is speaking to. Application also detects an object present in a surrounding. The main module can recognize and match scanned objects and person and provide information about them in form of speech.

I.II.OBJECTIVE

The objective is to give blind users the ability to move around in unfamiliar environment.

- 1) The visually impaired clicks a photo of the person standing in front of him through the application.
- 2) The application then scans the image and checks whether the image of that person is present in the database and also provides the information (if present) in form of speech.
- 3) If the image is not present in the database then the application will ask for a new entry.
- 4) They can also click a picture of surrounding if any object appears in front it recognize the object and provide information about it in the form of speech.
- 5) The application also recognize the text and read the information which is been scan by the application.
- 6) The application takes input in speech form and gives output in form f speech.

I.III. SCOPE

Further this application, with the help of Deep Machine Learning can be made able to perform all the tasks as soon as the visually impaired person opens up the mobile application. It would navigate the person throughout the device where other applications are stored and has many possibilities

II.FLOW CHART

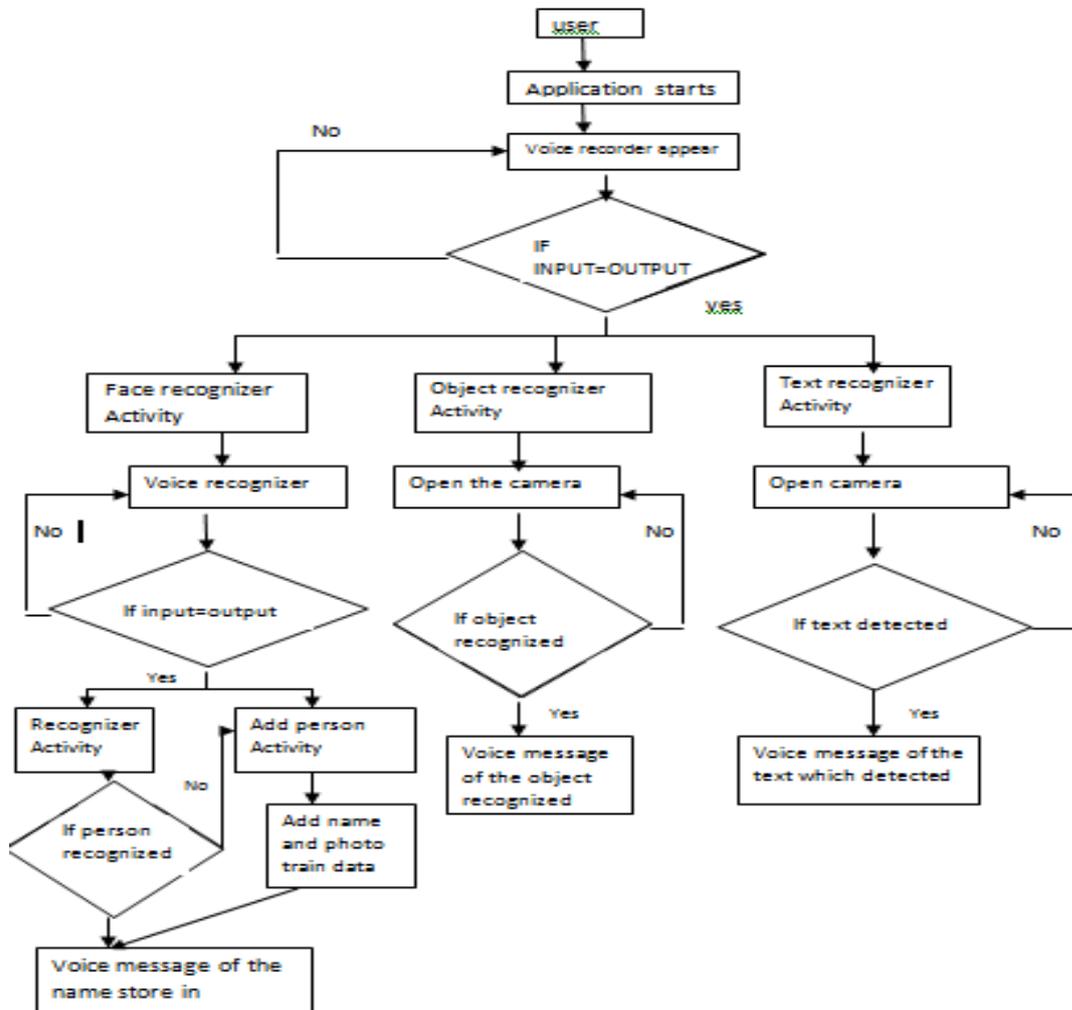


Fig-1.Flow chart

III. MODULES AND FUNCTION

III.I.Object Recognition module: Bazel is the primary build system for TensorFlow. *Bazel* is an open source variant of Google's in-house *build tool*. Bazel uses an abstract, human-readable language to describe the build properties of the project at a high semantical level. The bazel **Loads** the BUILD files relevant to the target and **analyzes** the inputs and their dependencies, applies the specified build rules, and produces an action graph. **Executes** the build actions on the inputs until the final build outputs are produced. Since all previous build work is cached, Bazel can identify and reuse cached artifacts and only rebuild or retest what's changed. Object recognition is based on **Feature Extraction**. They extract features from the input images and use these features to determine the class of the image. Tensors are just multidimensional arrays, an extension of 2-dimensional tables of data with a higher dimension. Implementations of the SURF, HOG, and FHOG feature extraction algorithms.

III.II.Face Recognition Module:Dlib is a modern C++ toolkit containing machine learning algorithms and tools for creating complex software to solve real world problems. It is used in both industry and academia in a wide range of domains including robotics, embedded devices, mobile phones, and large high performance computing environments. Dlib's open source licensing allows you to use it in any application, free of charge. Different Machine Learning Algorithms Reduced-rank methods, General

purpose multiclass classification tools, Structural SVM tools for sequence labeling Structural SVM tools for solving assignment problems „Conventional SMO based Support Vector Machines for classification and regression.

III.III. Text Recognition: Text recognition is the process of detecting text in images and video streams and recognizing the text contained there in. Once detected, the recognizer then determines the actual text in each block and segments it into lines and words. The Text API detects text in Latin based languages (French, German, English, etc.), in real-time, on device. The Text Recognizer segments text into blocks, lines, and words. Roughly speaking: a **Block** is a contiguous set of text lines, such as a paragraph or column, a **Line** is a contiguous set of words on the same vertical axis, a **Word** is a contiguous set of alphanumeric characters on the same vertical axis

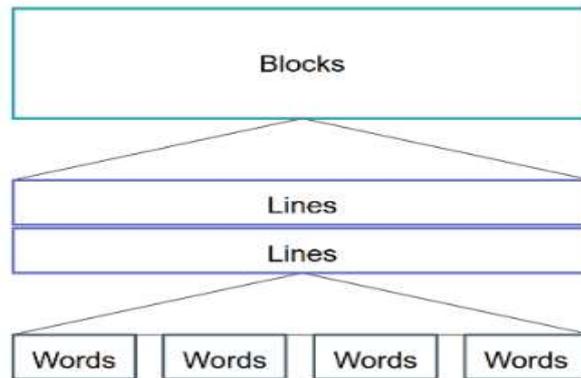


Fig-2. Flow Diagram for text recognizer

IV. ALGORITHM DEVELOPEMENT

IV.I.Camera Enabling: As we are having three modules ie. Face object and text recognizing all three works on different modules and hence each of them have their individual algorithm for enabling camera.

CameraView- CameraView is a well documented, high-level library that makes capturing pictures and videos easy.

- 1]Add<user permission for camera in Mainifest.
- 2]Import CameraView library of Dilib open source in java file.
- 3]Create a variable for CmeraView.
- 4]Link the Variable of cameraView from java file to the xml file.
- 5]Call for the camera permission and external storage.
- 6]Create a method name check permission.
- 7]Create a arraylist of string.
- 8]check weather permission granted.
- 9]Ture is return by metod to the list camera enables.

Camera2BasicFragment-The android.hardware.camera2 package provides an interface to individual camera devices connected to an Android device. This package models a camera device as a pipeline, which takes in input requests for capturing a single frame, captures the single image per the request, and then outputs one capture result metadata packet, plus a set of output image buffers for the request.

- 1]Add<user permission for camera in Mainifest.
- 2]Import Camera2 hardware and TextureView(which handles serval lifecycles) library in java file.

- 3] Create four methods of TextureView ie onSurfaceTextureAvailable, onSurfaceTextureChanged, onSurfaceTextureDestroyed, onSurfaceTextureUpdated.
- 4] call the openCmera(having two parameters width and height) method in onSurfaceTextureAvailable.
- 5] Create Camera Device which is useful when cameradevice changes its states
- 6] Call its three methods onOpen, onDisconnected, onError
- 7] In onOpen method release the lock first then set the cameraDevice to currenCameraDevice and create the preview.
- 8] In onDisconnected release the lock close the currentCameraDevice and make it to null.

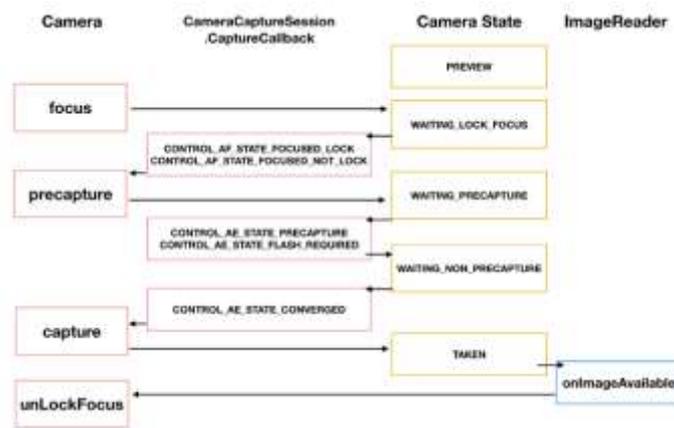


Fig-3. Block Diagram for camera enabling

IV.II. Text Recognize: This algorithm processes images and determines what text appears within them.

- 1] Create object TextRecognizer.
- 2] Check weather it is working.
- 3] call the detector process which reads text staright from the camera.
- 4] Create the sparseArray of TextBlocks .
- 5] call the handler and store the items into the sparseArray which are read by the detector.
- 6] Set the textview in xml file wit the items stored in sparseArray.

IV.III. Face Recognize : Dilib open source library is used. This application recognize the person and provide the name along with it .Also if the person data is not present the user is allowed to store information .We have to Algorithm use in this first is face recognition and second Adding person.

- 1] Create a TextToSpeech variable.
- 2] Call the voice recognizer.
- 3] Recognise is identified move to mOnClickListner method
- 4] Create FaceRec variable
- 5] Call an AsyncTask class and create Six methods for it.
- 6] Create an object which is use for storing images.
- 7] Create two folders in drawale → raw for face shape predictor and for face descriptor.

- 8]Create an arraylist which stores the name
- 9]Call for the camera permission
- 10]compare the name and match found then speak the message
- 11]Add person is identified move to different activity

IV.IV. Add Person

- 1]Call the recognizer.
- 2]Get the text and store in edit text field.
- 3]Call for recognizer again.
- 4]Take photo identified call the ImageCapture method.
- 5]Create a Method selectimage.
- 6]call the PackageManager and ask for the camera permission.
- 7]Create onActivityResult method which is called when the procedure starts.
- 8]Once camera activity started create a bitmap for the set of images which are to be captured
- 9]Get from gallery identified call for the media store
- 10]Once gallery activity started create a bitmap to store image which we select from the media store.
- 11]Call the AsyncTask class and the three methods .
- 12]set the message to detecting face in onPostExecute method.
- 13]Create a list for the faces to detect and store in an array
- 14]Return the message.

IV.V. Object Recognize: Using tensorflow object recognition is being carried out. The name and matching percentage of the object is displayed on the screen.

- 1]Create a class ImageClassifier.
- 2]Create an array in which images are stored.
- 3]Create a ByteBuffer which holds the images to feed into tensorflow.
- 4]call the inbuilt function comparator.
- 5]Compare the mapEntry of both the values
- 6]Initialize the image classifier.
- 7]Classify frame from preview stream.
- 8]Apply lowpassFilter at each stage.
- 9]copy the last stage filter output.
- 10]Close the Tensorflow to release resource.
- 11]Read labels list from the assets .

- 12]Memory map model file into assets.
- 13]Write Image data into ByteBuffer.
- 14]Print top k-labels
- 15]Call for the TextToSpeech and return the labels.

IV.VI. Text To Speech: Text to speech is becoming an integral part of many content driven mobile APPs. If your APP is a tutorial or content heavy, then it is high time you think about adding the Text-To-Speech (TTS) feature to it. Voice input based command and text to speech is a trend now. In this tutorial, we will explore how “Text-To-Speech(TTS)” feature can be implemented in an Android APP.

- 1) create an object of TextToSpeech class provide parameters to set the language, speed and pitch.
- 2) set speed of speech using textToSpeech.setSpeedRate() method
- 3) Setting up the pitch of speech
- 4) setting up the speech parameters, we have to invoke OnClickListener when button is clicked.

V. RESULT AND CONCLUSION

This is an innovative System for visually impaired people and acts as a voice assistant for them. This Application helps visually challenged people for detection and recognizing of face and object present in surrounding by making use of different custom layouts and using text to speech. The converted voice will help them to identify the object whose image is captured by them directly as input to software and can listen it. In this way, we can help to blind people in some sort of their activities such as identifying the objects, for example this application can also be used in day to day activities supposed if the person is not able to find his/her belongings he/she she can just start randomly take pictures and get information hence using this object is detected

V.I. GUI DESIGNS



Figure4.Splash screen GUI Design



Figure5.Three Modules Option



Figure6.Object Recognition Module

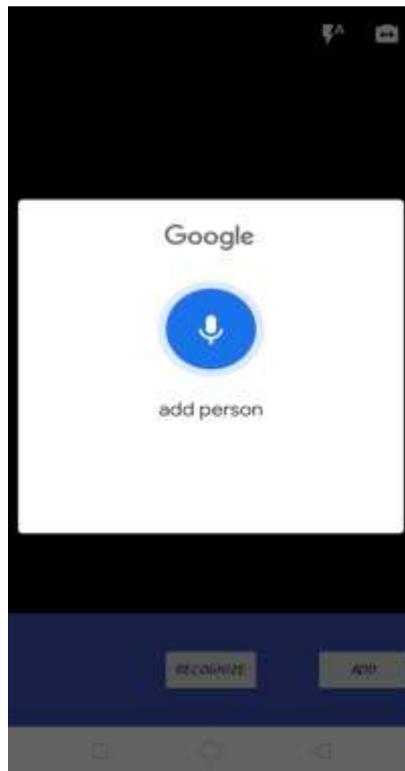


Figure7.Face Recognition Module

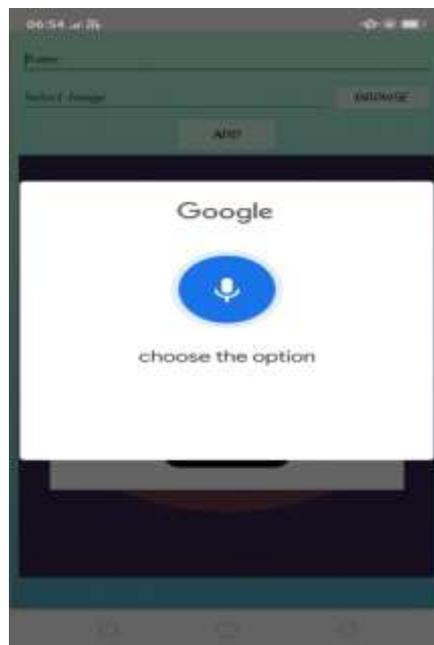


Figure8.Adding and Training Person face

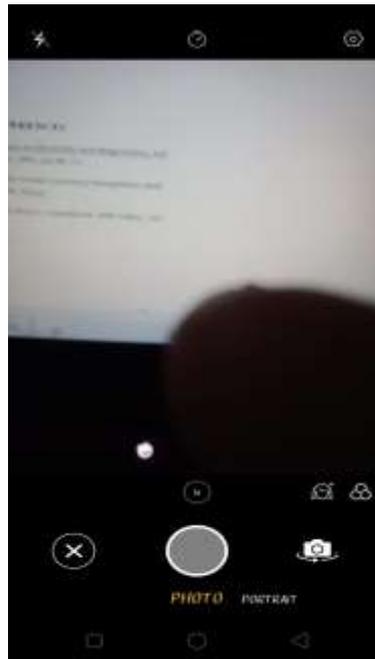


Figure9.Taking Photo for Training purpose



Figure10.Text Recognition Module

ACKNOWLEDGEMENT

It is pleasant task to express gratitude to all those who contributed in many ways to this project. First of all, we would like to thank our guide Prof. Deepti Vijay Chandran. This work would not have been possible without her guidance, support and encouragement . The application that we are trying to develop will advance the way of people that are going to use the computing environment. The application being developed would lead to a better solution for communication and daily living

for visually impaired people. It would be a fresh approach in the computing environment, and will lead to a successful and reliable approach.

REFERENCES

- [1] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford:Clarendon, 1892, pp.68–73.
- [2] An-Android-Application-for-Indian-Currency-Recognition-and-Authentication-for-Blind.pdfK. Elissa.
- [3] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [4] D. Pascolini and S. P. Mariotti, "Global estimates of visual impairment: 2010," British Journal Ophthalmology, 2011.
- [5] R. Jafri and H. R. Arabnia, "A Survey of Face Recognition Techniques," Journal of Information Processing Systems, vol. 5, pp. 41-68, 2009.

BIOGRAPHIES



Akash Bhardwaj is pursuing Bachelor degree (B.E.) in Computer Engineering from Smt. Indira Gandhi College of Engineering, Navi Mumbai.



Akash Prajapati is pursuing Bachelor degree (B.E.) in Computer Engineering from Smt. Indira Gandhi College of Engineering, Navi Mumbai.



Yashashree Bapat is pursuing Bachelor degree (B.E.) in Computer Engineering from Smt. Indira Gandhi College of Engineering, Navi Mumbai.

Prof. Deepti Vijay Chandran, M. E in Computer Engg. From Mumbai University. She is Asst. Professor with Smt Indira Gandhi College of Engg. Mumbai University, having 24 years of experience . Her area of interest includes Artificial intelligence, Data Science and Machine learning.