UNIVERSITY OF ENGINEERING AND TECHNOLOGY

PESHAWAR PAKISTAN

REAL WORLD MAPPING SYSTEM FOR VISUALLY IMPAIRED PEOPLE

By

Shah Raza 18PWCSE1658

Muzamal Ali 18PWCSE1665

Hudabia Faiz 18PWCSE1729

The undersigned certify that they recommend to the UET Peshawar for acceptance of this thesis for the fulfillment of the requirements for the degree stated.

Signature: Supervisor: ______________________

Engr.Nania Said

Signature:
Final Year Project Coordinator: ______________________

Dr. Muniba Ashfaq

Signature: Chairman: ______________________

Dr. Nasir Ahmad

Date: ______________________

Real World Mapping System for Visually Impaired People

By

Shah Raza

Muzamal Ali

Hudabia Faiz

A Thesis

Submitted to the UET Peshawar

as a Requirement for the Degree of

BACHELOR OF SCIENCE

DEPARTMENT OF COMPUTER SYSTEMS ENGINEERING

UNVIVERSITY OF ENGINEERING AND TECHNOLOGY,

PESHAWAR, PAKISTAN

JUNE 2022

DEDICATION

Every challenging project need commitment as well

as supervision of elders, particularly those who are

very close to our hearts.

Our courteous efforts we dedicate to our beloved

Father and Mother,

Whose affection, love, support, and prayers

of day and night Make us able to get such

success and honor

Along with all hardworking, devoted, inspiring, and respected

Teachers

ACKNOWLEDGEMENTS

First and foremost, praises and thanks to God, the Almighty, for His blessings throughout our final year project work to complete the A-EYE App (AI-Based Mobile Application for visually impaired People).

We all are greatly indebted to our FYP supervisor Engr. Naina Said for guiding us with a proper direction, guiding us professionally, giving us confidence from the very start, and moral support in many ways during the FYP period. She has taught us the methodology to tackle the FYP and conclude the work as clearly as possible. For us, it was a great honor to work and study under her guidance and friendly behavior.

We would also like to express our sincere gratitude to the University of Engineering and Technology Peshawar for letting us fulfill our dream of being a student here. We would also like to thank our Department (Computer Systems Engineering) for allowing us to write an undergraduate thesis.

Apart from these, we are all extremely grateful to our parents, brothers, sister, and beloved friends for their kindness, guidance, caring, and sacrifices in educating and preparing us for our bright future.

ABSTRACT

A-EYE provides a platform where the visually impaired users of our app can easily do their daily work without depending upon their sighted friends or family members. The App has voice instructions feature that helps the users navigate to different parts of the App without any difficulty. A-EYE provides three main features: Object Detection, Face Recognition, and Currency Recognition (Pakistani Currency). Apart from these main features, the App also has an emergency shake feature which sends the location of the user to their emergency contacts when the user shakes their mobile phone.

TABLE OF CONTENTS

Chapter Title **Page**

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
API	Application Programming Interface
App	Application
DB	Database
ML	Machine Learning Model
UI	User Interface
GUI	Graphical user interface
PKR	Pakistani Rupees
MSVI	Moderate and Severe Visual Impairment
COCO	Common Objects in Context
YOLO	You Only Look Once
VGG	Visual Geometry Group
TF	TensorFlow

Chapter 1

Introduction

Visually impaired people face a lot of difficulties in their daily life. Different technological intervention tries to assist them to face these challenging difficulties. Independent mobility in an unfamiliar environment is very difficult for people who are visually impaired. To keep a track of their routine environment, they usually rely on their sighted friends and relatives for assistance. The main idea of the project is to help these visually impaired people perform their daily activities without depending on someone else. We are proposing a real-time-based vision system using object detection, face recognition, and currency recognition for visually impaired people's assistance. A deep learning model is trained with multi-class indoor and outdoor objects which are highly relevant to visually impaired people. To bring more robustness to the training images, training images are augmented and manually labeled. The main purpose is to come up with a low-cost model which can detect and recognize surrounding obstacles and try to estimate the risk to the user, to facilitate visually impaired people in the challenging environment. Similarly, Face recognition and Currency Recognition also have their own models. We have tried to keep the accuracy of each model as high as possible.

1.1 Problem Statement

Visually impaired people encounter challenges and are at a disadvantage as visual information is what they lack the most. The visually handicapped can be helped with the use of innovative technologies. The concept is to make an Android and IOS mobile app that includes features such as voice assistance, Object Detection, currency recognition, face recognition, etc. The software can recognize items in the environment using different ML algorithms and inform the user using voice instructions.

1.2 Project Objectives

We present an App “A-EYE” to help ease the burdens of visually impaired people. It will significantly increase their productivity and perform their daily tasks easily. On other hand, it can also be very helpful for normal users as well. Here are some basic objectives:

- The implemented application should be user-friendly for anyone to use.
- Application should be used to recognize objects, faces, and currency notes.
- To bring ease and comfort to the life of visually impaired people through this system.

1.3 Project Motivation

Following are the motivational factors behind this project:

- In 2015, there were an estimated 253 million people with visual impairment worldwide. Of these, 36 million were blind and a further 217 million had moderate to severe visual impairment (MSVI). The prevalence of people that have distance visual impairment is 3.44%, of whom 0.49% are blind and 2.95% have MSVI. These people need to be provided with a system that can make their lives easier.
- In Pakistan, an average person cannot afford to have the luxury of high-tech expensive resources to help with their vision problems. We aim to develop a system that is very easy to access and much more affordable.

1.4 Approach

For this App, we are mainly using the different ML models for each feature that is included in our App. All of these models are then converted into their tflite versions and then they are deployed to the mobile app using flutter's tflite_flutter package. The inputs to these models are preprocessed using the tflite_flutter_helper package which has a lot of built-in functions that help us with image processing. We have used the local storage of the mobile phone to store the face embeddings and emergency contacts of the users. We preferred local storage over some databases because that way the users will not need an active internet connection for the App to work properly. All of the models are also in the assets of the Application. So, the App works completely offline.

1.5 Thesis Outline

Following is our thesis organization:

- Chapter 1 introduces the idea behind A-EYE, its motivation, objectives, and implementation approach briefly.
- Chapter 2 discusses the literature review.
- Chapter 3 discusses our workflow and methodology.
- Chapter 4 presents our implementation details.
- Chapter 5 concludes our work.

Chapter 2

Literature Review

2.1 Artificial intelligence (AI)

As we know that everything is revolutionized by Artificial intelligence (AI). AI is a technology with which we can create intelligent systems that can simulate human intelligence. The AI system does not require to be pre-programmed, instead of that, they use algorithms that can work with their own intelligence. It involves machine learning algorithms such as Reinforcement learning algorithms and deep learning neural networks. Web applications and software with artificial intelligence are primary components developed by AI development companies. Deep learning systems, data analytics platforms, machine learning algorithms, and conversational tools are examples of these projects. The overview of our project is mainly based on the application of AI. It means that there must be AI-based deployed components that will enhance the effectiveness of the required task.

Figure 2.1 Artificial Intelligence

2.2 Machine Learning (ML)

ML is about extracting knowledge from the data. It can be defined as ML is a subfield of AI, which enables machines to learn from past data or experiences without being explicitly programmed. ML enables a computer system to make predictions or take some decisions using historical data without being explicitly programmed. ML uses a massive amount of structured and semi-structured data so that an ml model can generate accurate results or give predictions based on that data. ML works on algorithms that learn on their own using historical data. It works only for specific domains such as if we are creating a machine learning model to detect pictures of dogs, it will only give results for dog images. But if we provide new data like cat image then it will become unresponsive. Machine learning is being used in various places such as for online recommender systems, Google search algorithms, Email spam filters, Facebook Auto friend tagging suggestions, etc.

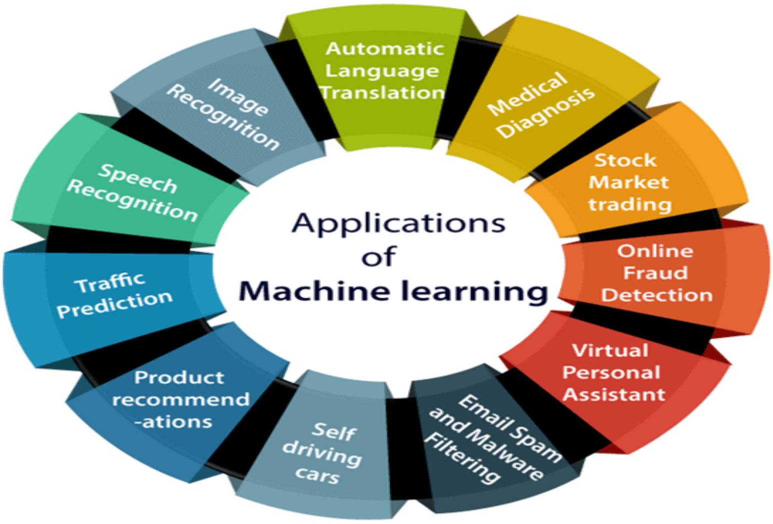

Applications of Machine learning
Automatic Language Translation
Medical Diagnosis
Stock Market trading
Online Fraud Detection
Virtual Personal Assistant
Email Spam and Malware Filtering
Self driving cars
Product recommend -ations
Traffic Prediction
Speech Recognition
Image Recognition

Figure 2.2 Applications of ML

2.3 Deep Learning

Deep Learning is a method of AI that resembles human brain function while data analyzing and patterns creating, for decision making. This data uses sound, text, and images. It is a subset of Artificial Intelligence and machine learning is based on ANN where the learning can be supervised or unsupervised from shapeless data. It is also known as a deep neural network. Deep learning uses the concept of neural networks like human brain neurons. Neural networks contain processing elements (neurons) that are highly interconnected between input and output, just like humans, they learn from practice. These layers do recognize features and manipulate them in a sequence of stages just as our brain does. A mathematical function of a neuron in a neural network that collects the data and then it classifies that data into a specific structure. Deep Learning and neural networks tend to be used interchangeably in conversation, which can be confusing. As a result, it’s worth noting that the “deep” in deep learning is just referring to the depth of layers in a neural network. A neural network that consists of more than three layers which would be inclusive of the inputs and the output can be considered a deep learning algorithm. A neural network that only has two or three layers is just a basic neural network.

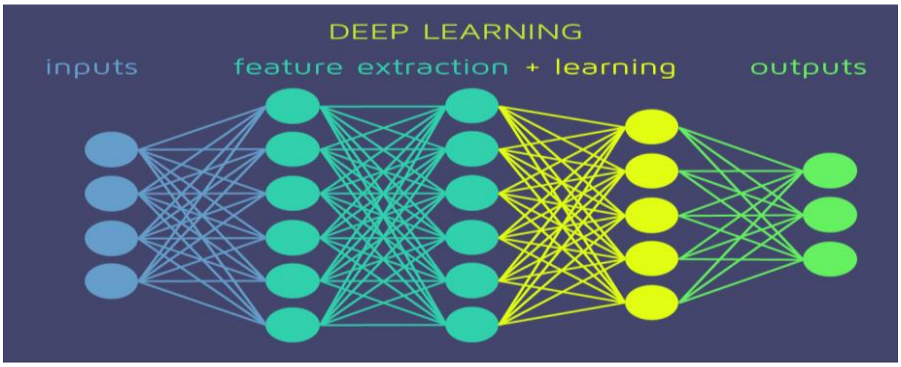


Figure 2.3 Layers of Deep Learning

2.4 Related Work

Seeing AI is a Microsoft research project that brings together the power of the cloud and AI to deliver an intelligent app designed to help you navigate your day. Point your phone's camera, select a channel, and hear a description of what the AI has recognized around you. Seeing AI can speak short text as soon as it appears in front of the camera, provide audio guidance to capture a printed page, and recognizes and narrates the text along with its original formatting. The app can also scan barcodes with guided audio cues to identify products, recognize and describe people around you and their facial expressions, as well as describe scenes around you using the power of AI. An ongoing project, the latest new ability to be added to Seeing AI's roster is identifying currency bills when paying with cash and describing images in other apps such as your photo gallery, mail, and Twitter [1].

Noteify is an artificially intelligent Indian currency detection app made for the visually impaired to check whether they have been handed the right amount of money and thereby, ensuring that they have not been cheated upon, giving the output as computer-generated audio [2].

Lumos is another app that has useful features for both visually impaired and deaf people.

2.5 Proposed Approach

We propose an application for both Android and iOS operating systems. We will develop the application in flutter and the models will be developed using Python and TensorFlow as shown in Figure 2.4. We have started work on the development of an android application and later same code will be simulated for iOS. We will be needing local storage for storing face embeddings and emergency contacts. By local storage, we mean the phone's internal storage. Every app gets a dedicated folder in the phone's internal memory when it is installed, this storage can be used to store data related to that app. We are making use of this local storage to make the App work completely offline.

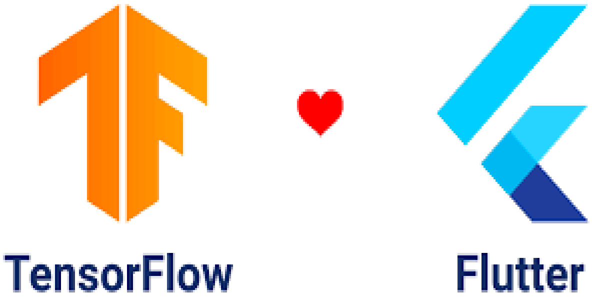


Figure 2.4 Project Developing Languages

The most common tools and services we will be using are the following:

1. Local storage: Database.
2. TensorFlow: For developing ML models.
3. Google’s ML Kit: For detecting faces.
4. Flutter TTS: For giving voice instructions to the users.

Python will be used for developing machine learning algorithms that will be used to recognize objects, faces, and Pakistani currency notes. Python is one of the versatile programming languages, covering a majority of fields such as web development mobile games, and different machine learning and deep learning field. Python will give access to excellent open-source libraries and frameworks for AI and machine learning (ML), accessibility, platform independence, and a large and cooperative community, making it distinct for machine learning and AI projects. Python provides access to different libraries for statistical and mathematical calculations, plotting, graph, and data visualization, which burnish its importance more for Machine learning use [3]

AI programming is different from the traditional software’s initiatives. To prefer AI is the technology stack, and more knowledge and mathematical skill to perform a machine learning project. Select a programming language for machine learning that is more robust and versatile, adaptable, and offers more libraries, and computational speed to implement your AI goals. Python has almost provided all these features with interactive tools and more attractive development environments. This is why Python is one of the most used languages for AI projects nowadays. Python community helps developers to be more productive and confident about the models they are developing. From the development perspective, Python's simplicity, and access to open-source libraries and different frameworks for AI and machine learning (ML) will make python a decent choice for development and maintenance. Flexibility, platform freedom, and a large community make it the best option for AI and ML model development. [4]

Chapter 3

Workflow and Methodology

This chapter gives an overview of our project giving details about the different aspects of the project, i.e., the machine learning part and application part of the project. Since our project mainly deals with Machine Learning Algorithms so it will be explained in-depth in the Implementation chapter.

3.1.1 Parts of the Project

Our project has the following six phases of development as shown in Figure 3.1.

1. User Interface (UI) development
2. Creating Machine Learning (ML) Models
3. Modifying UI for Application/Website
4. Involvement of Voice Instructions
5. Embedding ML in Application
6. Deploying Application

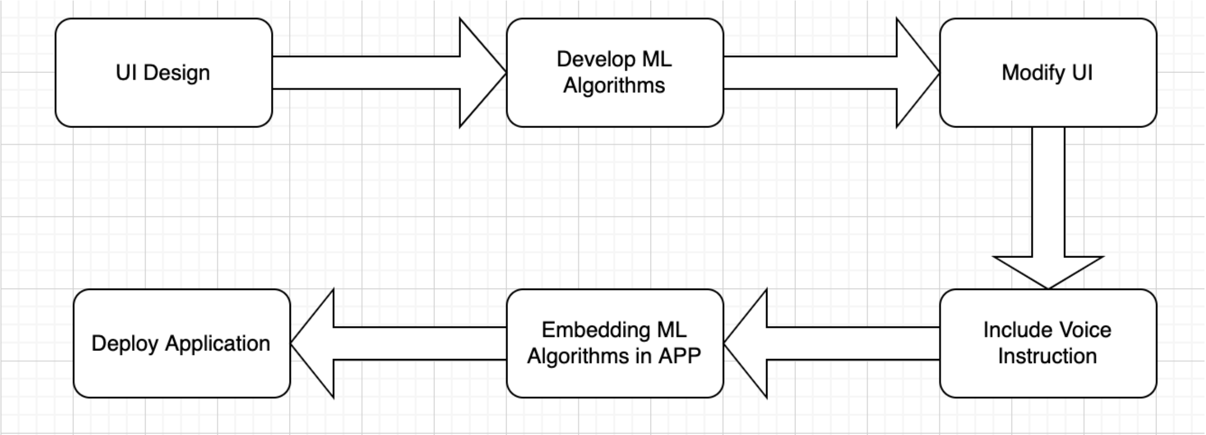


Figure 3.1 Project Workflow

3.1.2 User Interface (UI) Development

Before developing the front-end, we designed the basic screens of the user interface for A-EYE using Adobe XD. Figure 3.2 shows the basic prototype of our application.

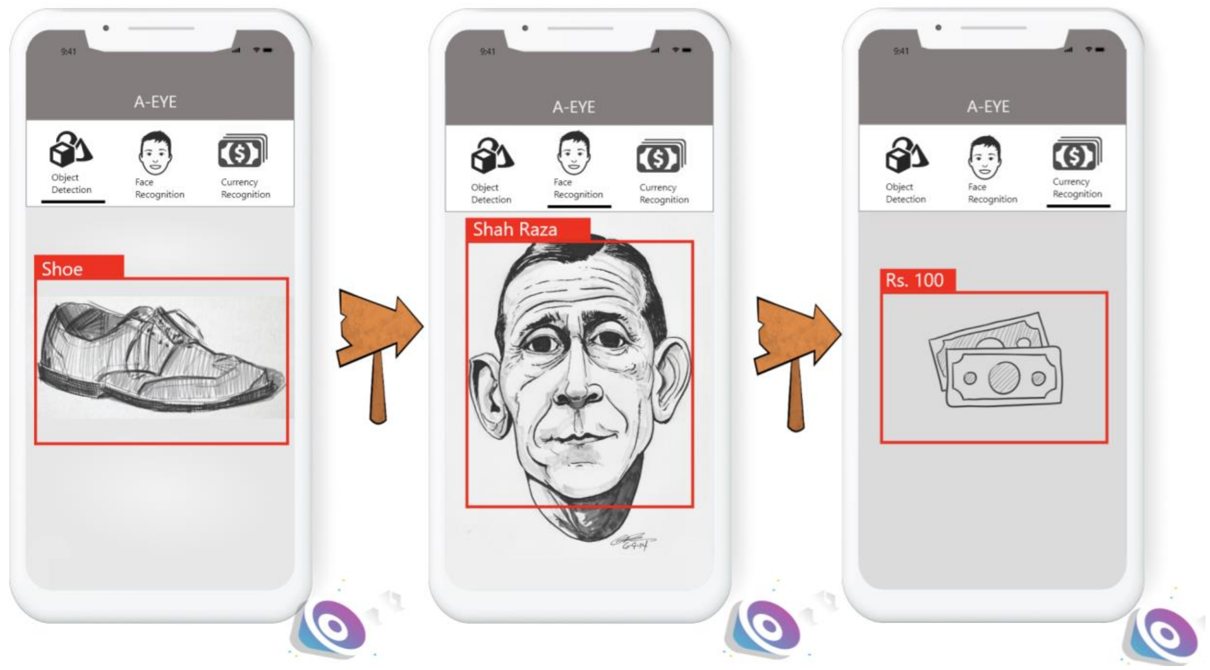
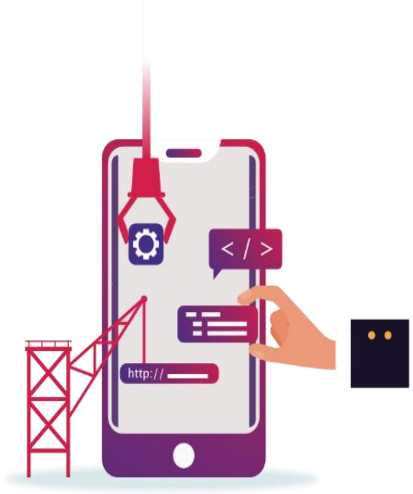


Figure 3.2 UI Screen Design

The complete prototype of the user interface design is done in Adobe XD. The XD allows us to create complicated user-friendly prototypes. Adobe XD offers device previews and online reviews of a prototype and integrates with creative cloud apps through Adobe XD's Creative Cloud Library support. The Adobe XD UI provides designers with the tools they need to accomplish their best work, from interface details such as iconography and typography to achieving the perfect balance of layout. Assets can be easily retrieved and stored in Adobe XD thanks to interactions with other creative cloud apps such as Photoshop and Illustrator [5].



3.1.3 Machine Learning (ML) Models

Machine learning can significantly help us with everyday tasks. By applying algorithms that train the machine itself, it can learn to differentiate between different objects, faces, and notes (PKR). Visually impaired users of our App can now experience the world like never before with the help of these ML models.

A-EYE is purely a Machine Learning App, which means that we provide Machine Learning as a Service. By using machine learning algorithms for the recognition of different things, our Machine Learning models will decrease the requirement for complex software programming.

Our machine learning solution will drastically reduce the programming time, making the app more efficient and useful.

As our project is AI-based, we have implemented ML models to make the lives of visually impaired people much easier. These ML models will help them recognize everyday objects such as a laptop, bottle, TV, Bed, Table, etc., and recognize the faces of their family members and friends. It will also help them with currency recognition.

The main goal of the ML models is to try and predict everything as accurately as possible because when it comes to problems like currency recognition, there is no room for less accuracy or errors. Three ML models are included in the assets of this App. We can say that these ML models are the life of this App.

3.1.4 Object Detection

Object detection is a computer vision technique that works to identify and locate objects within an image or video. Object detection draws bounding boxes around the detected objects. Object detection allows us to detect different objects in a given scene. Machine learning is a core part of our App. To enhance the user experience, it is necessary to recognize each object accurately. For this purpose, we used Machine Learning algorithms called Volo v4. Compared with the previous YOLOv3, YOLOv4 has the following advantages: It is an efficient and powerful object detection model that enables anyone with a 1080 Ti or 2080 Ti GPU to train a super-fast and accurate object detector. YOLOv4 is an object detection algorithm that is an evolution of the YOLOv3 model. The YOLOv4 method was created by Alexey Bochkovskiy, Chein-Yao Wang, and Houg-Yuan Mark Liao. It is twice as fast as EfficientDet with comparable performance. This YOLO algorithm is trained on the COCO Dataset. COCO is large-scale object detection, segmentation, and captioning dataset. It contains around 330,000 images out of which 200,000 are labeled for 80 different object categories.

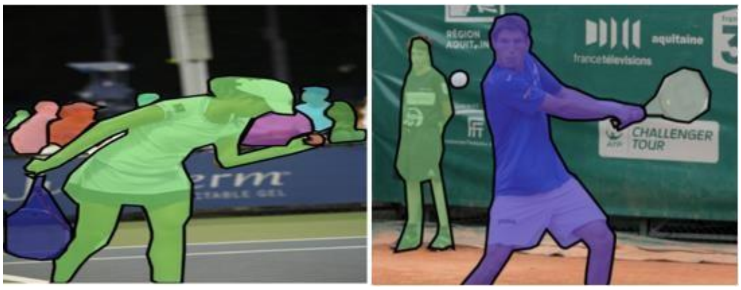
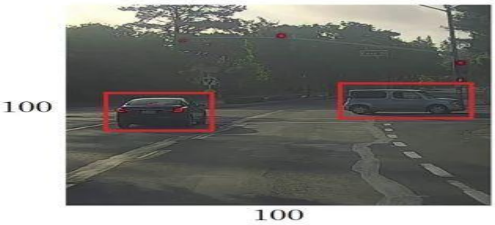


Figure 3.3 Object Detection

3.1.5 Working Principle of YOLO

Step One: YOLO first takes an input image



Step Two: The framework then divides the input image into grids (say a 3 X 3 grid)

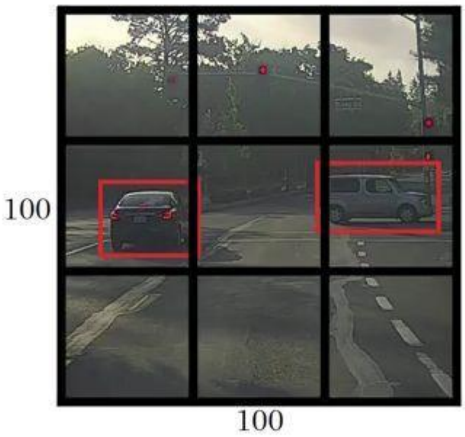


Figure 3.4 Working of YOLO

Step Three: Image classification and localization are applied on each grid. YOLO then predicts the bounding boxes and their corresponding class probabilities for objects (if any are found, of course)

3.1.6 Face Recognition

Face Recognition is a technology capable of matching a human face from a video frame against a database of faces. We have used the MobileFaceNet model for face recognition. MobileFaceNet is designed for achieving higher efficiency and accuracy for face recognition tasks in mobile and embedded devices MobileFaceNet is specifically designed for real-time face recognition. But before recognizing a face, we need to detect a face. So, for that purpose, we have used Google ML Kit's Face Detector. It is very fast and can be easily deployed in mobile Applications. The face embeddings of the users are stored in the local storage of the phone.

3.1.7 Working of FaceNets

FaceNet takes an image of the person's face as input and outputs a vector of 128 numbers which represent the most important features of a face. In machine learning, this vector is called embedding.

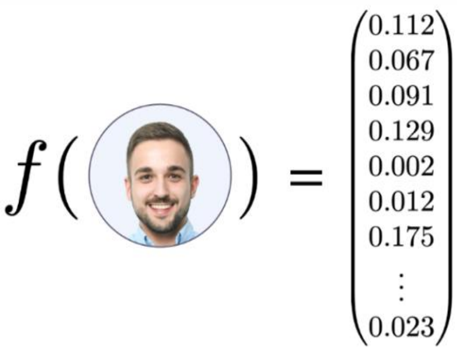


Figure 3.5 Face Embeddings

Embeddings are vectors and we can interpret vectors as points in the Cartesian coordinate system. That means we can plot an image of a face in the coordinate system using its embeddings. One possible way of recognizing a person on an unseen image would be to calculate its embedding, calculate distances to images of known people and if the face embedding is close enough to the embeddings of person A, we say that this image contains the face of person A.

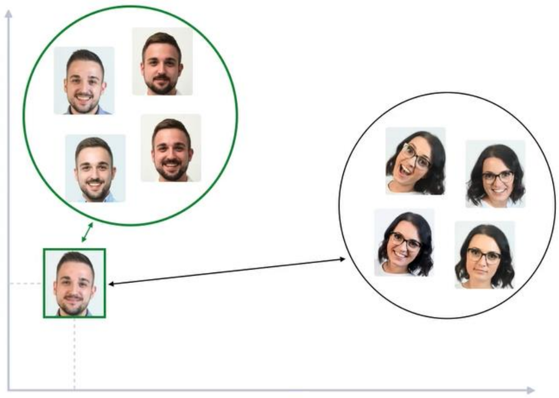


Figure 3.6 Plotting Face Embeddings

3.1.8 Currency Recognition

Currency Recognition is a technique that recognizes patterns in paper currency and tells the difference between different currency notes. We have used two datasets of Pakistani Currency. One dataset is taken from Kaggle, it has above 400 images for each currency note and around 3K images in total. The other one was provided to us by NCAI at the University of Engineering and Technology, Peshawar. We have trained the VGG19 model on these datasets for our currency recognition problem. VGG19 is an advanced CNN with pre-trained layers and a great understanding of what defines an image in terms of shape, color, and structure. VGG19 is very deep and has been trained on millions of diverse images with complex classification tasks. VGG19 is 19 layers deep. It is a very popular method for image classification due to the use of multiple 3×3 filters in each convolutional layer. An image of size 224×224 is inputted into this model and the model outputs the label of the object

(Currency notes in our case) in the image. Out of these 19 layers, 16 are used for feature extraction and 3 of them are used for classification.

3.1.9 Modifying UI for Application

The designed UI screens have been successfully implemented in the user interface development of the application as shown in Figure 3.7. We have redesigned a few things to make the user experience more enjoyable.

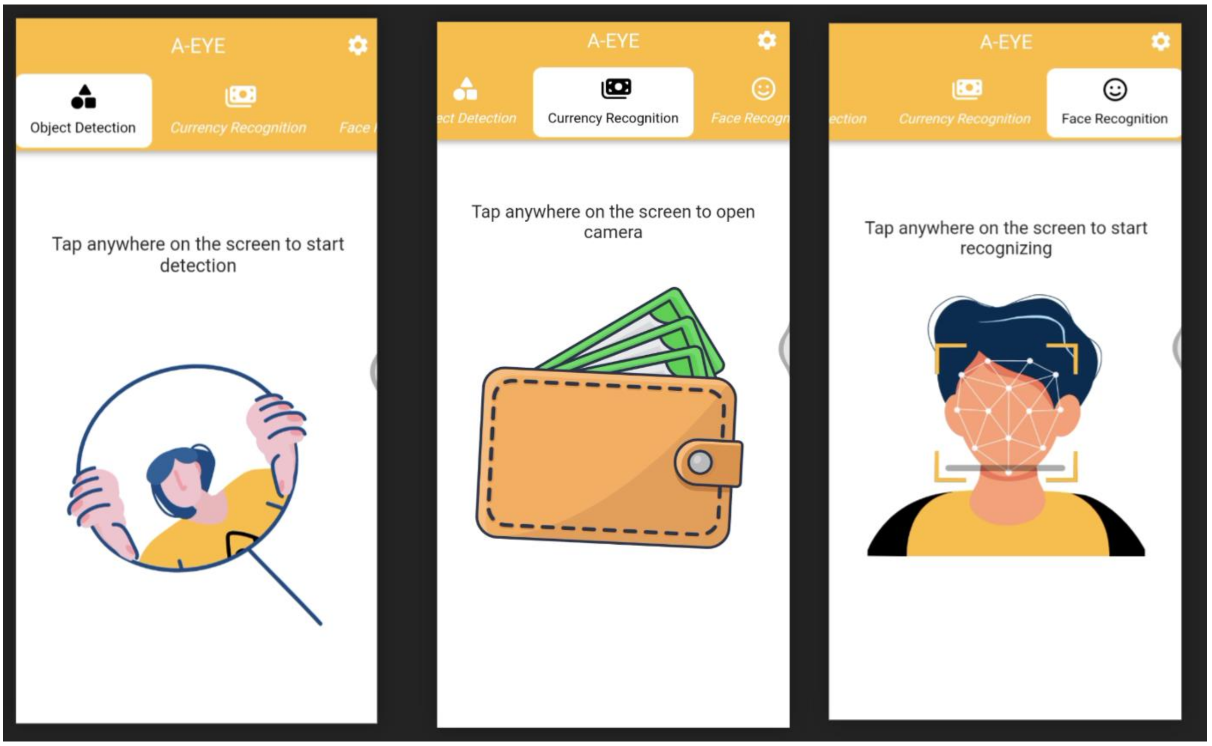


Figure 3.7 Modified UI for better experience

3.1.10 Voice Instructions

As our App's target audience is visually impaired, we need a medium through which we can easily tell them about the results that our ML models recognize and also help them navigate to different parts of the Application. For this purpose, we have used flutter's flutter_tts package. Flutter_tts is a flutter plugin for Text to Speech conversion. This plugin is supported on iOS, Android, Web, & macOS. We have used this plugin to help users navigate through the App and also tell them about the things that are detected.

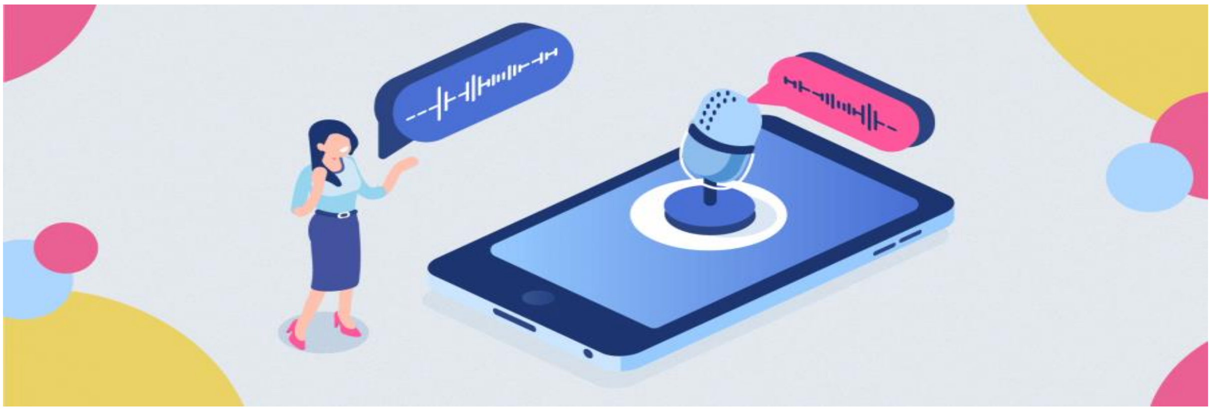


Figure 3.8 Voice Instructions

3.1.11 Embedding ML Model in Application

The trained ML models are embedded in the Application as shown in Figure 3.9. In this way, the Application will be able to recognize Objects, Faces, and Pakistani Currency Notes for visually impaired people and make them more independent of their sighted friends and family members. The Object Detection model will be able to classify 80 common objects in the COCO dataset. For the face recognition model to work, we need to first save a few faces in the phone's local storage. The model will compare the embeddings of an unseen image with the embeddings stored in the local storage. If any embeddings match the local database embeddings, then the model will recognize that face and the App will speak the name of that face. Similarly, the currency recognition model is trained for Pakistani Currency notes. Currently, the Pakistani Currency has seven notes. 10, 20, 50, 100, 500, 1000 and 5000 PKR. All seven of them can be easily classified by the ML model that we have embedded in our mobile Application.

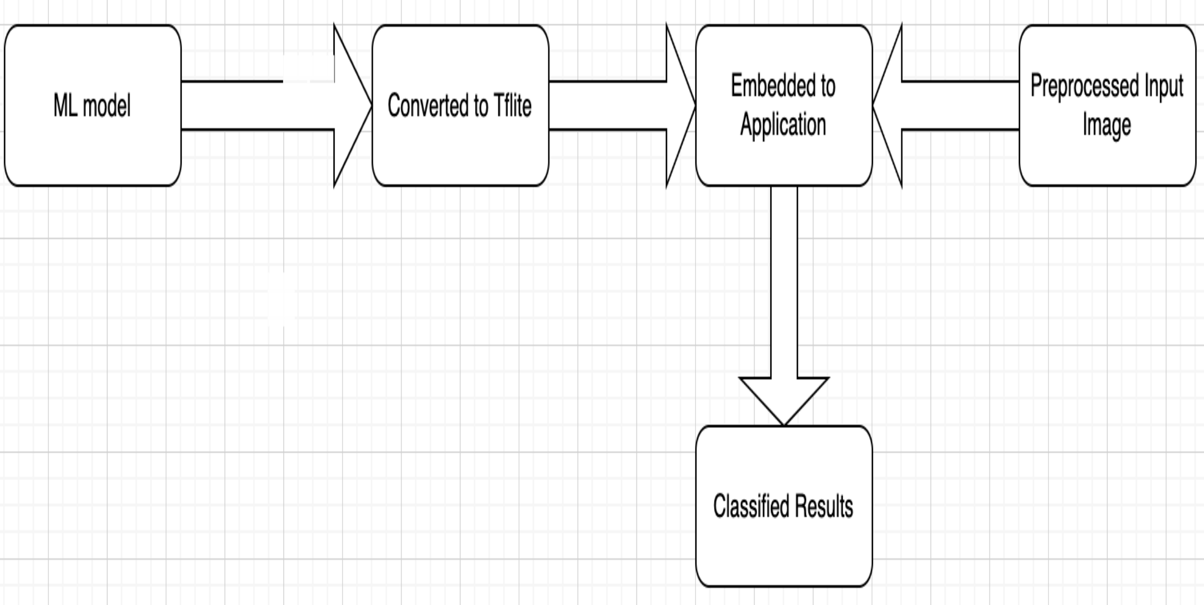


Figure 3.9 Embedding ML model in Application

3.1.12 Deploying Application

The process of getting a product code to your consumer is known as "deployment." It's significant since it affects both the speed with which a product responds to changes and the quality of those changes.

The make-up of an application contains different software packages in a specific environment, such as development or production, which is defined by application deployment. During development, an application passes through different physical phases. However, each application has unique, and our approaches are constantly improving, this is a standard process for designing mobile applications. The first step is the Idea that to feel the absence of something that is possible give them a shape of app or web, strategy is a phase where the team thinks of the need and requirements of the app, and design is giving a proper shape to their idea, development, in deployment developer start application development, and after launching phases are usually included in the mobile app development process [6]. Since flutter is a cross-platform framework, we can deploy our app in both the Google Play Store and Apple App Store.

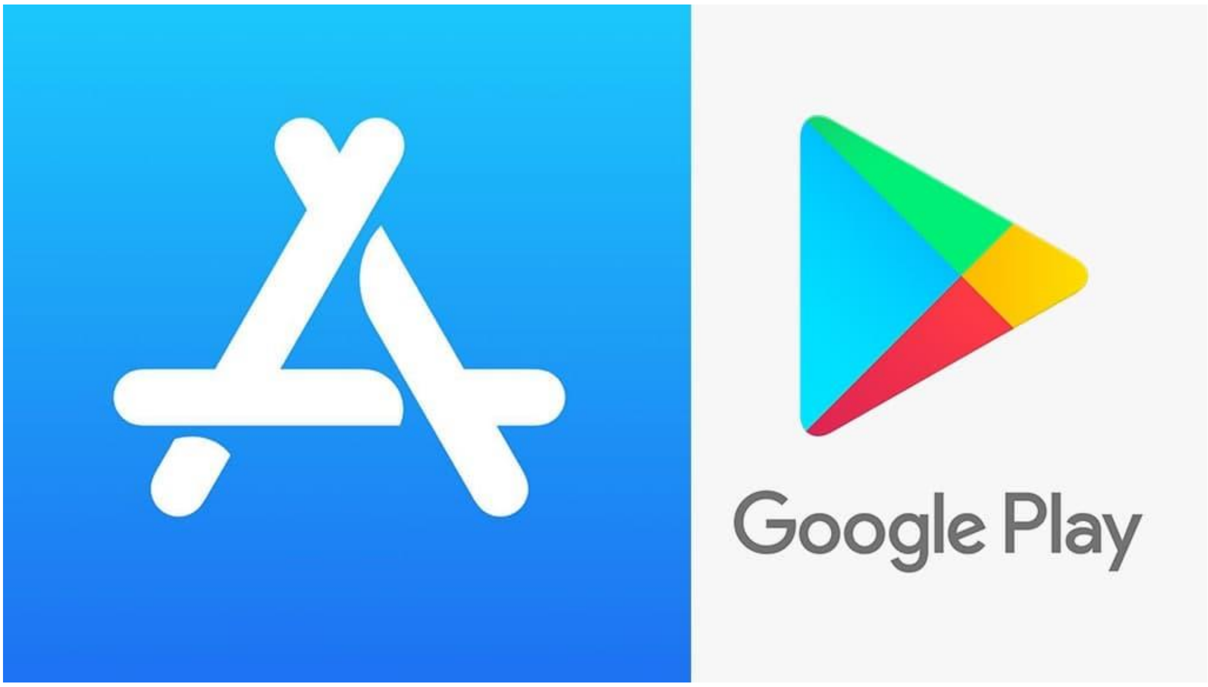


Figure 3.10 App Deployment

Chapter 4

Implementation

4.1 Overview

In this chapter, we are discussing the main implementation phases and the resource utilization while developing the application.

4.2 App development

This project is mainly focused on Application development. So, we will be mentioning every single resource that we are utilizing while building the mobile Application for this project.

4.2.1 Overview

We developed our project using the following three main resources.

- Flutter
- TensorFlow
- Google ML Kit

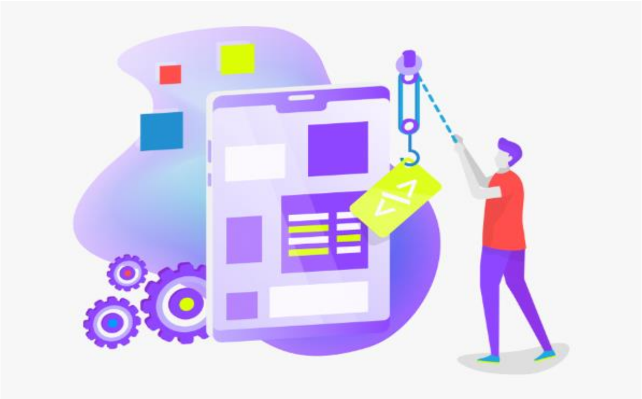


Figure 4.1 App Development

The mentioned resources are discussed in detail later on in this chapter.

4.2.2 Flutter

Flutter is one of the most powerful cross-platform application development platforms available. Since it's similar to native languages, has excellent performance, and has a plethora of off-the-shelf solutions, it's simple to learn for both native and cross-platform mobile application developers.

There are some reasons why we choose flutter to develop our project:

Flutter comes with a variety of widgets and ready-to-use solutions. This makes Flutter app creation simple and time-saving, which is ideal for companies who need their apps as soon as possible. There are currently six operating systems supported by the framework. With only one code base and a little tuning, a developer can release their app on six different platforms.

- Android
- iOS
- macOS
- Windows
- Linux
- Web

The code structure of flutter and native android is somehow much similar. The Flutter system was developed by google and provides the best features of the native android language. Hundreds of developers contribute to open-source technologies, resulting in a boom in cross-platform growth. The Flutter platform is one of the most rapidly developing of these innovations, and it is supported by Google, which means it has excellent documentation and is constantly evolving.

Flutter, like many other progressive languages, is built around widgets. Developers can see what they're doing with their code almost instantly with widgets, and they can save much time while designing simple UI designs for each screen and resolution. As Flutter has its widget toolkit, all of the components are natively rendered. This boosts performance and gives apps a more native feel.

4.2.3 Android Studio

Android studio is working on the idea of IntelliJ, which is used as an Official Integrated Development Environment (IDE) for the Development of Android App. In addition to IntelliJ's, Android studio's robust code editor and developer tools add additional capabilities and built-in features that improve performance when developing apps for Android devices. In addition, It is the official IDE for Android. It is primarily designed for Android to help you accelerate development and create high-quality apps for every Android device.

4.2.4 Python

Apart from the flutter, one of the main components of our project is Python. We have used Python for developing our Machine Learning Models. Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy-to-learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms and can be freely distributed.

Often, programmers fall in love with Python because of the increased productivity it provides. Since there is no compilation step, the edit-test-debug cycle is incredibly fast. Debugging Python programs is easy: a bug or bad input will never cause a segmentation fault. Instead, when the interpreter discovers an error, it raises an exception. When the program doesn't catch the exception, the interpreter prints a stack trace. A source-level debugger allows inspection of local and global variables, evaluation of arbitrary expressions, setting breakpoints, stepping through the code a line at a time, and so on. The debugger is written in Python itself, testifying to Python's introspective power. On the other hand, often the quickest way to debug a program is to add a few print statements to the source: the fast edit-test-debug cycle makes this simple approach very effective.

4.3 TensorFlow

TensorFlow is an end-to-end open-source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML-powered applications. TensorFlow offers multiple levels of abstraction so you can choose the right one for your needs. Build and train models by using the high-level Keras API, which makes getting started with TensorFlow and machine learning easy. If you need more flexibility, eager execution allows for immediate iteration and intuitive debugging. For large ML training tasks, use the Distribution Strategy API for distributed training on different hardware configurations without changing the model definition. Build and train state-of-the-art models without sacrificing speed or performance. TensorFlow gives you flexibility and control with features like the Keras Functional API and Model Subclassing API for the creation of complex topologies. For easy prototyping and fast debugging, use eager execution. TensorFlow also supports an ecosystem of powerful add-on libraries and models to experiment with, including Ragged Tensors, TensorFlow Probability, Tensor2Tensor, and BERT.

Figure 4.2 TensorFlow + Keras

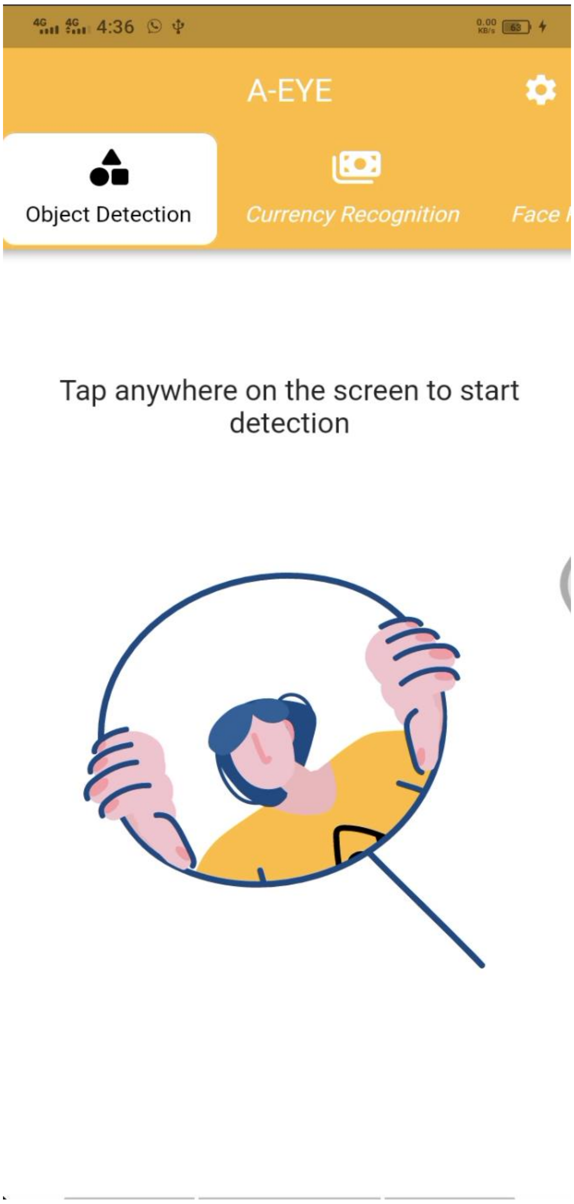
4.3.1 Automatic Volume Increase

When the app is opened, the volume of the phone is checked if it is less than 50% then the volume of the phone is automatically increased to 50%. This way the visually impaired person can listen to the volume instructions and results of the detected things.

4.3.2 Object Detection

The very first screen that appears on the app is Object Detection Screen. The app informs the visually impaired user about the current screen and how to navigate to another screen using Voice instructions. It also tells the user how to activate the Object Detection Model.

View:



Tap anywhere on the screen to start detection

Figure 4.3 Object Detection Home Screen

Voice Instructions: Object Detection Screen, tap anywhere to start detection or swipe left for Currency Recognition.

Camera View:

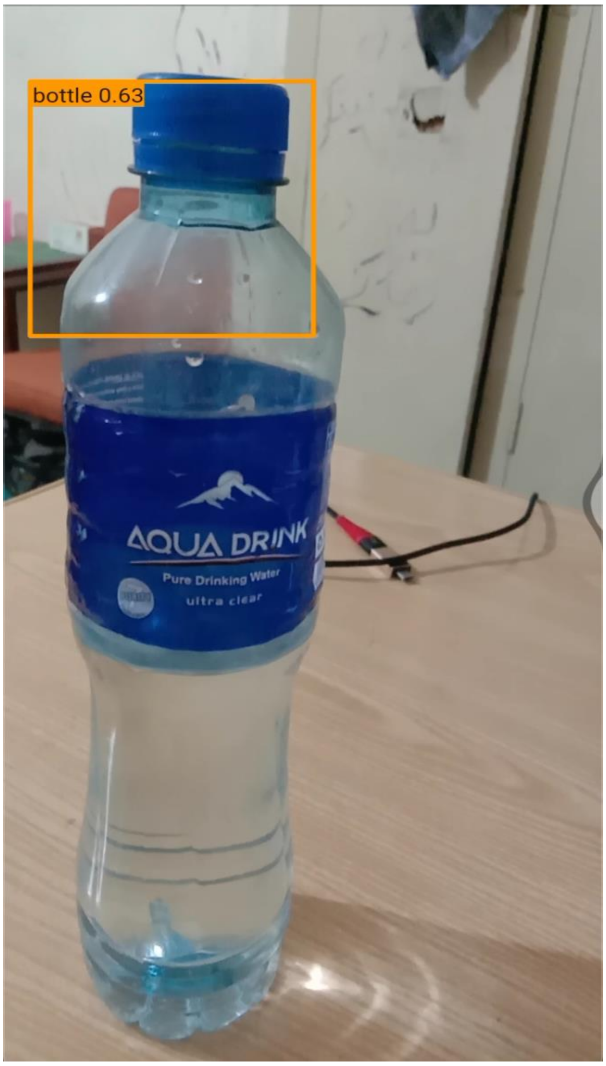


Figure 4.4 Object Detection Camera View

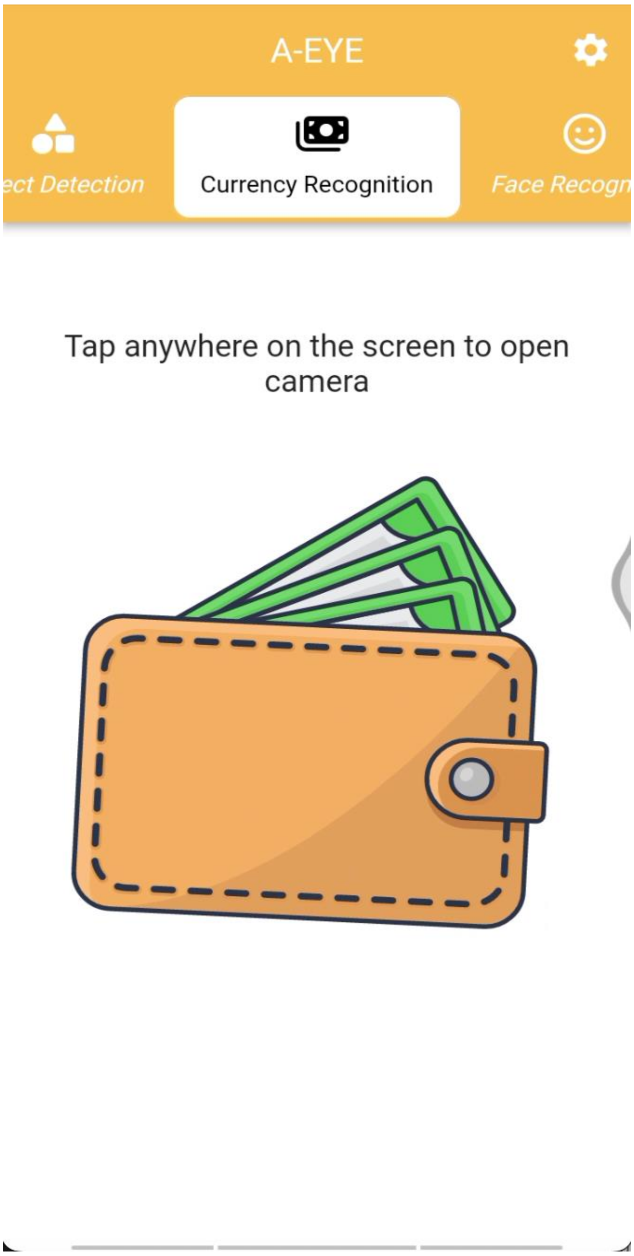
When the user taps on the Object Detection screen, a camera view is opened and it starts detecting the objects in front of the camera and the app starts to speak the results out loud to the visually impaired user.

Voice Instructions: Object Detection started, press and hold anywhere to exit Object Detection.

4.3.3 Currency Recognition

The second screen in the app is the Currency Recognition Screen. The app informs the visually impaired user about the current screen and how to navigate to another screen using Voice instructions. It also tells the user how to activate the Currency Recognition Model.

View:

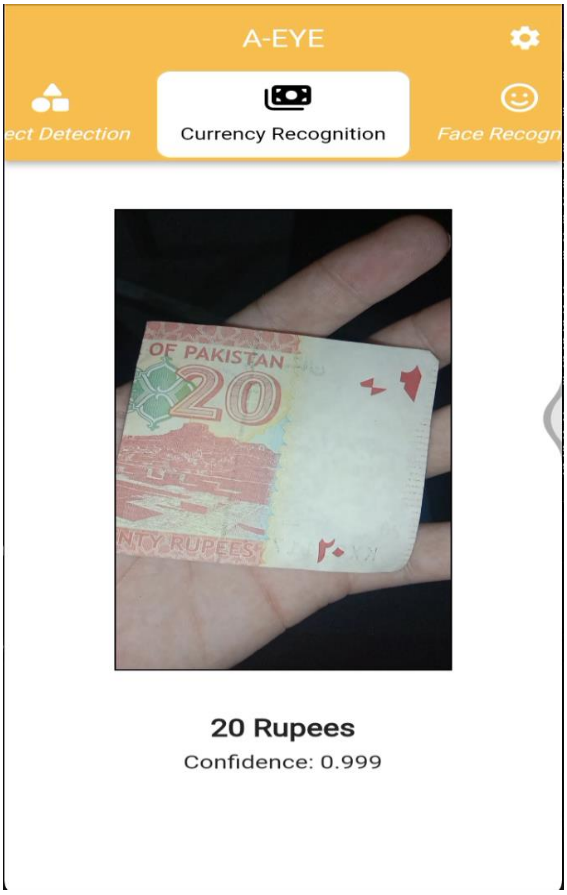


Tap anywhere on the screen to open camera

Figure 4.5 Currency Recognition Home Screen

Voice Instructions: Currency Recognition Screen, tap anywhere to open the camera or swipe left for Face Recognition.

Camera View:



A-EYE

ect Detection

Currency Recognition

Face Recogn

20 Rupees

Confidence: 0.999

Figure 4.6 Currency Recognition Camera View

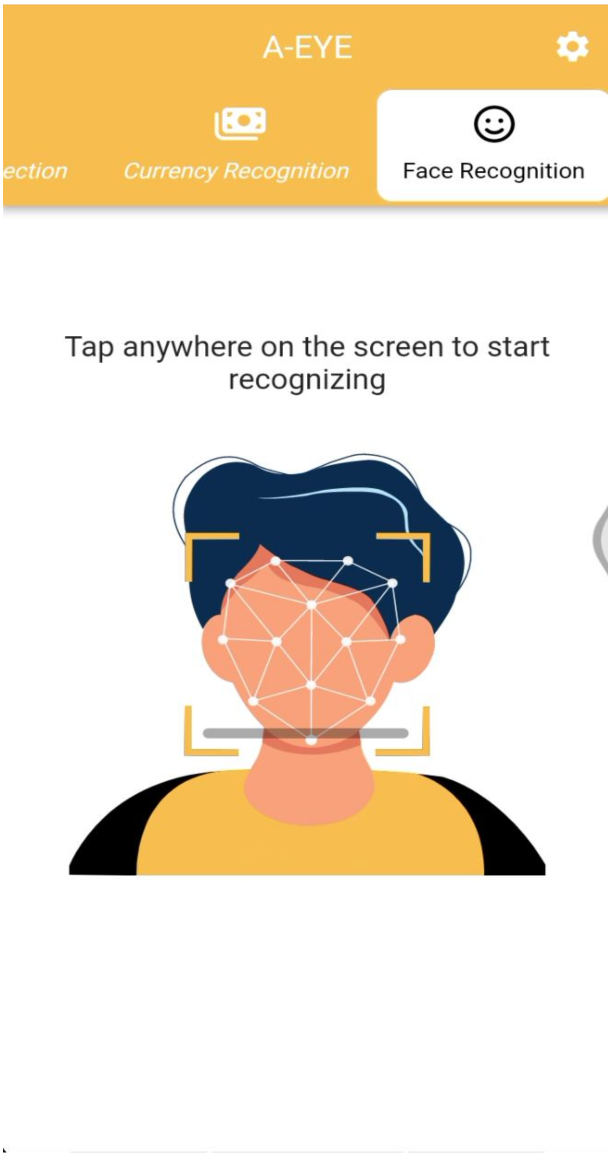
When the user taps on the Currency Recognition screen, a camera view is opened and the app asks the user to tap anywhere on the screen to take a picture of the currency note. When the picture is taken, the app comes back to the home screen with the results and speaks it out loud.

Voice Instructions: Tap to open the camera or long press to hear the results again.

4.3.4 Face Recognition

The third and final screen in the app is the Face Recognition Screen. The app informs the visually impaired user about the current screen and how to navigate to another screen using Voice instructions. It also tells the user how to activate the Face Recognition Model.

View:



Tap anywhere on the screen to start recognizing

Figure 4.7 Face Recognition Home Screen

Voice Instructions: Face Recognition Screen, tap anywhere to start recognizing or swipe right for Currency Recognition.

Camera View:

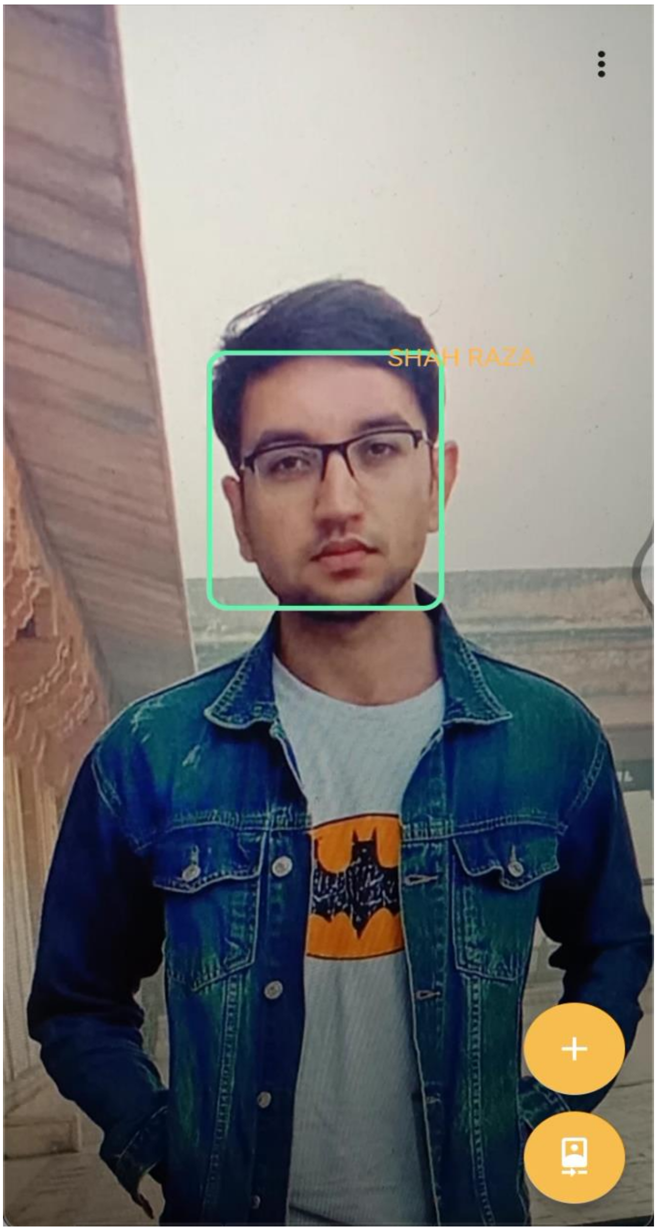


Figure 4.8 Face Recognition Camera View

When the user taps on the Face Recognition screen, a camera view is opened and the app starts to recognize faces in front of the camera and the app starts to speak the results out loud to the visually impaired user.

Voice Instructions: Face Recognition started, press and hold anywhere to exit Face Recognition.

4.4 Google ML Kit

ML Kit brings Google's machine learning expertise to mobile developers in a powerful and easy-to-use package. Make your iOS and Android apps more engaging, personalized, and helpful with solutions that are optimized to run on devices.

ML Kit's processing happens on-device. This makes it fast and unlocks real-time use cases like the processing of camera input. It also works while offline and can be used for processing images and text that need to remain on the device.

It has the following features:

- Barcode Scanning
- Face Detection
- Image Labeling
- Object detection and tracking
- Text recognition
- Digital Ink recognition
- Pose detection
- Selfie segmentation

We have used the ML kit for Face detection. Before we recognize a face, we need to detect a face and for that purpose, Google ML Kit's face detector is used. Once the face is detected, it is then fed to the face recognition model.

4.5 Emergency Shake

The app saves emergency contacts of the user and in case of emergency, if the user shakes their mobile phone, it will send the time and location of the user to the emergency contacts.

The message feature is added through shake detection because the visually impaired person will have to just shake their phone to send a text instead of manually typing a message.

4.6 Debugging A-EYE

We will be able to detect and remove existing and potential flaws (commonly known as 'bugs') in an App that can cause it to act unexpectedly or crash throughout the debugging process.

The Android Studio provides a debugger that almost debugs the App. For this, we need to select the device to debug on the App. Apart from this, we must set breakpoints in programming languages like C/C++, and Java Code. And later on, we have to examine the overall variables and evaluate the expressions at runtime.

4.6.1 Debugging pre-built APK

The Android Studio 3.0 and higher allow us to profile and debug A-EYE APK without having to develop them from an Android Studio project. However, we need to make sure we are using an APK with the debugging option enabled. For this, we need to navigate from the Android Studio Welcome screen, select Profile, or debug APK to begin debugging the APK. Alternatively, from the menu bar, select File > Profile or Debug APK if we already have a project open. Simply select the APK we want to import into Android Studio in the following dialogue window and click OK.

The unpacked APK files are subsequently displayed in Android Studio. This is not a fully decompiled collection of files, but it does include small files that make the .dex files more readable.

Chapter 5

Conclusions and Recommendations

This chapter provides a conclusion to the A-EYE Project and an overall summary. We developed a versatile system that helps visually impaired people with many tasks and make their lives better. In this project, we have explored every single implementation phase regarding the development of the A-EYE App and ML models for Object detection, Currency Recognition and Face Recognition. With this app, every visually impaired user will get a chance to become less dependent on other people around them. It means that A-EYE will try to make the experience for visually impaired people as easy as possible. While technology has helped humanity in almost every field, it also has a negative effect on eyesight, many people have any kind of visual problems nowadays. With A-EYE, these people get a chance to recognize things around them just like any other person.

Recommendations

Considering the scenario, nowadays it would be quite challenging to develop such a versatile app that could enable the entire visually impaired population to experience the world like everyone else but we can try to make it as better as possible. We can add more features such as color recognition and fruit classification etc. As of now, A-EYE can only recognize Pakistani Currency notes. It would be even better if we train the model such that it can recognize the currency notes of almost every other country. The best use of this idea will come to life when this idea is implemented in wearable glasses.

Results

The code is according to the model view controller (MVC) pattern. Which is considered one the most popular pattern in web and app development. S b which made them scalable and become business logical. In the model, every layer is separated. It has a separate presentation layer. And controller map all the layers. The app is very efficient from the performance point of view, all the listeners to text fields are disposed of when they are not needed. When we visit a screen, we initialize it at that time, when we switch to another screen the dispose function is called for the previous screen and it reduces the memory cache size. We are using the local storage of the phone as a database for fast and easy access without the need for any internet connection.

References

[1] "Seeing AI" [Online]. Available: https://www.microsoft.com/en-us/ai/seeing-ai.

[2] “Noteify: Indian Currency Recognition App” [Online]. Available: https://github.com/chandran-jr/Noteify

[3] “Why use Python for AI and ML”. [Online]. Available: https://steelkiwi.com/blog/python-for-ai-and-machine-learning

[4] “Python Community”. [Online]. Available: https://www.python.org/community

[5] N. Samrgandi, 1 1 2021. [Online]. Available: https://www.researchgate.net/publication/349087972_User_Interface_Design_Evaluation_of_Mobile_Applications..

[6] B. Ebrahim, "How to integrate machine learning models into your app?" 1 12 2019. [Online]. Available: https://medium.com/@BdourEbrahim/how-to-integrate-machine-learning-models-into-your-app-b77717e2702.

[7] "MIT library," [Online]. Available: https://libraries.mit.edu/scholarly/publishing/apis-for-scholarly-resources/.

[8] Ö. Çelik, "A Research on Machine Learning Methods and Its Applications," 9 2018. [Online]. Available: https://www.researchgate.net/publication/327547625_A_Research_on_Machine_Learning_Methods_and_Its_Applications.

[9] T. Rachad, "Intelligent Mobile Applications: A Systematic Mapping Study," 25 3 2020. [Online]. Available: https://www.hindawi.com/journals/misy/2020/6715363/.

[10] F. A. A. H. N. M. P. S. Lokhande, "Efficient Way of Web Development Using Python and Flask.," 1 2015. [Online]. Available: https://www.researchgate.net/publication/280154217_Efficient_Way_Of_Web_Development_Using_Python_And_Flask.

[11] [Online]. Available: https://enterprise-architecture.org/docs/application_architecture/application_deployments/.

[12] "APPLICATION DEVELOPMENT USING FLUTTER," [Online]. Available: https://www.irjmets.com/uploadedfiles/paper/volume2/issue_8_august_2020/3180/1628083124.pdf.

[13] L. Jupudi, "Machine learning techniques using python for data analysis in performance evaluation," 1 1 2018. [Online]. Available: https://www.researchgate.net/publication/324928841_Machine_learning_techniques_using_python_for_data_analysis_in_performance_evaluation.

[14] https://www.python.org/doc/essays/blurb/