CS 443 - Mobile Applications
Prof. Bo Sheng

Android Project: Smart Parrot
A text to speech android app with capability to speak from image, pdf and normal given text.

Members: Yash Mahant
   Venkat Sai Sravya Uppalapati

Date: 05/19/2021

Project Statement:

With the recent trends in audiobooks and podcast services there is significant growth in converting articles, books, blogs etc., into some form audios. Text to speech applications are changing profitability for what reason would you invest your energy stopped before a screen when you can perform multiple tasks and not overlook anything in the advanced world? The app will allow users to convert text to speech, pdf to speech and even image to speech. ML Kit is a mobile SDK that brings Google's machine learning expertise to Android and iOS apps in a powerful yet easy-to-use package. Whether you're new or experienced in machine learning, you can implement the functionality you need in just a few lines of code. There's no need to have deep knowledge of neural networks or model optimization to get started. On the other hand, if you are an experienced ML developer, ML Kit provides convenient APIs that help you use your custom TensorFlow Lite models in your mobile apps.
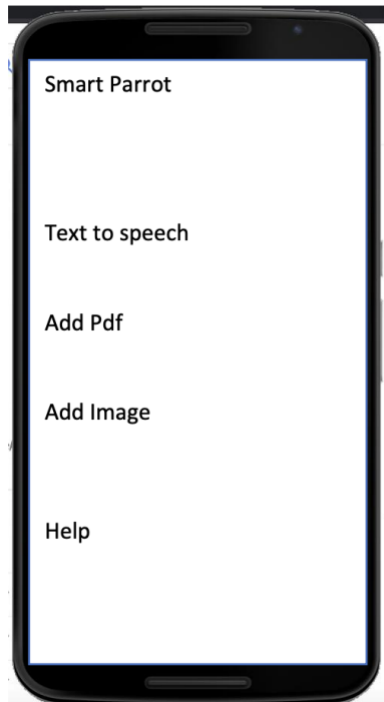

The whole idea of this app is to perform the three main functions,
1) Convert a given text to speech
2) Identify the texts from the pdf format files and convert it to speech
3) Identify the texts from the image file and covert it to speech
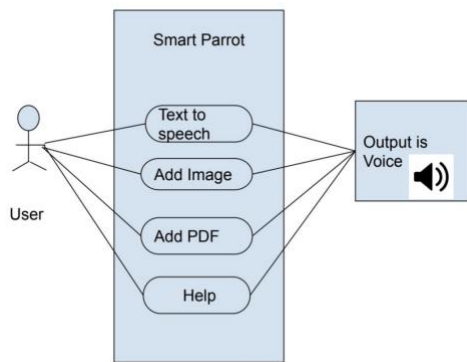
Although highly popular and successful apps like Audible exists in market but solution to all the speech related solutions are not available under one whole application. Like our app can turn texts from image into speech and even from given text. Also, this application can be of great use for people with vision disability not only that our app uses open source technology provided by Google that is Firebase ML kit.
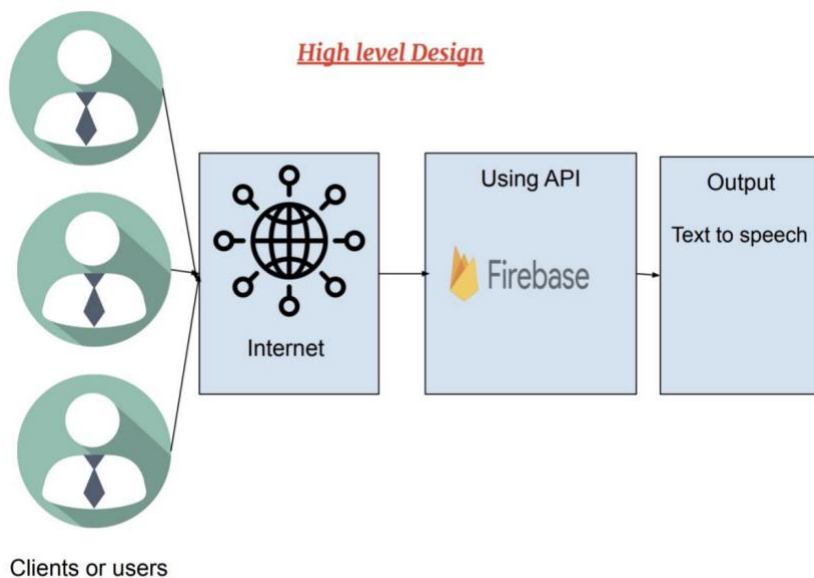
Application Design:

Our App was implemented by defining and using different classes like
- ImageActivity
  where the content is recognised and is converted to speech using firebase vision image.
- MainActivity
  In this class the major App activity is defined as a request for permission to allow camera and READ_EXTERNAL_STORAGE and WRITE_EXTERNAL_STORAGE.

- PdfActivity
  In this class the
  This fragment has a big {@ImageView} that shows PDF pages,
  {@link Button}s to move between pages. We use a
  {@link PdfRenderer} to render PDF pages as

- Text2Speech
  It converts the text Context to Listener.

We tested our app by trying all the possible ways to upload a document and tried Random pdf files.
We shared the app with our friends to test the app and it was working well.
Testing the functionality we have a sample unit test in our code which will execute on the development machine (host)
These methods helped us achieve functionality testing.

Performance of the Application
voices sound real clear, they have character, making them suitable for any application that requires speech output.

Usability of the App
Text to speech allows content owners to respond to the different needs and desires of each user in terms of how they interact with the content.
- Enhanced customer experience
- More autonomy for the digital content owner
- Increased web presence

Working ability of the App
We have never faced any issues on the performance and the working model of the application.
Text detection problems arise due to pixel quality of image.

We tried to solve the issue by using firebase ML kit.

References:

Firebase
https://firebase.google.com/docs
https://firebase.google.com/docs/ml-kit/recognize-text
https://firebase.google.com/docs/ml-kit/android/recognize-text

For zoom in zoom out of Pdf files
We have a reference from https://github.com/geraldPhiri/Android-Open-Lens-OCR-Text-Scanner.

Experiences and Thoughts

Initially we wanted to develop an application with features
- Converts text to speech
- Image to Speech
- Pdf to speech

But on further development along with above features we wanted to add features like upload a pdf.
For Image to capture from camera apart from choosing from gallery.
So these were the functions that are nice to have in our application.

We also have some Ideas to implement but due to the time limit we didn't go ahead.

Some of them are to apply ML techniques in your apps by bringing Google's ML technologies, such as
- Google Cloud Vision API
- TensorFlow Lite
- Android Neural Networks API together in a single SDK.

Whether you need the power of cloud-based processing, the real-time capabilities of mobile-optimized on-device models, or the flexibility of custom TensorFlow Lite models.

- Our speech recognition has problems identifying special characters and sometimes misreads Short Forms. Optical Character Recognition using firebase ML kit would solve the problem.
- Smart Parrot to accurately recognize text, input images must contain text that is represented by sufficient pixel data. Ideally, for Latin text, each character should be at least 16x16 pixels.

- So, for example, a 640x480 image might work well to scan a business card that occupies the full width of the image. To scan a document printed on letter-sized paper, a 720x1280 pixel image might be required.
- Poor image focus can hurt text recognition accuracy. If you aren't getting acceptable results, try asking the user to recapture the image.
- If you are recognizing text in a real-time application, you might also want to consider the overall dimensions of the input images. Smaller images can be processed faster, so to reduce latency, capture images at lower resolutions (keeping in mind the above accuracy requirements) and ensure that the text occupies as much of the image as possible
- The UI of Smart Parrot can also be designed better if we had some time on the deadline.

These are some of the possible enhancements that can be done on Smart Parrot.