Image to speech converter

Group Members:

Vaibhav Gaikwad
Uddesh Karda
Aniket Nighot

Mentor:

Mrs. Asma Parveen I.
Siddavatam

Introduction

Real world contains too many significant message and useful information cannot be ignored or left unread. Sometimes a signboard or any other notice could carry an important message or even danger notice that could be missed by visually impaired people. This application is mainly beneficial for visually impaired people to access printed text which may carry significant messages.

If the message is unreachable to mankind either due to biological barriers or linguistic barriers it might cause important information to be missed out which could lead to harm. Therefore this application will also be useful to the travellers and tourists, students, illiterate people to overcome the language barrier.

Objective

The aim of this project is to convert text in an image, taken by the user's smartphone camera into speech with increased computation speed and also keeping a high accuracy rate. The technology that allows us to convert text in images captured by an input device into an editable, searchable and reusable data is an optical character recognizer (OCR). Further this text generated can be converted into speech using the inbuilt android libraries of text to speech conversion.

Our approach to the solution

- Our solution includes the implementation of an android application which will be linked to a server.
- The user would click pictures or select pictures from gallery and send them to the server for further processing.
- At the server as soon as the image data is received, preprocessing would start.
- After the preprocessing phase, the processed image would be sent to the OCR engine for extraction of text.
- The user would then receive the extracted text, which the user could use in other features of the application.

Software Requirements

1. Android Studio
2. XAMPP
3. Python OpenCV libraries
4. Tesseract OCR engine
5. Yandex translator API
6. Android inbuilt TTS engine

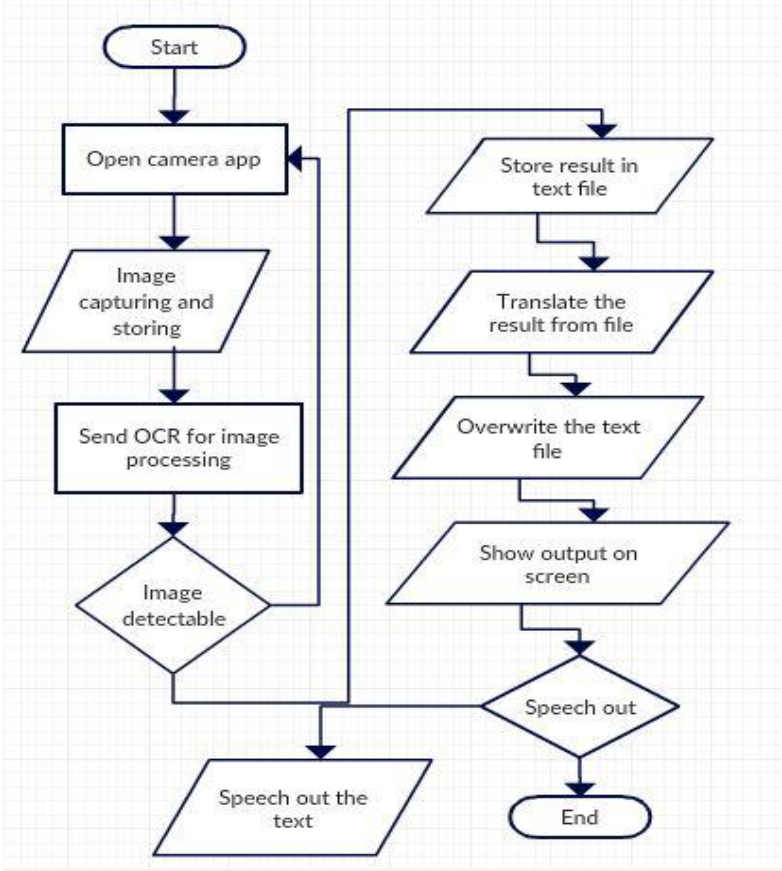
Languages used

- Python (Server)
- Java (Android)
- PHP (server and data transfer)
- JSON (server and data transfer)
- XML (server and data transfer)

Hardware Requirements

1. Android Mobile Device with auto-focus enabled camera with 3G data connectivity
2. Server

Flow diagram



Primary technology used

1. Image Preprocessing

- For the purpose of image preprocessing on the server OpenCV, a special python image processing library would be used.
- The techniques and methods that would be used for image preprocessing are:-
  1. Image scaling
  2. Image segmentation
  3. Image filtering and noise reduction
  4. Binarization or thresholding

2. Tesseract OCR Engine

- **Tesseract** is an optical character recognition engine for various operating systems.
- It is free software, released under the Apache License, Version 2.0, and development has been sponsored by Google since 2006.
- Tesseract is an OCR engine, not a complete OCR program.

- The following is a brief overview of how Tesseract works:

1. Outlines are analysed and stored
2. Outlines are gathered together as blobs
3. Blobs are organized into text lines
4. Text lines are broken into words
5. First pass of recognition process attempts to recognize each word in turn.

6. Satisfactory words passed to adaptive trainer
7. Lessons learned by adaptive trainer employed in a second pass, which attempts to recognize the words that were not recognized satisfactorily in the first pass
8. Fuzzy spaces resolved and text checked for small caps
9. Digital texts are outputted

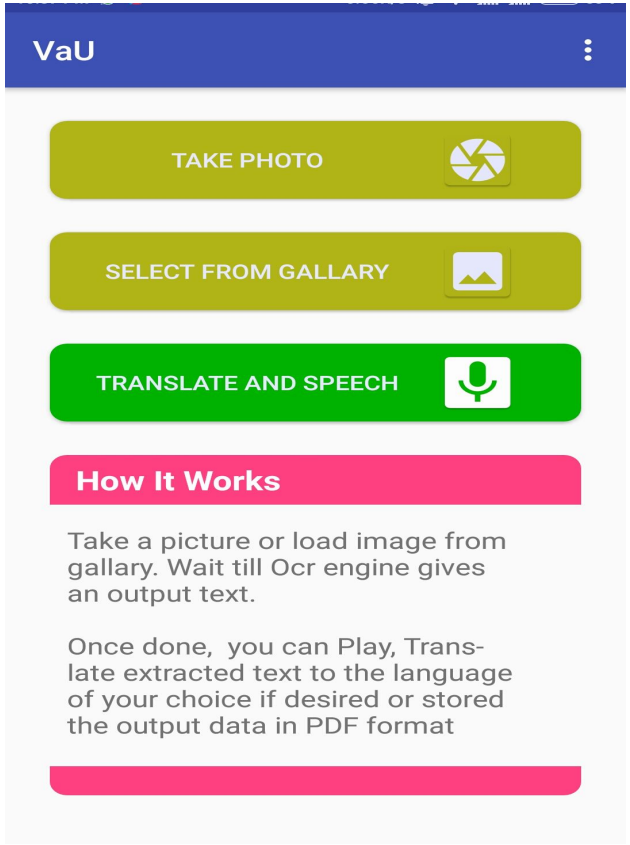
3. Translator

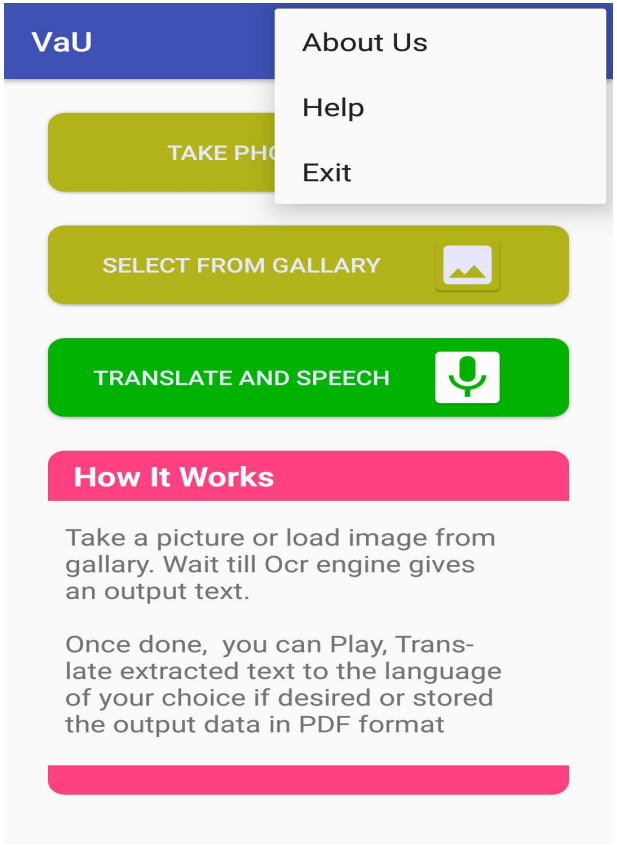
- There are various translators available in the market such as Google Translator, Bing Translator,
  etc.
- Google Translator and Bing Translators are paid API’s and provide access to only one user
  per access key.
- In our system Yandex translator will be used.
- Yandex is a translator that provides free API and multiple client access using one key.
- It also provides synchronized translation for 95 languages, predictive typing, dictionary with transcription, pronunciation and usage examples, and many other features.

4. Text to speech

- For text to speech, phone built-in feature would perform the speech out service.
- Android libraries such as android.text and android.speech will be used mainly for this purpose.
- Other available options were espeak, live-text- view and AndroidMary-TTS but the best option was the inbuilt text to speech libraries provided by google itself.

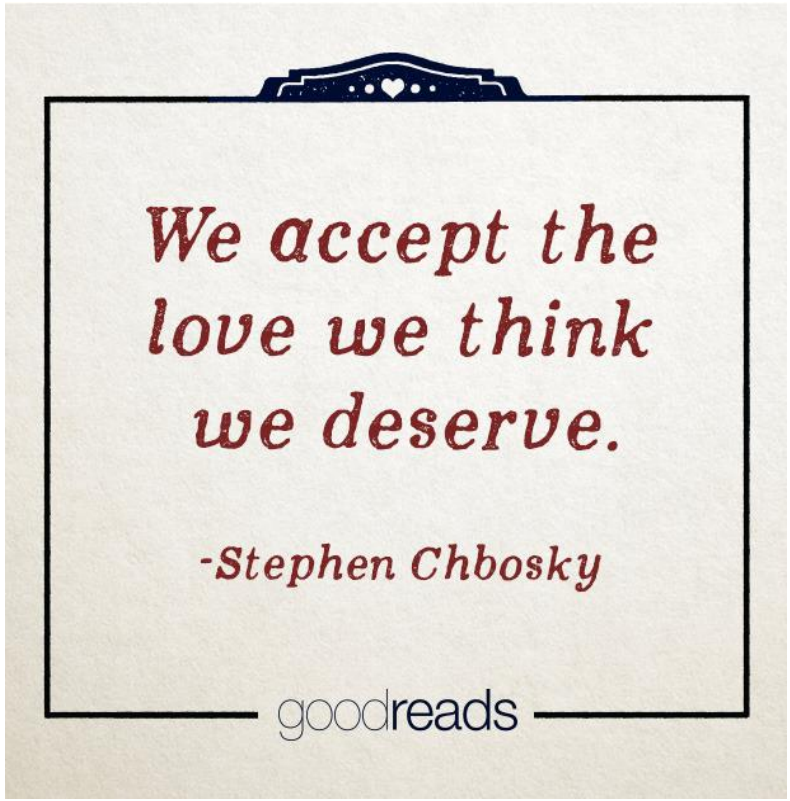
Application Interface



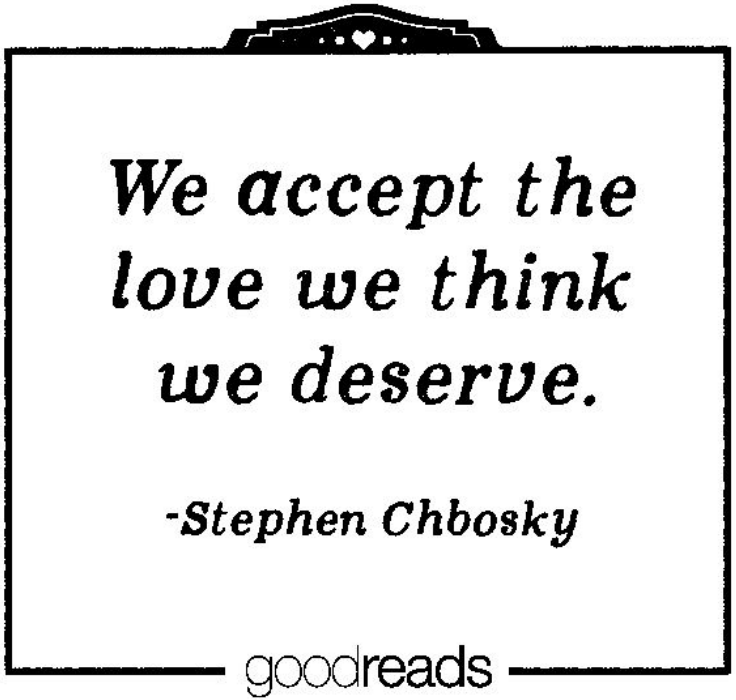


Output screenshots

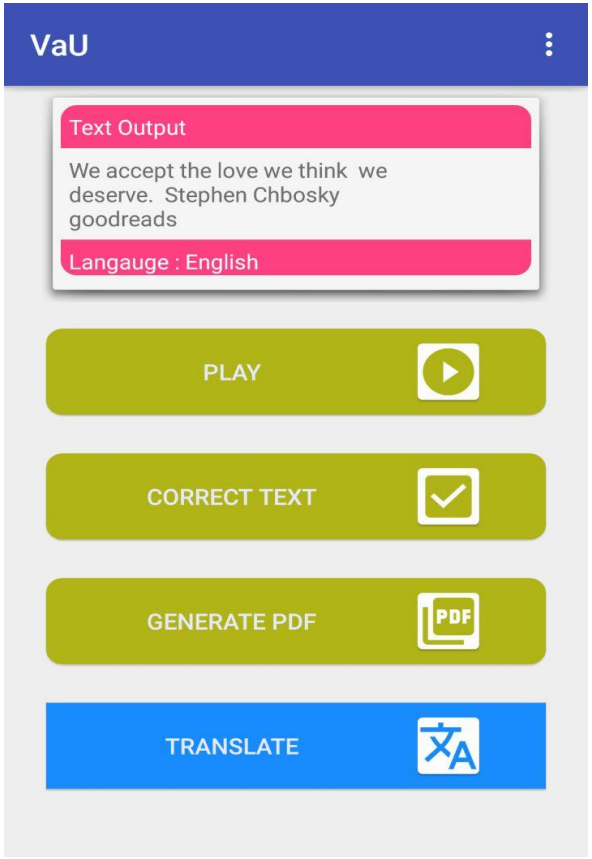
Input Image



1. Preprocessed Image-



2. Extracted text(OCR output)



3. Translation

Conclusion

- The implemented android application is able to recognize and extract the text in images with a good accuracy
- Moreover with the use of the android accessibility features such as talkback, the application proves to be useful in helping the blind or visually impaired people in listening to text that they were previously not able to read.
- Moreover the use of translator and text to speech feature makes it really easy for the tourists to not only understand but also communicate in different languages with native people.

Future Scope

- Many other languages can be added to the system.
- The application can also be extended to recognize mathematical and scientific documents.
- Also by adding templates for certain documents recognition can be made faster and easier.

References

1. Canedo-Rodriguez, S. Kim, J. Kim and Y. Blanco-Fernandez, 'English to Spanish translation
   of sign board images from mobile phone camera;, IEEE Southeastcon 2009, 2009.
2. OCR for Mobile Phones Kathryn Hymes and John Lewin
3. Teddy Mantoro, Abdul Muis Sobri, Wendi Usino, "Optical Character Recognition (OCR)
   Performance in Server-based Mobile Environment ;, 2013 International Conference on Advanced Computer Science Applications and Technologies
4. 2016 1st International Conference on New Research Achievements in Electrical and Computer Engineering Proposal for "Automatic License and Number Plate Recognition System for Vehicle Identification; by Hamed Saghaei
   Documentation:
5. OpenCV Documentation- https://docs.opencv.org/2.4/doc/tutorials/tutorials.html

Thank you