

“Image to speech Converter”



Group Members-

Aniket Nighot
Uddesh Karda
Vaibhav Gaikwad

Project Mentor-

Mrs. Asma Parveen

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. Also user can upload a pdf file which would be converted into speech. 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.

Existing System

- Text Fairy
- Accuracy Issue
- Missing characters Issue
- User Interface Issue

Our Solution

- We are making a android app which addresses above mentioned issues so that application will be usable to all.
- This app will directly convert any text image to speech.
- We are also including language translation feature which will translate the written text to selected language

Primary Study on Technology

1. Optical Character Recognition(OCR)

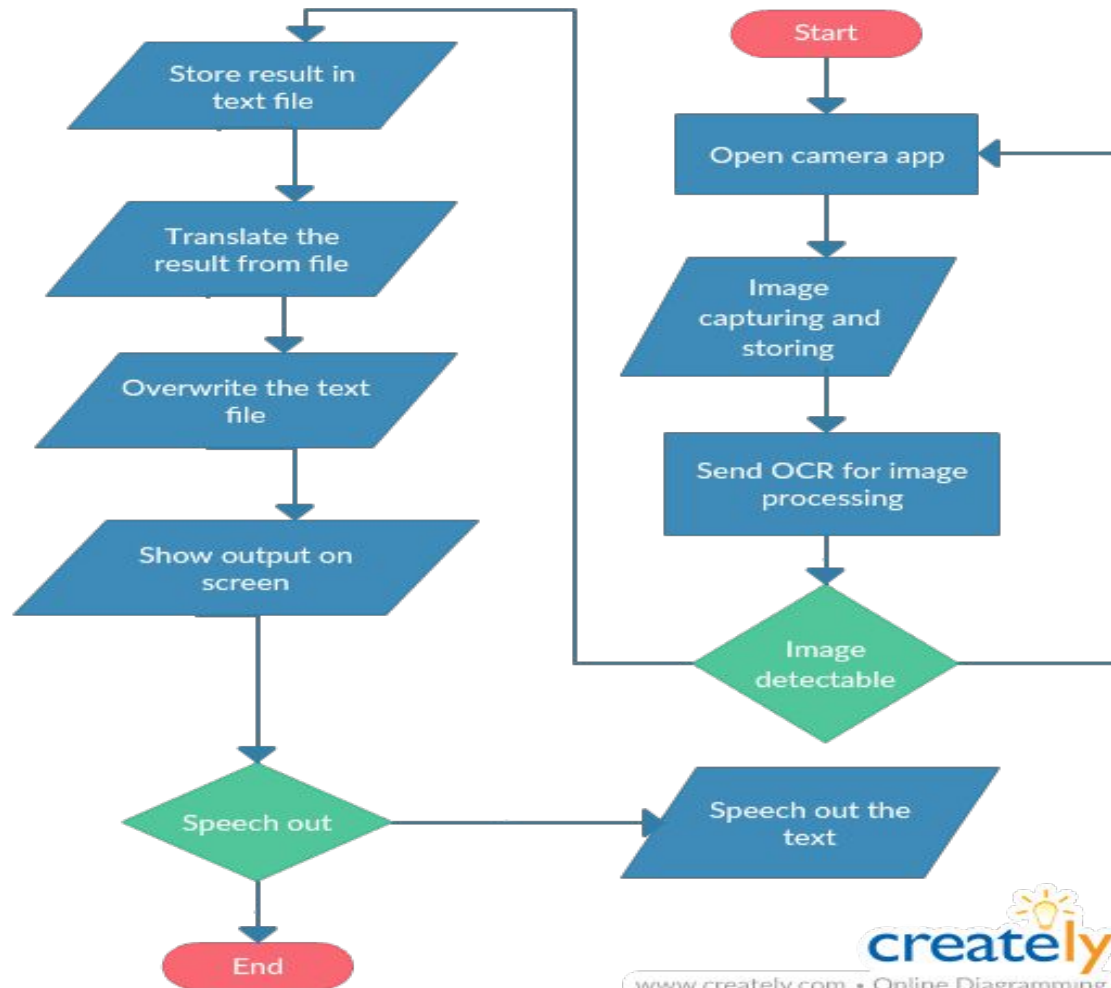
- OCR is a technology that allows users to convert text or documents in images captured by an input device into an editable, searchable and reusable data type for further image-processing.
- This technology enables a machine to recognize the characters automatically through an optical mechanism just like a human being use eyes to see an object in the world.

2. Translator

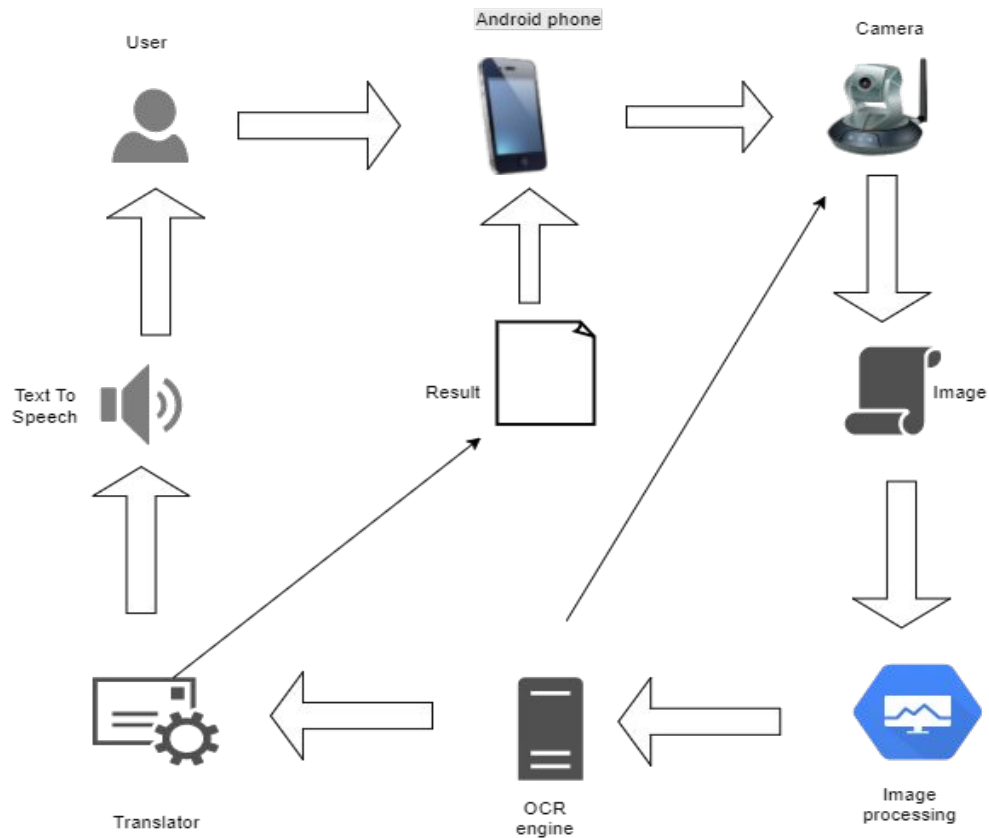
- For text translation, online translation service is chosen. There are many translators available for example NiuTrans, Systran, Google Translate, OpenLogos, Bing Translator, GramTrans, Babylon and etc.
- Out of many, three translators which are more suitable for mobile development compared which are Google Translate, Babylon, Yandex and Bing Translator.

3. 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



Architecture



Algorithms and Steps

A text detection / recognition / translation algorithm consists of following steps:

- 1) Morphological edge detection
- 2) Text feature filtering
- 3) Text region binarization
- 4) Optical character recognition
- 6) Text translation
- 7) Display of the translation

Hardware/software requirements

1. Tesseract OCR Engine.
2. Android Libraries for text to speech feature.
3. OpenCV library in Python for image preprocessing.

Conclusion

- Final deliverable of this project will be android application that will capture text, translate it and voice it out.
- In future, it can be upgraded with better OCR engine, translator services or even by multi supported text to speech engine. By doing so, it can significantly improve the performance of the system to ensure a better quality application.

References

Paper Title	Authors	Technologies
Detecting Text Based Image With Optical Character Recognition for English Translation and Speech using Android	<ul style="list-style-type: none">• Sathiapriya Ramiah• Tan Yu Liong• Manoj Jayabalan	Tesseract OCR engine, Bing translator and phones' built-in speech out technology.
Optical Character Recognition (OCR) Performance in Server-based Mobile Environment	<ul style="list-style-type: none">• Teddy Mantoro,• Abdul Muis Sobri• Wendi Usino	Server side processing OCRKit released for iPhone and iPad. Using a Text-To-Speech- IVONA

Medical Document Reader on Android Smartphone	<ul style="list-style-type: none"> ● Arrart Kongtalin ● Sutthipong Minsakorn ● Lalita Yodchaloemkul ● Sirasit Boontarak ● Sukanya Phongsuphap 	Tesseract OCR Engine
Proposal for Automatic License and Number Plate Recognition System for Vehicle Identification	<ul style="list-style-type: none"> ● Hamed Saghaei 	-
Optical Character Recognition Technique Algorithms	<ul style="list-style-type: none"> ● N. VENKATA RAO ● DR. A.S.C.S.SASTRY ● A.S.N.CHAKRAVARTHY ● KALYANCHAKRAVARTHI P 	Local techniques Neural Network

Implementation of Text to Speech Conversion	<ul style="list-style-type: none"> ● Chaw Su Thu Thu ● Theingi Zin 	WIN32 SAPI
English to Spanish Translation of Signboard Images from Mobile Phone Camera (optimization purposes)	<ul style="list-style-type: none"> ● Adrian Canedo-Rodríguez ● Jung H. Kim ● Soohyung Kim ● Yolanda Blanco-Fernández 	
Mobile Camera Based Text Detection and Translation	<ul style="list-style-type: none"> ● Derek Ma ● Qiuhau Lin ● Tong Zhang 	OpenCV with Tesseract OCR libraries