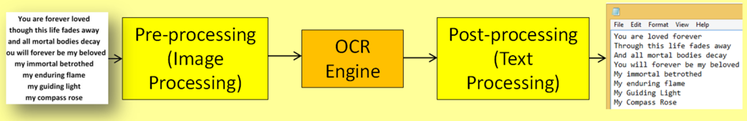
OCR-based Personal Assistant

Chaitanya Tejaswi, Shahnawaz Yousafzai
Department of Electronics & Communication, BVM Engineering College, VV Nagar
(Guide: Dr. Bhargav Goradiya, Head of Department)

OCR-PIA: Objective

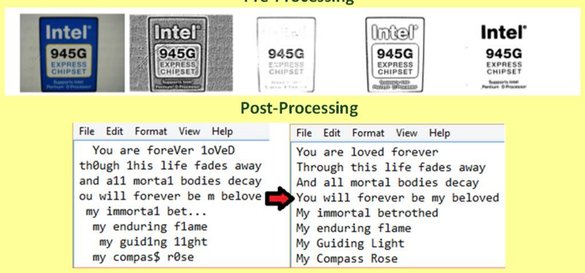
Modern-day applications integrate diverse functionality and package them in a standalone Android (*.apk*) file. One such application is an Optical Character Recognition (OCR) module, which interprets text from an input image, and displays it in raw text format. The nature of android makes it difficult for non-Java programmers to work with existing *apks*. The work to be carried out will thus involve programming an OCR system in a scripting language (such as *Python*), and implementing it on an Android base. Native functionality will be added as supported by the *Kivy framework*, which enables building standalone *apks* using Python scripts.

OCR System: Process Diagram



OCR: Problem Identification

- Generic OCR Engines assumed the image to be scanned (by a ***flatbed scanner***). So, the work was focused on developing robust ***OCR Engines***.
- Today, most images are captured using ***smartphones***. Implementing an OCR system for such images needs us to reconsider the defects generated in the process. So, the focus should be on effective ***Pre-Processing*** of images, while making use of existing OCR techniques.
- We must also note that OCR Systems are prone to make mistakes, and generate text that might not make much sense. So, additional ***Post-Processing*** of text will improve efficiency of the process.

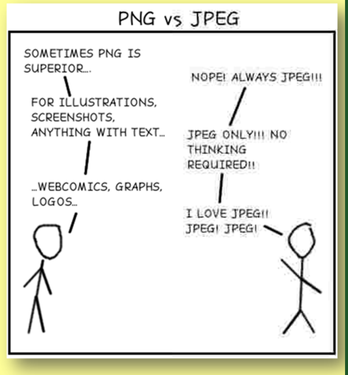


To JPEG or PNG, that is the question

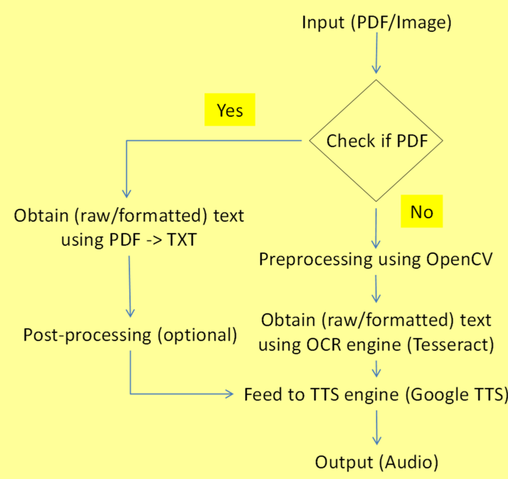
- JPEG uses lossy compression,PNG uses lossless compression.

WHAT DOES THIS MEAN?

- JPEG was designed for compressing full-colour/grayscale images of natural, real world scenes. It works well on photographs or naturalistic artwork, but not so good with lettering, cartoons, or line-drawings.
- Since human eyes perceive small colour-changes less accurately than small brightness-changes, JPEGs are great for us to look at. Plus, it saves space!
- But for OCR, we need high resolution image to start with, since we will lose pixels when processing it.
- PNG, which was meant for the Web, is better than GIF, and less complex than TIFF.
- But most smartphones save camera-generated images as JPEGs.

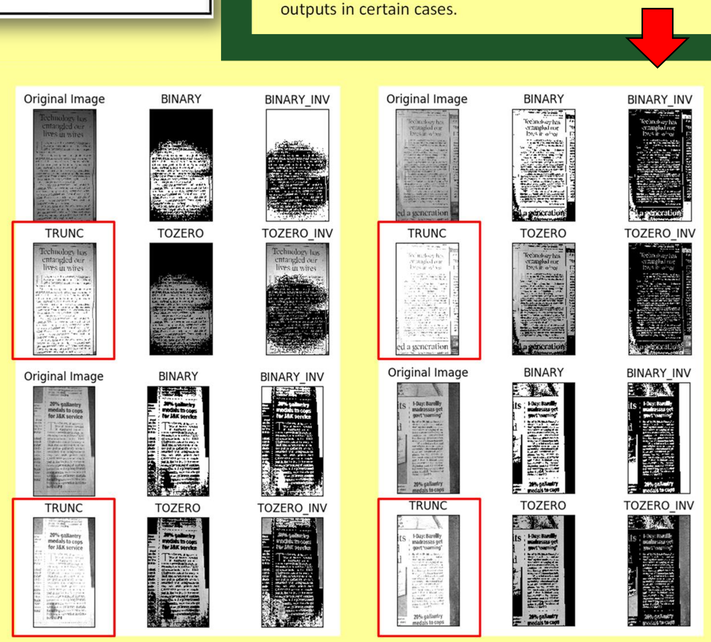


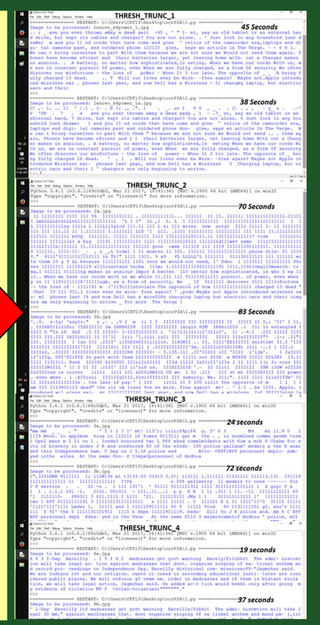
OCR-PIA: Flowchart



Some Results

We performed binarisation using ***Global-Thresholding*** methods in ***OpenCV***. But the time for conversion using ***Tesseract-OCR*** was found to be non-trivial. The results for ***Adaptive-Thresholding*** methods didn't yield any significant improvement. However, ***Otsu's Algorithm*** did provide quicker outputs in certain cases.





Steps Associated with Pre-Processing

Image processing
- Rescaling
- Binarisation (Thresholding)
- Rotation (De-Skewing)
- Border Removal
- Noise Removal

Page Segmentation

Dictionaries, Word lists & Patterns