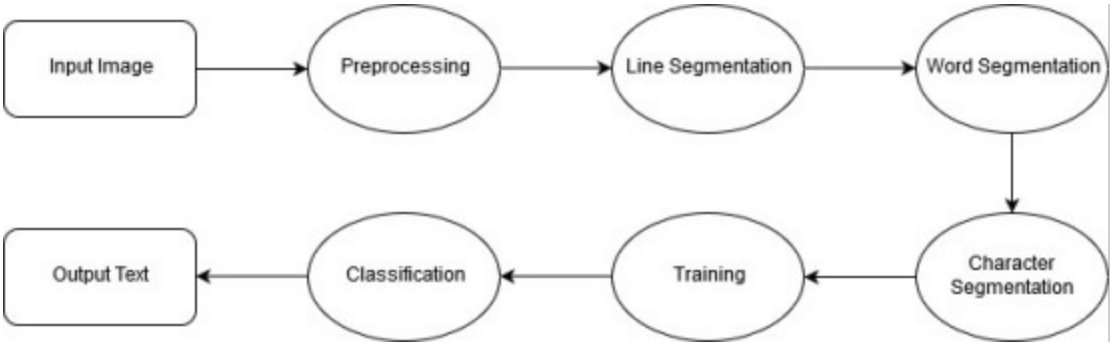
Project Pipeline



Preprocessing Module

We first convert the image from RGB level to Gray-scale level.
We use OTSU's Binarization. We use horizontal and vertical projection.
We tried noise removal and skew correction, but we didn't notice any increase in accuracy to justify the reduced performance.

Feature Extraction/Selection Module

We use dilation and contours to implement Line & Word Segmentation.
We use Baseline detection, cut point identification, and separation region filteration in character segmentation.

Performance Analysis Module

We compare the distance between output files and expected output with 'edit.py'.
We also print the time taken while generating the output.

Workload

Omar Mohammad

Character segmentation & Training

Walid Mohammad

Word & Character segmentation

Yahia Ali

Preprocessing, Line segementation