Information Extraction from Document Images July 2020

This project was aimed at streamlining the customer onboarding process via automation. As a part of the onboarding process the client received document images of government issued ID cards which were used to manually validate the information submitted by the customer during onboarding. The objective was to automate this process. In order to achieve this we leveraged the excellent Google Vision API for extracting textual information from the document images. In this project I worked with a data scientist from the client's team to develop the solution. My responsibilities were as follows:

- Preparing and labelling the custom image database
- Implementing an already existing training algorithm for Mask RCNN
- Cropping out ROI and correcting Skewness using OpenCV
- Preparing the cropped image and connecting to Google Vision API
- Parsing the response

The other data scientist in this project was responsible for deploying the solution for usage as well providing guidance and infrastructure throughout.

The major hurdle in this project was ensuring that all input images were properly cropped and skew-corrected. In order to accurately crop an image and perform proper skew correction we trained a Mask-RCNN on our custom labelled dataset to identify the region of interest. This ROI was then cropped from the input image and any skewness was corrected.

The resulting image was prepared for analysis using openCV's python implementation and then sent to the Google Vision API for information extraction. The response from the API was parsed to finally get a dictionary containing all the relevant information.