**OCR**

Any Typical machine learning OCR pipeline follows the following steps :

* Image preprocessing
* Text detection
* Text recognition

## Preprocessing

1. Remove the noise from the image
2. Remove the complex background from the image
3. Handle the different lightning condition in the image

01. Invert an Image: 9:47

03: Binarization: 13:33

04: Noise Reduction: 20:40

05: Dilation and Erosion: 28:33

06: Rotation and Deskewing: 35:07

07: Removing Borders: 42:18

08: Missing Borders: 49:09

09 : Transparency / Alpha Channel

### CRNN

This neural network architecture integrates feature extraction, sequence modeling, and transcription into a unified framework. This model does not need character segmentation. The convolution neural network extracts features from the input image(text detected region). The deep bidirectional recurrent neural network predicts label sequence with some relation between the characters. The transcription layer converts the per-frame made by RNN into a label sequence.

We can not expect the OCR model to be 100 % accurate. Still, we have achieved good results with the EAST model and Tesseract. Adding more filters for processing the image would help in improving the performance of the model.

### [Tesseract](https://nanonets.com/blog/ocr-with-tesseract/)

The capability of the Tesseract was mostly limited to structured text data. It would perform quite poorly in unstructured text with significant noise. Further development in tesseract has been sponsored by Google since 2006.

The latest stable version 4.1.0 is released on July 7, 2019. This version is significantly more accurate on the unstructured text as well.