Optical Character Recognition

OCR provides us with different ways to see an image, find and recognize the text in it. When we think about OCR, we inevitably think of lots of paperwork - bank cheques and legal documents, text present in number plate and street signs. In our case, we will try to predict the text of ID cards. Optical character recognition or OCR refers to a set of computer vision problems that require us to convert images of digital or hand-written text images to machine readable text in a form your computer can process, store and edit as a text file or as a part of a data entry and manipulation software.

To recognize the text of CNIC, use two pretrained model:

1. Easy OCR
2. Tesseract OCR

1) Easy OCR:

EasyOCR is implemented using Python and the PyTorch library. If you have a CUDA-capable GPU, the underlying PyTorch deep learning library can speed up your text detection and OCR speed tremendously.

As of this writing, EasyOCR can OCR text in **58 languages,** including English, German, Hindi, Russian, and more!

2) Tesseract OCR:

Tesseract is an open source text recognition (OCR) Engine, available under the Apache 2.0 license. It can be used directly, or (for programmers) using an API to extract printed text from images. It supports a wide variety of languages. Tesseract doesn't have a built-in GUI, but there are several available from the 3rdParty page. Tesseract is compatible with many programming languages and frameworks. It can be used with the existing layout analysis to recognize text within a large document, or it can be used in conjunction with an external text detector to recognize text from an image of a single text line. Before applying pretrained model of OCR do image processing.

Image Processing:

Images define the world, each image has its own story. It contains a lot of crucial information that can be useful in many ways. This information can be obtained with the help of the technique known as **Image Processing**. Image processing allows us to transform and manipulate thousands of images at a time and extract useful insights from them. It has a wide range of applications in almost every field.

For image processing use OpenCV, NumPy and PIL tool.

- ## OpenCV:

It stands for Open Source Computer Vision Library. This library consists of around 2000+ optimized algorithms that are useful for computer vision and machine learning.

- ## NumPy:

With this library you can also perform simple image techniques, such as flipping images, extracting features, and analyzing them.

Images can be represented by numpy multi-dimensional arrays and so their type is **ND Arrays**. A color image is a numpy array with 3 dimensions. By slicing the multi-dimensional array the RGB channels can be separated.

- ## PIL tool:

It stands for Python Image Library and **Pillow** is the friendly PIL fork by Alex Clark and Contributors. It's one of the powerful libraries. It supports a wide range of image formats like PPM, JPEG, TIFF, GIF, PNG, and BMP. It can help you perform several operations on images like rotating, resizing, cropping, grayscaling etc.

Comparison:

Easy OCR also recognize the line while tesseract is not. Easy OCR give the best result as compare to tesseract. In tesseract it's show special character and not recognize all the character correctly do the spelling mistake more while Easy OCR shown all the string with correct spelling and not any special character is present.

Result:

Easy OCR is **around 95% accurate** while Tesseract is **around 78.2% accurate.**