**ASSIGNMENT – 13**

**Name – Urmi Adak**

**Reg id – SIRSS1125**

**Text extraction from given images**

**Text recognition with machine learning**

We need to teach the computer to recognize what we know is text. The task is a bit simpler when we talk about high-quality, legible pictures, where the text is visible, and so are all the letters and digits. But what about pictures or scans of mediocre quality? This is where the challenge begins.

**OCR –** Optical Character Recognition

The most common text recognition technology is the OCR–Optical Character Recognition. OCR yields outstanding results only in specific use cases, but it is still considered challenging in general. Optical Character Recognition is a technology that enables you to convert different types of documents, such as scanned paper documents, PDF files, or images captured by a digital camera into editable and searchable data.

The most advanced OCR systems are focused on replicating natural human recognition. The OCR systems are based on three main rules–integrity, purposefulness, and adaptability. First, the observed object has always to be considered as one entity comprising many interrelated parts. In our case, the diploma is such an entity. Second, any interpretation of data must always serve some purpose. And finally, the OCR program has to be capable of self-learning.

* The usage of the OCR software

There are many types of OCR software that are used to serve lots of different intents. Right from “reading” the printed page from a book or a random image with text (for instance, graffiti or advertisement), up to reading street signs, car license plates, and even captchas. OCR software takes into consideration the following factors:

1. Text density. On a printed page, the text is dense. However, given an image of a street with a single street sign, the text is sparse. The OCR software has to recognize both.
2. Text structure. Text on a page is usually structured, mostly in strict rows, while text in the wild may be scattered everywhere, in different rotations, shapes, fonts, and sizes.
3. Font. While computer fonts are quite easy to recognize, handwriting font is much more inconsistent and, therefore, harder to read.
4. Artifacts. There are almost none of them on a perfectly scanned page, but what about outdoor pictures? In short, this is a completely different story, and you have to keep that in mind when using OCR.

* Real-world examples

Now, let’s consider some real-world, outdoor conditions examples: House numbers, road signs in a foreign country, and car license plates. House plates are extremely important, just to mention Google Street View and Google Maps. This is a massive source of tons of different house numbers. SVHN (Street View House Numbers) dataset incorporates over 600,000 digit images and is aimed at developing machine learning and object recognition algorithms.

Another widespread application of OCR is car license plate recognition. This also has a lot of possible applications, from police databases (data obtained from speed cameras) to private parking lots that open the barrier after a license plate and relevant payment is verified.

* The usage of the OCR software – Car Park

The car parks with text recognition and machine learning technology. One of the companies uses text recognition and extraction to manage cars driving in and driving out. When a vehicle approaches the barrier, the camera (similar to speed cameras) takes a picture of its license plate, sends it to the company’s central database, and the barrier automatically opens. When the text recognition part is done, the software extracts the car’s number plate and processes it into a plain, editable text, written in regular font.

* Machine learning text recognition in day-to-day situations– Car Park Payment

When a given car owner wants to leave the car park, they have to go to the ticket machine and choose their number plate from the list. Right after the payment, the barrier management software receives a signal that the given car can leave the parking lot. When the car approaches the barrier, its license plate is scanned again, and if the scanned number matches the already-paid numbers list–the barrier opens.

This is an example of how machine learning text recognition can be extremely helpful in day-to-day situations. The car owner doesn’t have to worry about a printed ticket and contrive where they should put it not to lose it. Everything happens within the software, and all the driver has to do is pay for their stopover.

* Machine learning text recognition and translation – Google Translate

When a person wants to identify or understand the texts written on the street they could easily use the help of the Google translate app.

This is another example of how machine learning text recognition can be extremely helpful in day-to-day situations. The person doesn’t have to worry about what is written on the signs when they are in a different country where the language might be an issue and not understandable to the person. Everything happens within the software, and all the person has to do is point the camera at the sign in the Google Translate App and wait to see the translation in the language they understand.

## ****Text extraction from images using machine learning****

With the text recognition part done, we can switch to text extraction. You see, at the end of the first stage, we still have an uneditable picture with text rather than the text itself. To solve this problem, the next step is based on extracting text from an image. Right after text recognition, the localization process is performed. All the related features about a particular image are gathered.

### ****REGION-BASED METHOD****

This method of text extraction uses a sliding window to detect text from any kind of image. This approach relies on several factors, such as color, edge, shape, contour, and geometry features.

### ****TEXTURE-BASED METHOD****

This method uses various kinds of texture and its properties to extract text from an image.

### ****HYBRID TECHNIQUE****

It’s the combination of the previous two techniques. First, the region-based approach is used to detect a text. Then, with the usage of the texture-based method, all the features are extracted from the text region.

### ****EDGE-BASED METHOD****

As its name indicates, this method is based on the detection of the edges of every letter and digit. This method is used to develop a high-level contrast between the text and the background.