**Handwriting Recognition using OCR - Optical Character Recognition**

**What is OCR?**

OCR, sometimes known as text recognition, is a term that is used occasionally. Data is extracted and reused from scanned documents, camera photos, and image-only PDFs by an OCR application. The original material can be accessed and edited by using OCR software, which isolates letters on the image, turns them into words, and then turns the words into sentences. Furthermore, it does away with the requirement for human data entry.

**Problem Statement:**

The optical character recognition (OCR) market is anticipated to be worth USD 13.38 billion by 2025, with a 13.7% year on year growth. The quick digitalization of business processes employing OCR to cut labor expenses and conserve valuable man hours is what fuels this rise. Although handwriting recognition (Handwriting OCR) or handwritten text recognition (HTR) is a crucial part of optical character recognition, it is still regarded as a difficult problem statement. Converting handwritten text into machine readable text is extremely difficult due to the wide variation in handwriting styles across individuals and the inferior quality of handwritten text compared to printed text. However, it's an issue in several sectors, including banking, insurance, and healthcare, needed to be resolved.

We will learn about the complexity of the task of handwritten text recognition in this project, as well as how deep learning techniques can be used to resolve it.

**Advantages of Handwriting Recognition using OCR:**

The fundamental advantage of optical character recognition (OCR) technology is that it makes text searches, editing, and storage simple, which simplifies data entering. OCR makes it possible for companies, people, and other entities to save files on their PCs, laptops, and other gadgets, guaranteeing ongoing access to all paperwork.

The following are some advantages of using this technology:

* Cost Cutting
* Speed up workflows
* Automate content processing and document routing
* Data centralization and security (no fires, break-ins or documents lost in the back vaults)
* Ensure staff members have access to the most recent and correct information to improve service

**Challenges in Handwriting Recognition** using **OCR:**

* Handwriting strokes have a great deal of variation and ambiguity from person to person.
* An individual's handwriting style also changes and fluctuates over time.
* Due to deterioration over time, the source document or image is of poor quality.
* While individuals do not have to write a line of text in a straight line on white paper, text in printed documents always sits in a straight line.
* Character separation and identification in cursive writing are difficult, contrary to printed text, where all the text sits up straight, text written by hand might vary in its rotation to the right.

**Deep Learning Techniques:**

Handwriting recognition was first approached using machine learning techniques like Hidden Markov Models (HMM), SVM, etc. After the initial text is pre-processed, feature extraction is done to extract important details about a certain character, such as loops, inflection points, aspect ratio, etc. To obtain the results, these created features are now sent to a classifier, such as an HMM. Due to their restricted learning ability and manual feature extraction phase, machine learning models perform very poorly. The feature extraction stage is not scalable because it differs for every single language. Deep learning has significantly increased the accuracy of handwriting recognition. In this project, I am going to cover about some of the leading studies on deep learning for handwriting recognition. Multi-dimensional Recurrent Neural Networks such as RNN/LSTM and Connectionist Temporal Classification (CTC) are used in this project.

RNN/LSTM can process sequential input to find temporal patterns and provide output in Handwriting recognition using OCR. They cannot, however, be directly applied to image data because they can only deal with 1D data. To overcome the above-mentioned problem, a multi-dimensional RNN/LSTM structure can be used.

The Connectionist Temporal Classification (CTC) algorithm is used to handle tasks like speech recognition, handwriting recognition, etc. where only the input data and the output transcription are available but no alignment information, such as how a specific region in the audio for speech or a specific region in the images for handwriting is aligned to a specific character, is provided. Since the amount of space each character requires in handwriting varies from person to person and occasionally, simple heuristics like giving each character the same area won't work.

The process of producing handwritten text that seems realistic can be used to supplement already available datasets. As is well known, gathering a large corpus of tagged handwriting images for several languages is a difficult undertaking because deep learning takes a lot of data to train. We can create training data for generative adversarial networks to solve this.

**Contributions:**

* In this project I am using ‘Keras’ which provides different preprocessing layers to deal with different modalities of data. My project involves preprocessing labels at the character level. For Example, if there are two labels, e.g., "cat" and "dog", then our character vocabulary should be {a, c, d, g, o, t} (without any special tokens). I used the ‘StringLookup’ layer for this purpose.
* In this Project I used the CTC loss as an endpoint layer.
* I used RNN Model with two max pools with pool size and strides 2. I have tested the model by increasing and decreasing the number of filters in the last layer from 64 to 128 before passing the output to the RNN part of the model.
* Utilized Evaluation metric ‘Edit Distance’ for evaluation of OCR Models. In this project, I have implemented it and used it as a callback to monitor our model.
* To get good results I have used 50 epochs while training the model.
* The prediction model I have utilized is fully compatible with TensorFlow Lite.

**Conclusion:**

Although there have been major technological advancements that aid in better handwritten text recognition, handwriting recognition is still far from a solved problem when compared to OCR and is hence not yet widely used in industry. Nevertheless, given the speed at which technology is developing and the arrival of models like transformers, we may anticipate that handwriting recognition along with OCR models will soon become the norm.

**Appendix:**

**Code URL:**

* https://github.com/Kruthitatavarthy11/ktatava\_64061.git

**References:**

* For a detailed understanding of the CTC loss, refer to [this post] (<https://distill.pub/2017/ctc/>)
* <https://nanonets.com/blog/handwritten-character-recognition/#summary>
* <https://www.ibm.com/cloud/blog/optical-character-recognition>