IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⋇ Impact Factor 7.39 ⋇ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11638

Review on Handwritten recognition Using IAM datasets

Dr. Maria Manuel Vianny[1], Harshitha K C[2], Keerthana L[3],Pavithra S, Varshitha YR[5]

Assistant Professor, Dept of CSE, Data Science, Jain University, Bangalore, India.[1]

Dept of CSE, Data Science, Jain University, Bangalore, India[2-5]

Abstract: Handwritten recognition is the ability of a computer to acquire and interpret handwritten entries from sources including paper files, images, touch displays, and different gadgets. Handwritten reputation is the maximum difficult undertaking because it's far a repeated painting that's written through human beings and inflicts terror. The handwritten reputation has been carried out in form of programs like Banking sectors, Health care industries, and lots of such organizations. Handwriting reputation is maximum generally utilized in modern-day cellular international is handwriting reputation as a right away enter to be transformed into texts. This is beneficial because it lets the person to speedy jot down numbers and names for contacts compared to inputting the equal records thru the onscreen keyboard. This paper consists of a survey on handwritten textual content reputation studies papers that have used IAM datasets.

Keywords: Handwritten recognition, IAM dataset.

INTRODUCTION –

Writing has a lengthy history. It commenced as easy pictographs drawn on a rock which slowly emerged into scripts. Handwriting is particular to each person and therefore there are more probabilities that now no longer all and sundry can interpret each handwriting. Every handwriting has special strokes and styles. In this contemporary international all of the files are digitally typed and saved however for the files that have already been handwritten this is tough to interpret. That's which HWR comes into the picture.

Handwritten reputation is the ability of a computer to acquire and interpret intelligible handwritten enter from information sources. Nowadays Handwritten Recognition is the most important tremendous and tough study area withinside the place of Image processing. The goal of the handwritten recognition machine is to enforce a person's pleasant pc assisted man or woman illustration to be able to permit a successful extraction of characters from handwritten files and to digitalize and translate the handwritten textual content into system readable textual content. Recognition of Handwritten English phrases was widely studied withinside the preceding years. Currently, numerous reputation methodologies are applied for the reputation of handwritten phrases. The application area of HWR is virtual file processing including mining records from information entry, cheques, programs for loans, credit score cards, tax, medical insurance paperwork, etc. During this survey, we gift a definition of contemporary studies paintings carried out for recognition of handwritten English phrases using IAM datasets. In a manuscript, there may be no restrictions on the writing techniques. Handwritten alphabets are complex to apprehend due to miscellaneous human handwriting approaches, the difference in duration and form of letters, and perspective. A form of reputation methodologies for handwritten English phrases is conferred right here along with their normal performance.

WORKING PRINCIPLE

Normally handwritten recognition is divided into six phases which are image acquisition, pre-processing, segmentation, feature extraction, classification, and post-processing.

A. Image Acquisition

Digitized/Digital Image is initially taken as input. The most common of these devices is the electronic tablet or digitizer. These devices use a digital pen. Input images for handwritten characters can also be taken by using other methods such as scanners, and photographs or by directly writing on the computer by using a stylus.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁒ Impact Factor 7.39 ⁒ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11638

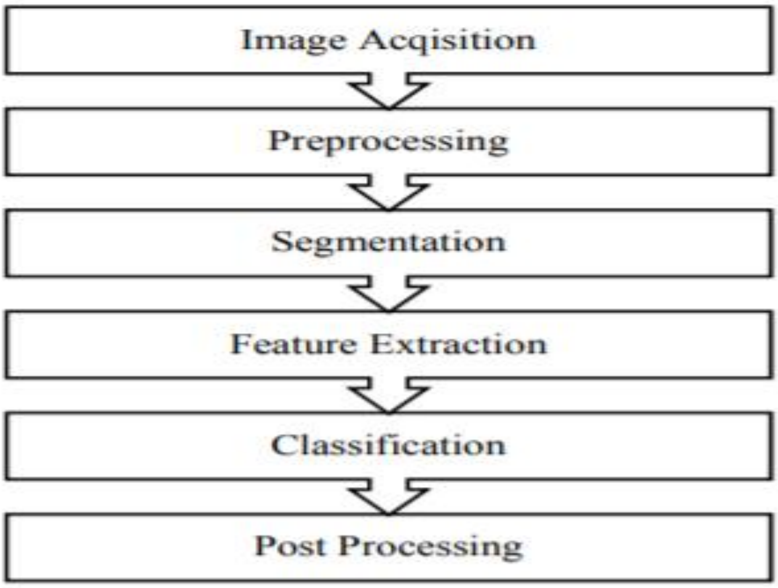

Figure 1: Block Diagram of Character Recognition

B. Preprocessing

Pre-processing is the fundamental segment of character recognition and it is crucial for an excellent recognition rate. The predominant goal of pre-processing steps is to normalize strokes and remove versions that might in any other case complicate recognition and decrease the popularity rate. These versions or distortions encompass the abnormal length of textual content, lacking factors at some stage in pen movement collections, jitter found in textual content, left or proper bend in handwriting, and choppy distances of factors from neighboring positions. Pre-processing consists of 5 common steps, namely, length normalization and centering, interpolating lacking factors, smoothing, slant correction, and resampling of factors.

C. Segmentation

Segmentation is carried out through the separation of the individual characters of a photograph. Generally, the file is processed hierarchically. In the first stage, strains are segmented with the use of a row histogram. From every row, phrases have extracted with the use of a column histogram and ultimately characters are extracted from phrases.

D. Feature Extraction

The predominant goal of the characteristic extraction segment is to extract the sample that's maximum pertinent for type. Feature extraction strategies like Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), Chain Code (CC), Scale Invariant Feature Extraction (SIFT), zoning, Gradient primarily based functions, Histogram is probably carried out to extract the functions of person characters. These functions are used to educate the machine.

E. Classification

When an input image is provided to the HCR machine, its functions are extracted and given as an entry to the educated classifier like a synthetic neural community or aid vector system. Classifiers examine the enter characteristic with the saved sample and discover the excellent matching magnificence for entering.

F. Post Processing

Post-processing refers back to the system of correcting misclassified consequences through making use of linguistic knowledge. Postprocessing is the processing of the output from form recognition.

Language records can boom the accuracy received through natural form recognition. For handwriting enter, a few form recognizers yield an unmarried string of characters, even as others yield some options for every man or woman, frequently with a degree of self-belief for every alternative.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⩫ Impact Factor 7.39 ⩫ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11638

LITERATURE REVIEW

This paper [1] **R. Reeve Ingle, Yasuhisa Fujii, Thomas Deselaers, Jonathan Baccash, Ashok C. Popat**
Discuss leverages the relationship between online and offline handwriting recognition by re-using an existing dataset of online handwriting samples to build a full HTR system. Their HTR system is trained on images that are rendered from the trajectory data of the online handwriting recognition system. To achieve sufficient accuracy on real images of handwriting, they have adopted an image degradation approach that is commonly used to generate realistic synthetic training images for OCR systems. After rendering clean handwriting images from
In trajectory data, the images are processed with an image degradation pipeline that applies realistic image transformations to the data. In addition to the collected online handwriting data. They also use a handwriting synthesis pipeline to enrich the variability in our training dataset to obtain better results. They are demonstrating the feasibility of this approach on Latin script, but we believe that this approach will work for other scripts as well. For the line recognizer they experimented with two different model architectures: They describe an LSTM-based architecture.
Similar to many of the state-of-the-art methods. However, a problem with recurrent architectures is that they do not train and run as easily on specialized hardware (GPUs, TPUs) as feed-forward networks. Thus They also propose a fully feed-forward network architecture that achieves comparable accuracy to our LSTM-based model. While the line recognition engine is often the primary focus of HTR research, it is just one component of a full-text recognition system. We describe the additional steps that were taken to integrate HTR support into an end-to-end text recognition system consisting of text detection, direction identification, script identification, and text line recognition models. Neural network models were implemented with TensorFlow.
A standard asynchronous stochastic gradient descent with multiple workers was used to train all models. The learning rate decayed exponentially through the training steps. The hyperparameters of the training were tuned on the development set if available; otherwise on a holdout set of the training data. The paper discusses in detail all advances in the area of handwritten character recognition. The most accurate solution provided in this area directly or indirectly depends upon the quality as well as the nature of the material to be read. Various techniques have been described in this paper for character recognition in the handwriting recognition system.

In Paper [2] **Annapurna Sharma, Dinesh Babu Jayagopi**
Recognition of cursive handwritten images has advanced well with recent recurrent architectures and attention mechanisms. Here, most of the works focus on improving transcription performance in terms of Character Error Rate (CER) and Word Error Rate (WER). Existing models are too slow to train and test networks. Furthermore, recent studies have recommended models be not only efficient in terms of task performance but also environmentally friendly in terms of model carbon footprint. Reviewing the recent state-of-the-art models, it recommends considering model training and retraining time while designing. This becomes challenging for handwriting recognition models with popular recurrent architectures. It is truly critical since line images usually have a very long width resulting in a longer sequence to decode. In this paper, they have considered the IAM dataset, the HW-AES dataset for English, and the RIMES dataset for French handwriting recognition. Here, they have presented a fully convolution-based deep network architecture for cursive handwriting recognition from line-level images. The proposed model is a combination of 2-D convolutions and 1-D dilated noncausal convolutions with Connectionist Temporal Classification (CTC) output layer. A data augmentation pipeline is further analyzed while model training. A comparison is done with state-of-the-art models with different architectures based on Recurrent Neural Networks (RNN) and their variants. This shows their model has fewer parameters and takes less training and testing time, making it suitable for low-resource and environment-friendly deployment.
In this paper, they have presented a CNN and dilated TCN-based deep network for offline handwriting recognition. The model structure is refined empirically. They further showed experimental results including data augmentation and image down sampling/scaling. The work showed that the proposed CNN-DTCN structure along with the CTC cost function enables faster handwriting recognition. Their model has fewer parameters and takes less training and testing time as compared to the state-of-the-art architectures which are mainly recurrent networks, viz MDLSTM, CNN-LSTM, etc. This in turn could lead to not only less carbon footprint but also better experience model uses for low-resource devices. An analysis of baseline models is also performed to show that the recognition performance is better than the other state-of-the-art methods without data augmentation. Though the model with data augmentation does not beat the recent state-of-the-art results extensive studies and experiments can be investigated with TCN and its variants to enable faster and more robust handwriting recognition.

In this paper [3] **Vasiliki Tassopoulou George Retsinas Petros Maragos**
Their objective is to enhance handwritten text recognition with N-gram target sequence decomposition (Unigram) and Multitask learning. The goal is to convert a sequence of features, extracted from an image to a sequence of text.
In this paper, they utilized a Multi-task Learning scheme, which is a subfield of machine learning in which multiple learning tasks are solved at the same time while exploiting commonalities and differences across tasks.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⧖ Impact Factor 7.39 ⧖ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11638

They trained the model using n-gram information, implicitly, in the training process, that performs decompositions of the selected target sequence with target units of different Granularity, while the final recognition is performed using only the unigram output. N-gram helps the machine understand a word in context. N-gram can be divided into Unigram, Bigram, and Trigram according to the number of text present.

The architecture used in this paper

- CTC - It converts a sequence of visual features into a set of target units.CTC decomposes the target sequence into unigram character-level target units and subsequently, during the decoding process, the transcribed text is going to be synthesized by them.
- Single-task Learning
- i) Optical Model ii) Column-wise Max-pooling iii) Sequential Model
- MULTI-TASK LEARNING Methodology
- BMT – create a different branch for each scale in a parallel manner is the Block Multi-task Architecture.
- HMT – a higher level of n-grams corresponds to more context-rich information than the previous level.

During the decoding process, they considered two alternatives: greedy and Beam Search decoding.

The greedy decoding algorithm has the disadvantage that, in the case of most character-level target units of fine coarse granularity, any word can be created, which is potentially negative as the formed word may be non-existent. Therefore they use the CTC Beam Search algorithm, a dynamic algorithm that allows taking into account external language information via word, i.e. Language Machine LM.

Their multi-task in Beam search decoder approaches outperforms the closest to HMT architecture in the greedy scheme, by an absolute 2.52 % in the WER. While maintaining the inference time the same as the single-task model.

In paper [4] **R . Parthiban , R . EzhilArasi , D . Saravanan**

Their objective is to make such an OCR that gives us an impressive recognition, and exactness of manually written text using RNN to improve accuracy. This study was performed on English handwritten text from the IAM image dataset and also from various sources. This study contains a dataset from IEE and also various datasets from different sources. According to this study, an input image is resized to 30*30 pixels and converted to a grayscale structure of size 128*32. Randomly split data into training and testing sets, the characters are anticipated precisely at the position. A technique used is a Recurrent neural network and the LSTM Neural network requires past words to foresee the sentence and RNN comprehends the issue using its hidden layer. First words get transformed into machine-readable vectors, then the RNN process the sequence of vectors one by one it passes the previous hidden state to the next step of the sequence. When (input) value is less than 1 input value vanishes also called vanishing gradient. Physically assigning w value for each would be time-consuming. Thus LSTM (Long short-term memory) is used to solve the "vanishing gradient" issue and it is done by replacing RNN hidden layers with LSTM cells.LSTM works in 4 types of gates such as Forget gate-->helps to reduce irrelevant information by passing hidden layer and current output to sigmoid function. Input gate--> previous hidden state and current input passes through sigmoid function and also passes to tanh function helps to regulate the network. Cell state-->acts as a transport highway, pointwise multiplication, and pointwise addition take place. Output gate-->The output is the hidden state. We can use the hidden state for predictions.

In conclusion this proposed framework the handwritten English document is scanned and recognized optically. By using the RNN algorithm output are not as printed text with 90% accuracy. To additionally improve the accomplished exhibition, there is a need of examining the issue further for discovering better arrangements by planning new engineering for English content.

REFERENCES

[1] Ingle, R. Reeve, et al, "A scalable handwritten text recognition system" International Conference on Document Analysis and Recognition (ICDAR), IEEE, 2019.

[2] Annapurna Sharma, Dinesh Babu Jayagopi, "Towards efficient unconstrained handwriting recognition using Dilated Temporal Convolution Network", Expert Systems with Applications, 164, Elsevier, 2021.

[3] Vasiliki Tassopoulou (Jan 11, 2021) "Enhancing Handwritten Text Recognition with N-gram Sequence Decomposition and Multitask Learning" 25th International Conference on Pattern Recognition, Underline Science Inc. 2021.

[4] R. Parthiban, R. Ezhilarasi and D. Saravanan "Optical Character Recognition for English Handwritten Text Using Recurrent Neural Network" International Conference on System, Computation, Automation, and Networking (ICSCAN), 2020.

[5] Rajib Ghosh, "A Recurrent Neural Network-based deep learning model for offline signature verification and recognition system", Expert Systems with Applications", 168 Elsevier, 15 April 2021.