SindhUniv. Res. Jour. (Sci. Ser.) Vol. 51 (03) 477-482 (2019)

http://doi.org/10.26692/sujo/2019.03.76

SINDH UNIVERSITYRESEARCH JOURNAL (SCIENCESERIES)

Automated Handwritten Character Recognition using Deep Learning

M. HUMAYUN, R. SIDDIQUI, S. S. ZIA++*, M. NASEEM*, I. MALA**, A. NAZIM

Department of Software Engineering, Bahria University Karachi Campus, Karachi-Pakistan

Received 15th October 2018 and Revised 12th July 2019

Abstract: Computer recognize the human written character, words documents comes under the definition of Handwriting Recognition. It's been and is still an area under research. It has gained prominence due to numerous applications. Like office automation, historic documents preservation, help for the blind etc. In this paper we provide in detail the existing study in Offline Handwritten Characters Recognition. The techniques used in this area of research. We provide a meticulous literature review on different techniques used in Offline Handwritten Character Recognition (HWCR) in different languages.

Keywords: BLSTM, Hybrid, CNN, HWCR, Offline, Online RNN

1. INTRODUCTION

Conversion from handwritten paper to digital format needs lot of manual work. Therefore, there is a dire need to automate this task so that a lot of time and resource may be saved. It is a difficult problem to solve (Toledo *et al.* 2017). This conversion is commonly known as writing recognition. It can further classified into handwritten words and handwritten character acknowledgement. Handwritten character recognition is recognizing the characters in natural handwriting either by online or offline handwritten character recognition (HWCR). Offline handwritten recognition as a difficult task, it doesn't have any information expect the matrix of pixels (Li, and Qiu 2016). This paper particularly focusses on the offline character recognition. The basic steps in offline handwriting recognition includes five major steps, Pre-processing, Dissection, Feature Mining, Cataloguing and Post-processing i.e. recognition.Pre-processing includes Noise elimination, Skew recognition/correction, Binarization (Castillo *et al.* 2015).

Offline works on the already written content. The offline may have individual characters or cursive or combination of cursive and separate characters.The words either be formed from the cursive handwritten characters or from the individual (block) characters (Sharma, *et al.* 2013).HWCR is a difficult chore due to numerous reasons like: Lettering Styles of every individual is different, not only this every writer has different writing styles under various mental and emotional state plays an important role. Similarity of structures of different characters causes confusion. The dataset's quality and quantity required for the training is vital.The researches in handwriting recognition has been done in different global and regional languages for decades, like in English, Arabic, Chinese (He *et al.* 2015), Bangla, Urdu, Georgian etc. but is still an open area of research (Soselia *et al.* 2018).

This paper is organized as follows. Section 2 presents the literature review of offline handwriting recognitionwhile Section 3 presents the methodology of offline handwriting recognition. Section 4, shows the comparative analysis of handwriting character recognition using Deep Learning classifiers. Section 6 presents the conclusions and future directions of the researchers in this dimension.

2. LITERATURE REVIEW

The handwriting recognition is a broad domain that falls under the category of OCR and is distributed into two sorts online and offline handwriting recognition. In provides a brief overview of the researches done using different features extraction method and classification method and its results (Castillo, *et al.* 2015).

(Patel, *et al.* 2012) focuses on offline handwriting recognition but the assumption is using blank papers for collecting data and uses multi resolution for feature withdrawal and Euclidean Distance Metric for identification/ prediction.Competitive Neural Trees is proposed as a learning technique for Myanmar language and concludes to, still have issues in handwriting recognition (Htike and Thein 2013). Neural networks and decision trees and other machine learning techniques are intended to identify the pattern and classification.

In the offline handwritten images the major factor which effects the accuracy of recognition is the noise. There are many reasons due to which the noise is introduced into image either be capturing or transmission of images, noises strongly effect the results

++ Correspondence author: Syed Saood Zia, saood_zia@hotmail.com,
*Department of Software Engineering, Sir Syed University of Engg. & Technology, Karachi- Pakistan.
**Department of Electrical Engineering, Usman Institute of Technology, Karachi- Pakistan.

of recognition. Multiple noise removing systems are proposed to diminish the noise in the imageries, noises like Gaussian Sound, Impulse noise, Shot noise, Uniform noise, there are multiple types of noises falling in the categories in linear and non-linear filters (Mythili and Kavitha 2011). KFCM algorithm is proposed for removing the noise forming cluster the pixels according to the text, background and noise according to the features (Sober and Levin 2017) focusses on the completion of the characters and strokes which may cause the noise and effect the recognition accuracy.

Over the past few years, with the institution of Deep learning, many researches have remained done in Handwriting recognition using deep learning especially different architectures of Artificial Neural Networks were intended for image processing, including Convolution Neural Network (CNN), Recurrent Neural Network (RNN) and different architectures of CNN including ResNet (He *et al.* 2016), VGGNet (Simonyan and Zisserman 2014), AlexNet (Krizhevsky, *et al.* 2012), GoogleNet (Szegedy *et al.* 2015) were used.

3. OFFLINE HANDWRITING RECOGNITION METHODOLOGY

For offline character recognition the methodology which will be followed is as under:

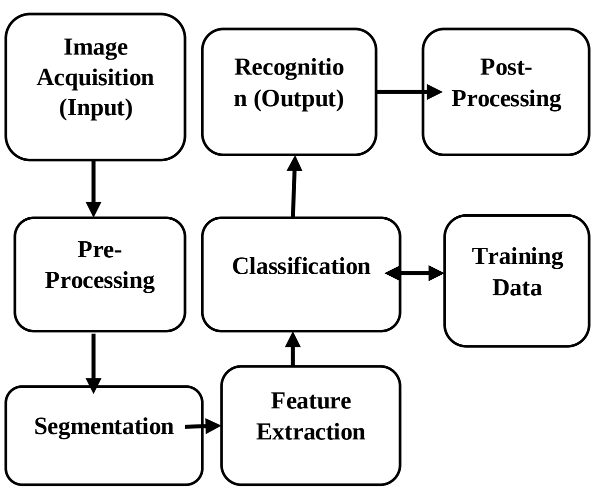

Fig.1. Flowchart of Offline Handwritten Character Recognition

A. Image Attainment

Image attainment is obtaining an image of the document through scanners and camera. The image will be exposed to the basic steps for HWCR.

B. Preprocessing

The aquatinted image is then subjected to the primary steps to make the input image prepared to the next steps. Some of the basic and necessary sub steps is Binary Conversion, Noise Removal, and Normalization.

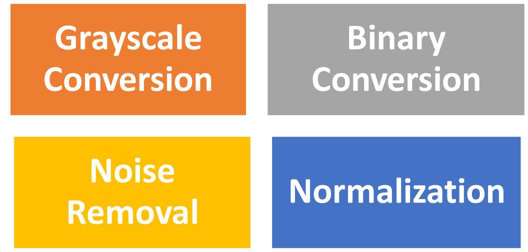

Fig.1. Preprocessing Steps

Grayscale and Binary Conversion

The image captured is RGB image, converted into grayscale. The gray scale is then altered into binary image. This process is normally baptized Binarization or Digitization of images.

Noise removal

Noise removal means that the extra pixels in the image that will act as a hurdle in image recognition will be removed. Appropriate algorithm is to be used. Usage of such algorithms to augment the eminence of the image and to make the imageries in accordance with the training data.

Normalization

The methods like Skew detection and correction, skeletionization etc. comes in the normalization of the input image.

C. Segmentation

The segmentation defines the state of the image in which the characters or words as the objects are defined in the image. The segmentation is distributed into two classifications, word dissection and character dissection. Segmentation divides the imageries, upon the category, into words, lines or characters. Correct segmentation increases the recognition rate. Common problems in segmentation are division of a character, segmentation of overlapping characters into one segment. Correct segmentation leads to greater chances of correct recognition. But for the correct segmentation it is necessary to recognize what you want to segment. It is even more difficult for the cursive handwriting recognition (Toledo *et al.* 2017).

D. Feature Extraction

The alternate representation of image is features. Extraction of expedient information from the input image is feature extraction. There are multiple features but choosing the right features that gives the true representation of image is important. There are numerous feature extraction techniques. The classifier is trained using the features and used for the groupings. Feature Extraction is the supreme noteworthy step in any pattern recognition problem, where the patterns are the images in this consequence (Naz *et al.* 2017). The features can be extracted manually but it is difficult and expensive step as it consumes time, human resources and energy (Granet *et al.* 2018), it can be automatically done. Representing the images in terms of features reduces the learning and computation capabilities (Naz, *et al.* 2017).

E. Classification

The input data and the classifier will be used for the recognition purpose. The input image will classified into class. Every class represent every symbol of that language. The classes will be having the samples in it with which the recognition will be performed.

F. Recognition

The recognition is about making the decision about the input image are following:

- What the image is?
- Which class does that character image belongs to?

4. COMPARATIVE ANALYSIS OF HWCR USING DEEP LEARNING CLASSIFIERS

The comparative analysis for handwritten character recognition is shown in **(Table 1).**

Table 1. Comparative Analysis with respect to Classifiers for HWCR

Classifier	Segmentation	Reference
CNN	Words	(Krishnan and Jawahar 2016)
		(Dilipkumar 2017)
	Characters	(Kaur and Rani 2017)
		(Zhang 2015)
RNN	Words	(Toledo et al. 2017)
		(Shkarupa, Mencis, and Sabatelli 2016)
		(Granet et al. 2018)
HYBRID (CNN + RNN)	Words	(Naz et al. 2017)
	Characters	(Yin et al. 2017)
	Characters	(Yang et al. 2017)

A. CNN

(Krishnan and Jawahar 2016) anticipated a scheme to measure similarity between two handwritten documents enthused by (Jaderberg *et al.* 2014). To identify the plagiarism, identify the keywords in old documents, analyzing person's notes, marking of the answer sheets, even in the health care for prescriptions etc. The proposed methods is named as Measure of Document Similarity (MODS). Similarity in the context of handwritten document is explained as the proportion of the text reused in the two contestant documents that are paralleled. (Krishnan and Jawahar 2016) focuses on the structural evaluation of the documents word scattering. Many approaches were available but the results were not scalable for the multiple writer problem. (Krishnan and Jawahar 2016) describes the proposed architecture by using CNN at the word level, as the feature descriptor. The experimental results concluded that, with the large amount of data the results can be improved. The architecture fails to work well with the documents with the graphics in the documents and has low performance for the mathematical equations and expressions.

(Dilipkumar 2017) handwriting data using limited real world data to improve word-level classification. Two datasets are used in this project. The first consists of maintenance records from Boeing aircrafts (private), and the second is the IAM Handwriting Database (publicly available). The architecture used for the word classification is CNN.

(Kaur and Rani 2017) have presented a scheme for offline handwritten Gurmukhi character recognition grounded on CNN classifier that classify a character based on the three features like Zoning, Horizontal Peak Extent and Diagonal. The Recognized system consists of the different Stages like digitization, preprocessing, feature extraction and classification.

(Zhang 2015) distinct that CNN has provided better results for handwriting recognition. The images in the experiment had to go under preprocessing, having the steps of grayscale conversion, resizing. (Zhang 2015) conducted the training of CNN with two possibilities, one with the dataset splitting into training and testing and other with the whole dataset and evaluate their results with different CNN configuration (number of layers in CNN) to examine the accuracy of the CNN, with different amount of filters in the CNN effect the performance of CNN using ReLUlayer with 0.5 epochs.

B. RNN

(Toledo *et al.* 2017) proposes a method using Deep learning (CNN) with attribute embedding to sequence learning. It is a two-step approach. First step is the attribute embedding and then to the BLSTM-CTC. In the proposed system, Pyramidal Histogram of Character (PHOC) is used for attribute embedding in the attribute space. In the space the words are categorized by the attributes. The PHOC is used to describe the words by the presence of a character. It is done by the PHOCNET to produce a sequence of PHOC. It is then forward into the BLSTM-CTC comprising two layers. Upon research (Toledo et al. 2017)clinched, the Datasets have the CET of 7.32 and 0.83% of Washington and Esposalles dataset. It also eliminate the need of lexicon for the attribute.

(Shkarupa, *et al.* 2016) proposed two methodologies of LSTM RNN are tested for word recognition. One methodology is using CTC (Connectionist Temporal Classification) and other is the sequence to sequence learning. The test is done on the unsegment words as input and decoding string on the output.

(Granet e*t al.* 2018) defines the architecture to decipher the problem of building the facts set for training by using the transductive transfer learning. The system is built upon the BLSTM-CTC. The aim is to transfer the knowledge from the modern data to recognize the historical data. Another aim is to transfer data from one language to another language. (Granet et al. 2018)is based on the no or few noted data. The training is done using the three datasets using the Italian comedy daily routine documents. The network constructed is based on the feature extraction and handwriting recognition.

C. HYBRID (CNN + RNN)

(Naz, *et al.* 2017) has proposed an architecture, conjoining convolution and recursive neural network (RNN). The CNN is used to extract low level features, apply the processes to the raw pixels and share the weights. RNN is used to extract the contextual features extractions and learning. The input in the RNN is the skeleton image as input and the output by the Connectionist Temporal Classification (CTC) layer.

(Yin, *et al.* 2017) methods providing auspicious result, used is either convolution neural network (CNN) andRNN with Long Short Term Memory (LSTM). (Yin. *et al.* 2017)focuses on the inherent characteristics of the text recognition and computer vision for acting like human vision and cognition mechanism for the understanding text. (Yin, *et al.* 2017) is using the CNN feature map with batch normalization, the proposed architecture detects and recognizes the text in one go by using the sliding window to extract features. Architecture encompasses to the three categories. Implicit segmentation, explicit segmentation and holistic methods. Implicit segmentation uses the window approach to slice out the images and then label the word sequence. In the explicit segmentation includes character and word segmentation. In the holistic method, the input images is considered as the input without segmentation for the classification. The transition layer, combines back the characters predicted into the predicted word or character sequence having the CTC and decoding. The architecture is inspired by the modern research in the human cognition mechanism, which includes saccades and fixation.

(Yang *et al.* 2017) present the iterative refinement module that can be pragmatic to the output feature maps of any prevailing CNN in direction to further increase classification accuracy. Iterative refinement module that receipts as input the feature maps learned by a CNN followed by Batch Normalization and ReLU (Nair and Hinton 2010). The iterative refinement module is executed with an attention-based recurrent neural network (RNN).

5. CONCLUSION

From the research papers it is concluded, for conducting the research the researchers have either used the online available datasets for their concerned languages or they can developed their own dataset. Quantity and quality of the dataset matters. If the data, is collected and personalized dataset is generated then the processes used in the collection of the data matters. More data more it contributes in the recognition rate, more variation it provide the classifier about a class regardless of the classifier used. Learning rate, number of filters also contributes in the recognition rate.

REFERENCES:

Castillo, E., Z. Grande, A. Calviño, W. Y. Szeto, and H. K. Lo. (2015). “A State-of-the-Art Review of the Sensor Location, Flow Observability, Estimation, and Prediction Problems in Traffic Networks.” *Journal of Sensors*.

Dilipkumar, D. (2017). “Generative Adversarial Image Refinement for Handwriting Recognition.” Student. Farahmand, Atena, Hossein Sarrafzadeh, and Jamshid Shanbehzadeh. 2017. “Noise Removal and Binarization of Scanned Document Images Using Clustering of Features.

Granet, A., E. Morin, H. Mouchère, S. Quiniou, and C. Viard-Gaudin. (2018). “Transfer Learning for Handwriting Recognition on Historical Documents.” In *ICPRAM*, 432–39.

He, K., X. Zhang, S. Ren, and J. Sun. (2016). “Deep Residual Learning for Image Recognition.” In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–78.

He, M., S. Zhang, H. Mao, and L. Jin. (2015). “Recognition Confidence Analysis of Handwritten Chinese Character with CNN.” In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, 61–65.

Htike, T. and Y. Thein (2013). “Handwritten Character Recognition Using Competitive Neural Trees.” *International Journal of Engineering and Technology* 5 (3): 352Pp.

Jaderberg, M., K. Simonyan, A. Vedaldi, and A. Zisserman. (2014). “Synthetic Data and Artificial Neural Networks for Natural Scene Text Recognition.” *ArXiv Preprint ArXiv:1406.* 2227Pp.

Kaur, H and S. Rani. (2017). “Handwritten Gurumukhi Character Recognition Using Convolution Neural Network.” *International Journal of Computational Intelligence Research* 13 (5): 933–43.

Krishnan, P. and C V Jawahar. (2016). "Matching Handwritten Document Images." In *European Conference on Computer Vision*, 766–82.

Krizhevsky, A., I. Sutskever, and G. E. Hinton. (2012). "Imagenet Classification with Deep Convolutional Neural Networks." In *Advances in Neural Information Processing Systems*, 1097–1105.

Li, Y., Z. Li, and Q. Qiu. (2016). "Assisting Fuzzy Offline Handwriting Recognition Using Recurrent Belief Propagation." In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–8.

Mythili, C, and V. Kavitha. (2011). "Efficient Technique for Color Image Noise Reduction." *The Research Bulletin of Jordan, ACM* 1 (11): 41–44.

Nair, V. and G. E. Hinton. (2010). "Rectified Linear Units Improve Restricted Boltzmann Machines." In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 807–14.

Naz, S., A. I Umar, R. Ahmad, I. Siddiqi, S. B. Ahmed, M. I. Razzak, and F.l Shafait. (2017). "Urdu Nastaliq Recognition Using Convolutional--Recursive Deep Learning." *Neurocomputing* 243: 80–87.

Patel, D. K., T. Som, S.l K. Yadav, and M. K. Singh. (2012). "Handwritten Character Recognition Using Multiresolution Technique and Euclidean Distance Metric." *Journal of Signal and Information Processing* 3 (02): 208Pp.

Sharma, N., T. Patnaik, and B. Kumar. (2013). "Recognition for Handwritten English Letters: A." *International Journal of Engineering and Innovative Technology (IJEIT)* 2 (7).

Shkarupa, Y., R. Mencis, and M. Sabatelli. (2016). "Offline Handwriting Recognition Using LSTM Recurrent Neural Networks." *28th BNAIC*, 88–95.

Simonyan, K and A. Zisserman. (2014). "Very Deep Convolutional Networks for Large-Scale Image Recognition." *ArXiv Preprint ArXiv:* 1409.1556.

Sober, B and D. Levin. (2017). "Computer Aided Restoration of Handwritten Character Strokes." *Computer-Aided Design* 89: 12–24.

Soselia, D., M. Tsintsadze, L. Shugliashvili, I. Koberidze, S. Amashukeli, and S. Jijavadze. (2018. "On Georgian Handwritten Character Recognition." *IFAC-PapersOnLine* 51 (30): 161–65.

Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. (2015). "Going Deeper with Convolutions." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1-9.

Toledo, J I., S. Dey, A. Fornés, and J. Lladós. (2017). "Handwriting Recognition by Attribute Embedding and Recurrent Neural Networks." In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 1:1038–43.

Yang, X., D. He, Z. Zhou, D. Kifer, and C. L. Giles. (2017). "Improving Offline Handwritten Chinese Character Recognition by Iterative Refinement." In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 1: 5–10.

Yin, F., Yi-Chao Wu, Xu-Yao Zhang, and Cheng-Lin Liu. (2017). "Scene Text Recognition with Sliding Convolutional Character Models." *ArXiv Preprint ArXiv:1709.*1727.

Zhang, Y. (2015). "Deep Convolutional Network for Handwritten Chinese Character Recognition." *Computer Science Department, Stanford University*.