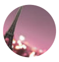
A State of Art Approaches on Handwriting Recognition Models

Conference Paper · March 2019

DOI: 10.1109/ICONSTEM.2019.8918925

2 authors:



Vinotheni Coumar
Pondicherry Engineering College

S. Lakshmana Pandian
Pondicherry Engineering College

A State of Art Approaches on Handwriting Recognition Models

Vinotheni C
Dept. of Computer science and engineering
Research Scholar
Pondicherry Engineering College
Puducherry, India
vinotheni95@pec.edu

Lakshmana Pandian *S*
Dept. of Computer science and engineering
Associate Professor
Pondicherry Engineering College
Puducherry, India
lpandian72@pec.edu

***Abstract*— Presently, handwriting recognition (HWR) become an open research issue to recognize the handwritten words out of a scanned document which is unconstrained. In spite of the extensive implication of computer in office, handwriting still stays as a significant mode of annotating and apprehending textual data. In recent days, there is an exponential growth in the number of researches carried out in this field because of its benefits. On the past decade, great efforts are paid to the online hand-drawn content interpretation from handwritten text. Keeping this in mind, in this paper, we made a review of existing HWR model in different aspects have been reviewed.**

***Keywords*— HWR; Pattern recognition, CRF; Lexicons**

I. Introduction

Handwriting Recognition (HWR) is a challenging issue because of the large data variability. It is an open research issue to recognize the handwritten words out of scanned documents which is unconstrained. Document complex structures hinder sections into words and lines. For recognizing word, the handwriting variability implicates other challenging level additionally. Figure 1demonstrate the issue of word recognition. In the classification below, word recognition can be categorized. With the top ranked matching lexicon, lexicon driven method employs a length lexicon that is fixed and links every word image in document. On robotic recognition of characters, Lexicon independent techniques and Segmentation free/Segmentation based model are used by employing its global features, segmentation free methods try and distinguish the whole word image.

At the same time, segmentation-based methods, depend on word breaking image into smaller sections recognizable as characters and link a character label to these sections. By employing a time series models like Hidden Markov Models (HMMs), contextual dependencies among nearest segments are used. Many accomplishments have been attained in this field employing lexicon driven segmentation depended methods. HMMs can be employed in segmentation-based methods to capture the spatial dependencies over nearest neighbors as they were generic models, and try to build joint probability model of label and data. In contrast to this Conditional Random Fields (CRFs) [1] model, the conditional distribution does not create any considerations over the data distribution. CRFs can obtain sum of feature functions unlike HMMs and every feature function can employ the whole sequence of input data. A basic restriction of maximum entropy Markov models (MEMMs) can be avoided by CRFs and other discriminative Markov models depending on directed graphical models that can be influenced to states with some successor states. The model HMMs obtains the transition to the similar label of the character otherwise to the subsequent word character and it can be employed for recognizing free word segments.

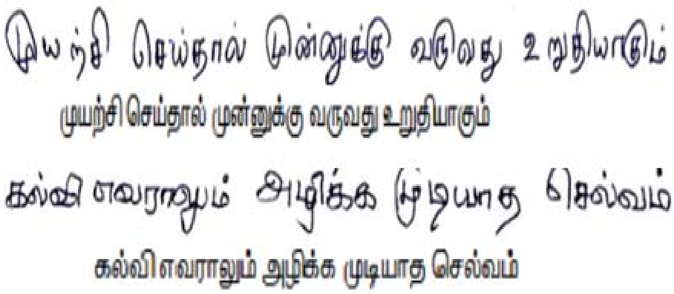


Fig. 1. Word recognition problem

At word level, CRFs were employed in [20] for identification however this method tends to a huge parameter counts as the model has been each word state parameter in the transition and lexicon attributes for each probable couple of adjacent words in lexicon. Depending on the segmentation of character, we employ a lexicon driven method, HMMs is searched to be inappropriate for this job due to its is instinctively wrong to employ character labels transition while comparing word image over a word that is given within lexicon. To relate the adjacent character labels with the feature of respective segments of handwritten character, at the same time CRFs employ transition characteristics. In this paper, with regard to the above, we have done a review of various HWR models. Depending on different perspectives, models available are reviewed.

The upcoming part of the survey is set as below: a review of ML method for DM prediction takes place in section 2 and conclusions are drawn in section 3.

II. Problem Domain

In spite of the extensive implication of computer in office, handwriting still stays as a significant mode of annotating and apprehending textual data. Computer-assisted handwriting is a highly important element of common interface among the business world and electronic media. Due to the huge compactness of personal digital assistants which is now a niche market, the major computing stream is known as Tablet PCs. A tablet PC is major input computer which is portable with a digital stylus and sensitive screen. The OS of that PC is built with element which capable to manage digital ink. And it is significant to know the differences among the character-recognition interfaces and digital ink. When it is a simple machine intelligence form that can recognize each letter of an alphabet which the keyboard might be exchanged through the same however more compactable, pen and tablet, a distinct form of input is demonstrated by the digital ink. For readers and for author, it exists in documents until enviable. Instead of initially converting semantically focused form like an ASCII string, words might replace, inserted or deleted. While the end copy is created, conversion might gradually happen.

For the applications of digital ink, there is a good balance such as one among the semantic substance and graphical form. From the pen immediacy input, one might get advantage, which it is infinite letters of word and it is informal by nature when at the other hand the machine can go through the ink structure which is capable to make changes at particular portions of it. The penetration depth requires being excellent, under a huge self-contained unit level like words and lines. At the same time, it is not treated really as handwriting, if no examination is performed on the ink input however it may be a common freehand graphical input. Subsequently computer support is much constrained (with the robotic formatting, placement and association in document).

Few studies aim at one phase of pen input semantic penetration: the word level. 'Word' refers to a collection of pen strokes which comprise a significance of lexical, that is, one which demonstrates a unique symbol or human language which might be employed. There is focus to segment robotically the sampled trace of a pen (digital int) into word segments which is truly based on spatiotemporal relations among continuous strokes, ignoring any meaning. This well-known issue in the research of handwriting recognition, it is seen as a precursor to whole character identification. [2] demonstrate that ‘‘to segment the signal into meaningful units the acquired data is generally preprocessed prior to any recognition".

By the survey [2], the word segmentation study history is constrained. With few assumptions that are given to stroke timing, employing convex shells, the major achievement is shown in this domain and are restricted to straightforward geometric segmentation. It is evidenced in [2], there are not prior proposals further in this field. One might contemplate further that what the consequence for no progress in this study is. Because of the individual writer’s idiosyncrasies and the fact that these methods fails to adopt more temporal signals and subtle structural that might strengthen the segmentation basis, our experience in this study demonstrates that simpler segmentation techniques are error pruned. Through the implication of hierarchical agglomerative clustering, an extensive robotic context of handwritten document structural analysis is presented that tries to enhance structure identification.

Although they are liable to write idiosyncrasies when being insensible to some repeating language features employed by the writer, thus these are choice of our interesting methods. The writing variability depends on one's inherent diversity and style of writing which it might lead to adaptive solution. To particular pen trace ad hoc metric, the solution might not be limited however might hold a huge collection of reasonable measures by considering the prime characteristics and any of the secondary one that might be important. Without priori evidence or formal basis, these characteristics are proposed still depending on plausibility. For a pen gesture recognizer, through [15] we have been guided by temporal and geometric classification.

III. Review of HWR Models

On the past decade, great efforts are paid to the online hand-drawn content interpretation from handwritten text [1,3], symbols and gestures, to structured scientific notations [4], sketches and diagrams or tables. The issue of examining free-form online documents, unconstrained has gaining an growing focus, and techniques have been projected for content block segmentation processing, segmentation of text lines otherwise for retrieving document structure.

A series of points assigned over the pen trajectory called as online note-taking document over the surface, where points are planned in strokes. For dividing a document, strokes provide a common unit and it is commonly considered which a stroke might get affected a unique kind as customers always lift up the pen out of the surface while text to non-text content [5] switching. For document understanding, separating the textual task out of non-textual strokes is assumed as a main issue in the study, which it might be assumed in isolation otherwise merged with the inter-related tasks of structure analysis and segmentation.

Based on how contextual data is used, many features of methods for non-text/ text stroke classification can be demonstrated. Without simple inclusion of local context description or any contextual data in this part, we initially demonstrate methods for stroke classification which is isolated. To support the stroke classification task, we project the methods of structured prediction, where the interactions among elements are used. We make difference over depending on document temporal structure and depending on spatial structure.

A. Isolated stroke classification

For segmenting online strokes as non-text or text, [6] initiated a local method in handwritten documents. Out of every stroke, two features are derived, without assuming the interaction with other strokes. With homogeneous regions and high accuracy (97%), a prediction is made by linear classifier of every stroke in unique like tables and text blocks are detected in next step. Over a difficult IAM-OnDo database, the similar features are implied with SVM (support vector machine) and it performed at a lower rate of 91.3%. Out of the geometrical shapes otherwise polygons, other researchers have modeled method based on entropy considering that high entropy of the online pen trajectory make difference in text writing. For stroke classification that is isolated, spectral features have been used with linear classification.

In online signal, for mode detection, [7] employed eight features set with nearest-neighbor classifier. In such circumstances, the elements of non-textual are categorized additionally into subgroups assumed as a various mode. With assuming few contextual data, it is useful to segment individually the strokes, and this decision problem always not to be given a solution undoubtedly. Based on whether it is positioned in a text line of drawing part, the similar stroke shaped as a similar circle can be assumed to be a non-text stroke or text stroke. From the nearest strokes, [8] projected to derive features. It was demonstrated that local context feature extraction enhances the prediction of stroke.

Many other study findings [5] make sure the observations. During stroke label prediction, the significance of assuming contextual knowledge is clearly highlighted. Researchers has modulated the stroke classification issue as structured prediction issue, that is, through designing combination of stroke labeling out of a document beyond including local context features from nearest one.

B. Exploitation of temporal context

In an online document, an accurate way to design interaction among stroke is to use the temporal data. The nature of online data recommends managing stroke labeling as a sequence labeling issue, where every stroke has to be labeled as non-text or text. The document can be classified into sub-series of strokes which are considered to split the similar kind due to its temporal distance which is lower than empirical threshold. Through using a richer context, the classification might be worked at subseries phase. The suitable strategy for segmentation is not usually clearly addressed and the consideration of label consistency on sub-series is frequently unchecked in practical datasets.

For sequence prediction, in a drawing series, for designing interaction among consecutive strokes, [9] modeled a hidden Markov model which is a probabilistic framework. The dependencies show the reality that two strokes that are written consecutively are probable to be of similar kind. Through a multi layer perceptron, emission probabilities for every stroke are estimated with 11 input attributes and the transition probabilities are computed from training data.

In HMM model, Labeling strokes mainly do well with MLP classifier the independent labeling. With bipartite HMM formulation, the designed model can be enhanced, wherever the gaps among strokes are assumed as additional hidden states and observations and are merged for modeling transitions among non-text and text states. HMMs disadvantage is that it considers the independence among assumptions that avoids from for prediction assuming local context of stroke label. The benefit in modeling dependencies among labels of nearest strokes temporally is reduced through weaker description limitation for every prediction of stroke label.

Depending on the bidirectional long short-term memory (BLSTM) neural networks (NN), [10] projected the most recently a mode detection method. BLSTM is kind of recurrent NN which have been implied effectively to predicting sequence. BLSTM is implied over document demonstrated as a feature vectors stream derived out of the sampled pen trajectory points. Through a bidirectional temporal context, label prediction is influenced as the stream of data is fed to both backward and forward network. The training step tunes the attributes for contextual interactions by the memory blocks mechanism for long short-term memory. To establish label over stroke level, a voting method is implied to the output point wise labeling at test time. This method gains most precise predictions. The benefits are trainable and it does not depend on heuristics in addition it needs any previous empirical segmentation. But, other context potential sources are avoided if the temporal context.

C. Exploitation of spatial context

In an online document, the strokes spatial distribution clearly offers valuable data for classifying stroke. In Japanese hand-drawn documents, [11] merged for stroke classification integrate spatial interactions in a Markov random field model. From the SVM classifier output that are trained independently, the interaction potential and observation potential function of MRF model are estimated. On a hidden Markov model, experiments show the MRF model superiority, therefore gaining significance of interactions among spatially nearest strokes. Because of the underlying independence considerations, but the MRF model does not allow enough local context utilization for stroke labels prediction as a common model.

D. Towards integration of multiple contexts

In online documents, merging spatial data and temporal data for stroke classification has been projected. The combination of successive spatial and temporal stroke collection for document segmentation into same blocks which can be divided as graphics or text. On classifying stroke, the contextual data cannot be trained and many attributes have to be adjusted manually, making it a questionable that it is autonomous.

E. Conditional Random Fields (CRF)

In past decade, Conditional Random Fields (CRF) [12] became popular models for sequence modeling, as they are discriminative models and they do not depend on similar restrictive considerations. In automatic language processing domain, to process symbolic data, the actual CRF framework was projected. A main disadvantage affecting CRF is its inability in processing numerical data as it can only process discrete rates

With a view to design the dependency among classes, while focusing numerical data, they are introduced commonly at a second model phase when raw numerical data are examined by a classification phase like Artificial Neural Networks (ANN) for instance in the Automatic Speech Recognition (ASR) domain. CRF model are constrained to observation series in spite of its capability to agreement with symbolic data, that is., to offer a label to every frame series. As a result, by combing language and/or lexicons, CRF is not capable to combine high level knowledge with HMMs. Against HMMs, a second constraint CRF is the need of comprising ground truth data at frame level therefore avoiding employing embedded training. By integrating the benefits of generative and discriminative together in early nineties, hybrid structures have been projected. Through merging ANN with HMM, they were modeled at the beginning for ASR. For HWR, hybrid model also been projected.

To classify and examine the frame level local observations, these models employ ANN discriminative stage, where the generative stage of HMM is dedicated to the combination of high level data like language models and lexicon and so on. For local posteriors estimated through ANN stage, the Gaussian Mixture Models (GMM) of the HMM stage is substituted. for sequence classification Bilateral Long Short Term Memory (BLSTM) neural networks merged with Connectionist Temporal Classification (CTC) stage is a powerful hybrid architecture as an alternative. With a discriminative classification stage created of a simpler logistic classifier, a structure merges a low level frame modeling stage that is efficient with the model ability at long time dependencies. For HWR and ASR, this structure has proven to do well. For language processing tasks, CRFk were actually modulated because of its theoretical features and implied in domain in that the numerical data ability is significant. Like ASR or (GR) [9], CRF model has been applied to application domains with a view to process this information. Few attempts have been suggested on employing CRF models in the HWR field.

To recognize character sequence, a CRF model is introduced. On an already segmented character series, this method is implicated in series, which is not designed through CRF stage. A combination of CRF and Deep Neural Network (DNN) which is called as non linear HCRF model is projected for recognition and segmentation of characters. At low level, the deep structure enhances discrimination when HCRF enables huge modeling level. Many of the existing works in survey have been build hybrid models that maps with observable raw data as low level input. To the next phase of hybrid architecture, Neural Networks like BLSTM, DNN, MLP models gives high informative feature level. Through the information that is given through primary phase, the second phase is dedicated to the examining contextual hypothesis. It is commonly depending on a generative model that establish limitations like language models and/or lexicons. In many circumstances, HMMs are executed however CTC which is a dynamic programming phases have proven to be probable alternate structure. At both high and low level phases, HCRF comprise specificity to be discriminative. However, it is constrained to the sequence labeling task that is trained for. During the time of decoding, it cannot implement high information level like language model or lexicon.

[13] introduced a hybrid CRF-HMM model for HWR with an intention to integrate the CRF discriminative abilities with the modeling abilities of the HMM. The CRF is applied to the discrimination of low level frame representations, whereas the HMM perform a lexicon-driven HWR. The low level frame representation is indicated by gram codebooks and HOG descriptors. The method undergoes training and testing on the open access HWR dataset.

[14] presented an efficient autonomous CRF based inferencing model to recognize characters. Here, a word is a series of linked characters and is attained by the use of various binarization models and diverse probable series are assumed by the use of a tree structure. CRF make use of contextual information for learning efficient fundamental series and identifies the possible labeling of the series by the use of many hypothesis tree to generate the proper series of alphabets. This method is highly applicable for reduced printed document images as it assumes many substitute hypotheses for proper decisions.

In [15], various models are explored to improve the results of the dependency models on discrete features for HWR. Hidden Markov Models has been employed for HWR. CRF enables the dependency and investigated the usage. It is believed that the initial try is to employ HWR. It is shown that the entire HWR process performs well than the HMM on the open access dataset.

[16] intends to compare the results of the discriminative and generative recognition mechanisms that are explained by generatively-trained hidden Markov modeling (HMM), discriminatively-trained conditional random fields (CRF) and discriminatively-trained hidden-state CRF (HCRF). By learning samples attained from two different database undergo training and employed an HMM classification method. For enabling the HMM classifier for efficiently reject the mistake and out-of-vocabulary segmentation. It enhances the models with adaptive threshold methods. In addition, the performance of these three mechanisms is totally validated by the use of

diverse dataset. The HWR outcome of the words as well as letters is projected, with the advantages as well as disadvantages for each method is emphasized.

F. Other Methods

[17] developed effective multi-dimensional long short-term memory (MDLSTM) based approaches for models for neural HR specially intended to eliminate computation wastages produced from padding. The presented example-packing method will replace the wasteful stacking of padded examples with effective tiling in a 2-dimensional grid. For word-based NHR, it produces a speed enhancement of factor 6.6 over an already effective baseline of minimal padding for every batch individually. For line-based NHR, the savings are highly modest. Besides, a detailed validation of the IAM dataset also takes place and a comparison with the state of art methods. The effective NHR method and reusable method explained offers various approaches for realizing effective methods for the omnipresent scenario of variable-length inputs in deep learning.

Regardless of being an important language globally, this study developed a Bengali handwritten numeral recognition (BHNR) [17]. The available approaches are mainly based on feature extraction and older machine learning algorithms. The current trend in machine learning is deep neural network particularly by the use of Convolutional Neural Network (CNN) that shows significant performance with precise results. The presented method recognizes numerals with high degree of accuracy approximately 96% even in most challenging noisy conditions. At the beginning, 72000+ specimens from NumtaDB (85000+) has been employed to train 17000+ specimens and also employed as test dataset. These enhancements in results from difficult situations have been noted during the training of different specimens.

A deep learning based online handwriting system [18] is presented which has the ability to support 102 languages. It is entirely replaced our previous Segment-and-Decode-based system and minimized the error rate by 20%-40% relative for most languages. In addition, recent methods on IAM-OnDB dataset is tested and compared to one another. It results to 10 times faster recognition when compared to previous method.

The IAM-OnDB dataset [19] is the commonly employed evaluation dataset for online handwriting recognition. It has a total of 298 523 characters in 86 272 word instances from a dictionary of 11 059 words written by 221 writers. The standard IAM-OnDB dataset is employed for separation: one training set, two validations sets and a test set with 5 363, 1 438, 1 518 and 3 859 written lines, correspondingly. The weights of the decoder are tuned by the validation set with 1 438 items and reported the error rates on the test set. A detailed analysis of the number of layers and nodes per layer for raw as well as curve input formats to determine the optimal size of the bidirectional LSTM network (see Fig. 2, Table 1). Experiments are carried out with no extra feature functions (Fig. 2, solid lines), and compute the performance with altered weights for language models and character classes. It is observed that for every input format, 3 or 5 layers outperforms more shallow networks, and using more layers gives hardly any improvement. In addition, the use of 64 nodes per layer is enough, since wider networks offer slight enhancements. At the end, a comparative results with the existing methods takes place on Table 2. It offers better results over the compared methods.

Table 1 Comparison of character error rates on the IAM-OnDB test set for different LSTM layers configuration

input	lstm	64 nodes	128 nodes	256 nodes
raw	1 layer	6.1	5.95	5.56
	3 layers	4.03	4.73	4.34
	5 layers	4.34	4.20	4.17
curves	1 layer	6.57	6.38	6.98
	3 layers	4.16	4.16	4.83
	5 layers	4.02	4.22	4.11

Table 2 Error rates on the IAM-OnDB test set

System	CER[%]	WER[%]
Frinken et al. BLSTM [20]	12.3	25.0
Graves et al. BLSTM	11.5	20.3
Liwicki et al. LSTM [21]	-	18.9
this work (curve, 5x64, no FF)	5.9	18.6
this work (curve, 5x64, FF)	4.0	10.6
our previous BLSTM [22]	8.8	26.7
combination	-	13.8
Segment-and-Decode [22]	4.3	10.4
this work (production system)	2.5	6.5

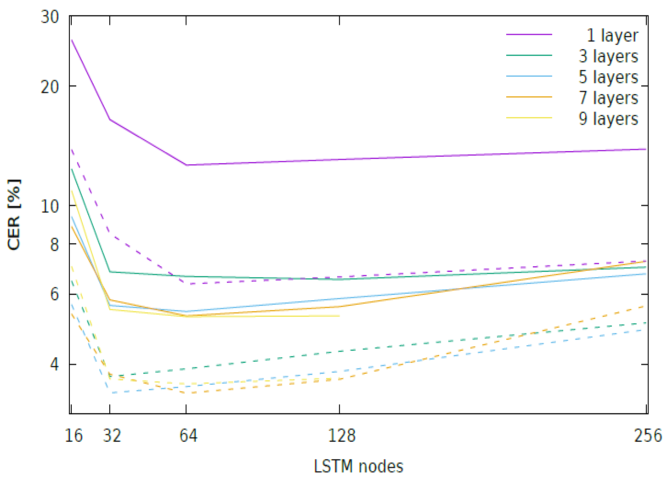


Fig. 2. CER of models trained on the IAM-OnDB dataset

IV. Conclusion

On the past decade, great efforts are paid to the online hand-drawn content interpretation from handwritten text HWR is a challenging issue because of the large data variability. It is an open research issue to recognize the handwritten words out of scanned documents which is unconstrained. In this paper, with regard to the above, we have done a review of various HWR models. Depending on different perspectives, models available are reviewed. In future, we try to develop a new HWR model using CRF particularly for Tamil language which will be useful to recognize the Tamil characters from handwriting.

REFERENCES

[1] Plamondon, Réjean, and N. Sargur Srihari, "Online and off-line handwriting recognition: a comprehensive survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63-84, 2000.

[2] Yoshua Bengieo, Yann Le Mc Cun and Chris Burges, Le Rec: "ANN/HMM hybrid for on-line handwriting recognition", Neural Computation, vol.5, pp. 1289-1303, Nov 1995.

[3] C.L. Liu, Stefan Jaeger and Masaki Nakagawa, "Online recognition of Chinese characters: the state-of-the-art", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 2, pp. 198-213, 2004.

[4] Mouchere, Harold, Christian Viard-Gaudin, Dae Hwan Kim, Jin Hyung Kim and Utpal Garain, Crohme2011: Competition on recognition of online handwritten mathematical expressions, *In Proceedings of the 11th International Conference on Document Analysis and Recognition*, pp. 1497-1500, 2011.

[5] E. Indermühle, Analysis of Digital Ink in Electronic Documents, Ph.D. Thesis, University of Bern, 2012.

[6] K. Jain, M. Anoop Namboodiri and Jayashree Subrahmonia, Structure in on-line documents, *In Proceedings of the 6th International Conference on Document Analysis and Recognition*, pp. 844–848, 2004.

[7] D. J. M. Willems, S. P. Rossignol and L. G. Vuurpij, Features form detection in natural online pen input, *In Proceedings of the 12th Biennial Conference of the International Graphonomics Society*, pp. 113–117, 2005.

[8] Grosicki, Emmanuèle and Haikal El Abed., ICDAR 2009 handwriting recognition competition, *In Proceedings of the IEEE Tenth International Conference on Document Analysis and Recognition*, pp. 1398-1402, 2009.

[9] Bishop, Christopher, Markus Svensen and Goeffrey Hinton, Distinguishing text from graphics in on-line handwritten ink, In Proceedings of the IEEE 9th International Workshop on Frontiers in Handwriting Recognition, pp. 142-147, 2004.

[10] Indermühle, Emanuel, Volkmar Frinken and Horst Bunke, Mode detection in online handwritten documents using BLSTM neural networks, *In Proceedings of the IEEE 13th International Conference on Frontiers in Handwriting Recognition*, pp. 302-307, 2012.

[11] X.D. Zhou, and C.L. Liu, Text/non-text ink stroke classification in Japanese hand- writing based on markov random fields, *In Proceedings of the IEEE 9th International Conference on Document Analysis and Recognition*, vol. 1, pp. 377–381, 2007.

[12] Quattoni, Ariadna, Michael Collins and Trevor Darrell, Conditional random fields for object recognition, In Advances in neural information processing systems, pp. 1097-1104, Dec 2005.

[13] Bideault, Gautier, Luc Mioulet, Clément Chatelain and Thierry Paquet, A hybrid CRF/HMM approach for handwriting recognition, *In Proceedings of the International Conference on Image Analysis and Recognition*, pp. 403-410, Springer, Cham, 2014.

[14] Ray, Anupama, Ankit Chandawala and Santanu Chaudhury, Character recognition using conditional random field based recognition engine, *In Proceedings of the IEEE 12th International Conference on Document Analysis and Recognition*, pp. 18-22, Aug 2013.

[15] Feng, Shaolei, R. Manmatha and Andrew McCallum, Exploring the use of conditional random field models and HMMs for historical handwritten document recognition, *In Proceedings of the 2nd International Conference on Document Image Analysis for Libraries*, 2006, pp. 1-8, Apr 2006.

[16] Elzobi, Moftah and Ayoub Al-Hamadi, Generative vs. Discriminative Recognition Models for Off-Line Arabic Handwriting, Sensors 18, no. 9, pp. 2786, Sep 2018.

[17] G.M. Wenniger, L. Schomaker, A. Way, "No Padding Please: Efficient Neural Handwriting Recognition". arXiv preprint arXiv:1902.11208. Feb 28, 2019.

[18] R. Noor, K.M. Islam and M.J. Rahimi, "Handwritten Bangla Numeral Recognition Using Ensembling of Convolutional Neural Network". In *2018 21st International Conference of Computer and Information Technology (ICCIT)*(pp. 1-6). IEEE, December, 2018.

[19] V.Carbune, P. Gonnet, T. Deselaers, H.A. Rowley, A. Daryin, M. Calvo, L.L. Wang, D. Keysers, S. Feuz, and P. Gervais, "Fast Multi-language LSTM-based Online Handwriting Recognition", 2019, *arXiv preprint arXiv:1902.10525*.

[20] V. Frinken, S. Uchida, "Deep BLSTM neural networks for unconstrained continuous handwritten text recognition", *In: ICDAR*, 2015.

[21] M. Liwicki, H. Bunke, J.A. Pittman, S. Knerr, "Combining diverse systems for handwritten text line recognition", Mach. Vis. Appl. Vol.22, no.1, pp.39–51, 2011.

[22] D. Keysers, T. Deselaers, H. Rowley, L.L. Wang, V. Carbune: "Multi-language online handwriting recognition", IEEE Trans. Pattern Analysis & Machine Intelligence Vol.39, no.6, pp.1180–1194, 2017.