International Journal of Engineering Research and Management (IJERM) – Volume 2 Issue 5- September 2016

Recognition of Handwritten Devanagari Words using Recurrent Neural Network: A Review

Subhash V. Pingale[1], Rajan S. Jamgekar[2], Rutuja N. Benkar[3]

[#]ME Computer (Engineering), Dept. of Computer Science & Engineering, Solapur University,
SKN Sinhgad College of Engineering, Korti, Pandharpur, Maharashtra, India
[1]sub.pingale83@gmail.com
[2]rs.jamgekar@gmail.com
[3]rutujabenkar06@gmail.com

***Abstract*— Handwritten Word Recognition is an important problem of Pattern Recognition. In India, more than 300 million people use Devanagari script for documentation. There has been a significant improvement in the research related to the recognition of printed as well as handwritten Devanagari text in the past few years. Though Devanagari is the script for Hindi, which is the official language of India, its character and word recognition pose great challenges due to large variety of symbols and their proximity in appearance. Offline handwritten recognition system for Devanagari words is still in developing stage and becoming challenging due to the large complexity involvement. The difficulty of segmenting overlapping characters, combined with the need to exploit surrounding context, has led to low recognition rates for even the best current recognizers. Most recent progress in the field has been made either through improved pre-processing or through advances in language modelling.**

Recurrent neural networks, specifically designed for sequence labelling tasks where the data is hard to segment and contains long-range bidirectional interdependencies. In addition, to demonstrate the network's robustness to lexicon size, measure the individual influence of its hidden layers, and analyse its use of context.

Keywords—*Devanagari script, Offline handwritten recognition, Online handwritten recognition, Recurrent Neural Network*

I. Introduction

Pattern Recognition is important Component of intelligent system. Pattern Recognition concern with Description or Classification of measurement. Handwritten Word Recognition is an important problem of Pattern Recognition. Handwriting recognition can be divided into two categories: Offline and On-line. This division is done on the basis of representing data to the system. For the first kind of recognition system user's handwriting is digitized by a scanner or a camera at a later time and the data is presented to the system as an image while for the second kind, handwriting is digitized by a tablet and stylus at the time of writing and strokes are captured as they are being formed by sampling the pen's position at evenly spaced time intervals. A pressure sensitive switch on the tip of the pen detects pen-up/ pen-down position and discriminates stroke segmentation.

In online recognition, a time series of coordinates, representing the movement of the pen tip, is captured, while in the offline case, only an image of the text is available. Because of the greater ease of extracting relevant features, online recognition generally yields better results. So far, the majority of systems have tackled the easier of the two problems, namely, the on-line problem where the time ordering of strokes is available as well as pen up/down information; overlapping strokes can easily be distinguished and stroke positions are accurately known. On the other hand, offline systems have to cope with the large varieties of pen type, wide strokes which overlap and a lack of ordering information. Thus, the research in Offline recognition is more complex than the On-line version. The main disadvantage of on-line handwriting recognition is that the writer is required to use special equipment. So offline handwriting recognition is preferred.

Recurrent neural networks (RNN) have been successfully applied for recognition of cursive handwritten documents, both in English and Arabic scripts. Ability of RNNs to model context in sequence data like speech and text makes them a suitable candidate to develop OCR systems for printed Devanagari scripts. A regular recurrent neural network (RNN) is extended to a bidirectional recurrent neural network (BRNN). The BRNN can be trained without the limitation of using input information just up to a preset future frame. This is accomplished by training it simultaneously in positive and negative time direction. Bidirectional Long Short Term Memory (BLSTM) architecture was employed to recognize printed Devanagari text.

II. Literature Review

Recognizing lines of unconstrained handwritten text is a challenging task. The difficulty of segmenting cursive or overlapping characters, combined with the need to exploit surrounding context, has led to low recognition rates for even the best current recognizers. Relatively little work has been done on the basic recognition algorithms. Indeed, most systems rely on the same hidden Markov models (HMM) that have been used for decades in speech and handwriting recognition, despite their well-known shortcomings. [1] Proposes an alternative approach based on a novel type of recurrent neural network (RNN), specifically designed for sequence labeling tasks where the data is hard to segment and contains long-range bidirectional interdependencies. The experiments on two large unconstrained handwriting databases achieve word recognition accuracies of 79.7 percent on online data and 74.1 percent on offline data.

A scheme for offline Handwritten Devanagari Character Recognition is proposed in [2], which uses different feature extraction and recognition algorithms. This system assumes no constraints in writing style, size or variations. First the character is preprocessed and features namely: Chain code histogram, four side views, shadow based are extracted and fed to Multilayer Perceptrons as a preliminary recognition step. Finally the results of all MLP's are combined using weighted majority scheme. This system is tested on 1500 handwritten Devanagari character database collected from different people. It is observed that the system achieves 98.16% recognition rates.

[3] Developed an effective method for recognition of isolated Marathi handwritten numerals written in Devanagari script. Fourier Descriptors that describe the shape of Marathi handwritten numerals are used as feature. 64 dimensional Fourier Descriptors represents the shape of numerals, invariant to rotation, scale and translation. Three different classifiers, namely, nearest neighbourhood (NN), K-nearest neighbourhood (KNN) and Support Vector Machine (SVM) are used independently in order to recognize test numeral. These classifiers are trained with 64 dimensional Fourier Descriptors (FD) of training samples. The proposed system is experimented with a database of 13000 samples of Marathi handwritten numerals using fivefold cross validation method for result computation. An overall recognition rate of 97.05%, 97.04% and 97.85% are obtained for NN, KNN and SVM respectively.

The work reported in [4] presents a two-stage classification approach for handwritten Devanagari characters. The first stage is using structural properties like shirorekha and spine in a character. A differential distance based technique is designed to find a near straight line for shirorekha and spine. The second stage exploits intersection features of characters, which are then fed to a feed forward neural network (FFNN) for further classification. This approach has been tested for 50000 samples and got 89.12% success.

HMM are used to recognize the on-line and off-line handwritten texts acquired from a whiteboard and text document. Recognizing handwritten devanagari words using recurrent neural network 415 pseudo characters in [5]. The word level recognition is done on the basis of string edit distance. The system is based on the combination of several individual classifiers of diverse nature. Recognizers based on different architectures (hidden Markov models and bidirectional long short-term memory networks) and on different sets of features are used in the combination. In order to increase the diversity of the underlying classifiers in cursive handwriting recognition, commercial recognition systems have been included in the combined system, leading to a final word level accuracy of 86.16%. This value is significantly higher than the performance of the best individual classifier (81.26%).

The details of many handwritten Devanagari numeral, character, and word recognition systems are summarized in following Table 1.

TABLE 1: **LITERATURE SURVEY TABLE**

Data Set (size)	**classifiers**	**Results (%)**
13,040 handwritten lines, 86,272 characters 11,050 distinct English words. [1]	HMM (Word accuracy) RNN (Word accuracy) RNN (Character accuracy)	64.5% 74.1% 81.8%
1500 handwritten Devanagari basic characters. [2]	Multi-Layer Perceptron (MLP)	98.61%
13000 samples of Marathi handwritten numerals. [3]	Neural Network (NN) K-Nearest Neighbour (KNN) Support Vector Machine (SVM)	97.05% 97.04% 97.85%
50000 samples of Devanagari characters. [4]	Neural Network (NN)	89.12%
13,00 written lines, 86272 characters and 11,050 distinct cursive words. [5]	BLSTM HMM Combining BLSTM and HMM	81.26% 79.2% 86.16%

From the literature survey it is came to know that recurrent neural network gives better results than other techniques. Recurrent neural network have greater representational power and the ability to perform intelligent smoothing by taking into account syntactic and semantic features. Also there are different techniques for feature extraction. Lots of work has been done on Devanagari numerals, characters and very few were done on words. Recognition rate of handwritten Devanagari numerals is very good, but for words it is less.

III.COMPARATIVE STUDY OF DIFFERENT HANDWRITTEN WORD RECOGNITION METHODS

There are two most popular approaches for handwriting recognition: Hidden Markov Models (HMMs) and BLSTMs (Bidirectional Long-Short Memory). Both approaches are efficient for recognizing line of text since they are segmentation-free. A sliding window algorithm is used for the segmentation free word recognition. In HMMs, frame dependency with neighbouring ones is taken into account by 1st order regressions. But in the bi-directional architecture of

BLSTMs, frame dependency with long-distant ones on its left and right, is taken into account by using the forward and backward recurrent connections. Since frame dependency is implemented in very different ways for HMMs and BLSTMs, it is of interest to compare the feature extraction parameter setting obtained for each approach. HMMs are based on the stochastic modelling of observed frames while BLSTM are based on recurrent neural networks. Both take as input frame sequences extracted from a sliding window.

A. *HMM Model*

In HMM modelling, characters are represented as a succession of states with left-right transitions and a self-transition. In each state, observations follow a Gaussian mixture probability density function. Such approach leads to increasing number of parameters, compared to a context-independent HMM. The state-tying is based on decision-tree clustering. These decision trees are based on the right and left context of a central character. A decision tree is built for each central letter and node splitting is decided according to two parameters: cluster minimal occupancy and likelihood maximization. This approach presents the advantage of allowing trigraph models. Cluster minimal occupancy and threshold both have been optimized on a distinct validation database. This approach led to improved recognition performance so this system participated in Arabic & Latin handwritten word recognition competitions. To model full text lines with HMMs, concatenate word models with space models in-between. Word models themselves are made up of the succession of its compound character models. Performance also depends on feature extraction parameters. Decoding is done by a Viterbi token passing algorithm through lattices. These lattices can include language model probabilities. Pruning is performed at token and word levels.

B. *BLSTM Model*

Use a BLSTM (Bidirectional Long-Short Term Memory) with one hidden layer containing 100 blocks. In BLSTM, classical neural network units are replaced with LSTM (Long Short-Term Memory). Those blocks can keep information through more than 1000 time samples: LSTM includes a memory cell and multiplicative logical gates which are specifically designed to memorize or forget relevant information through time. Those gates can pass or block signals. BLSTMs are bidirectional recurrent networks: They consist of the coupling of 2 recurrent neural networks. They take a mono-dimensional signal as input, sequence of frames from 1 to T and introduce two contexts in the image, past and future. Thus, contextual information is taken into account from both left-to-right and right-to left handwriting directions.
So select this RNN architecture rather than a Multidirectional RNN since its input is similar to HMMs that is frames are extracted by a sliding window. Sharing the same inputs makes it easier to compare recognizers. Here consider a network with only one hidden layer. The hidden layer is made of two independent neural layers: the forward layer which takes the original frame sequence as input from 1 to T and the backward layer which takes the reversed sequence from T to 1 as a input in Fig.1. Hence, the value of an output unit at time step t is the sequential combination of the outputs of the forward and backward hidden layers at this time step t. forward and backward layer are processed independently through training and decoding phases. A backward-forward algorithm referred to as CTC (Connectionist Temporal Classification) token passing that takes the posteriors as input and provides a sequence of words given to dictionary and a language model.

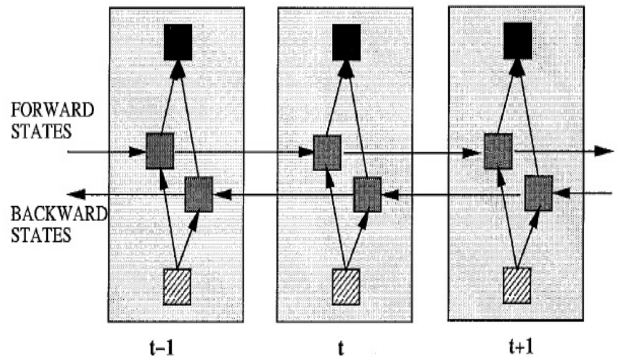

Fig. 1 BRNN unfolded through 3 times steps

IV. SCOPE

The proposed system will be useful to recognize the simple Devanagari handwritten legal amount words in Bank Sector. It will be used to recognize such handwritten legal amounts present on Indian bank cheques. It will also useful in Cheque Truncation. Cheque Truncation speeds up the process of collection of cheques resulting in better service to customers, reduces the scope of loss of instruments in transit, lowers the cost of collection of cheques, thus benefitting the system as a whole. Handwritten recognition systems are also used in Forensic Department & in Manual Form filling Process.

V. CONCLUSION

As Handwritten Devanagari Script Recognition has a huge scope in many areas, the researchers should use the most efficient techniques to get the desired results. Recognizing the handwritten Devanagari word is difficult task. Thus this system is implemented for increasing accuracy rate by using RNN classifier.

References

[1] Alex Graves, Marcus Liwicki, Santiago Fernandez, Roman Bertolami, Horst Bunke, and Jurgen Schmidhuber, "A Novel Connectionist System for Unconstrained Handwriting Recognition", IEEE transactions on pattern analysis and machine intelligence, vol. 31, no. 5, may 2009.

[2] S. Arora, D. Bhattacharjee, M. Nasipuri , D.K. Basu2 , M.Kundu, L.Malik, "Study of Different Features on Handwritten Devanagari Character", Second International Conference on Emerging Trends in Engineering and Technology, ICETET-09, 2009 IEEE.

[3] G. G. Rajput, S. M. Mali, "Fourier Descriptor based Isolated Marathi Handwritten Numeral Recognition", International Journal of Computer Applications (0975 – 8887) Volume 3 – No.4, June 2010.

[4] Sandhya Arora, Debotosh Bhatcharjee, Mita Nasipuri, Latesh Malik, "A Two Stage Classification Approach for Handwritten Devanagari Characters", International Conference on Computational Intelligence and Multimedia Applications, 2007 IEEE.

[5] Marcus Liwicki, Horst Bunke, James A. Pittman, Stefan Knerr, "Combining diverse systems for handwritten text line recognition" , Springer-Verlag 2009, 23 July 2009.

[6] Pooja Agrawal, M. Hanmandlu, Brejesh Lall, "Coarse Classification of Handwritten Hindi Characters", International Journal of Advanced Science and Technology Vol. 10, September, 2009.

[7] R. Jayadevan, Satish R. Kolhe, Pradeep M. Patil, and Umapada Pal, "Off-Line Recognition of Devanagari Script: A survey", IEEE TRANSACTION ON SYSTEMS, MAN, AND CYBERNETICS - PART C: APPLICATIONS AND REVIEWS, VOL. 41.NO.6, NOVEMBER 2011.

[8] Umapada Pal, Sukalpa Chanda Tetsushi Wakabayashi, Fumitaka Kimura, "Accuracy Improvement of Devanagari Character Recognition Combining SVM and MQDF", In Proc. 11^{th} ICFHR, pp.367-372, 2008.

[9] Satish Kumar," Performance Comparison of Features on Devanagari Hand-printed Dataset", International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009.

[10] Marcus Liwicki, Alex Graves, Horst Bunke J¨urgen, Schmidhuber, "A Novel Approach to On-Line Handwriting Recognition Based on Bidirectional Long Short-Term Memory Networks", Proc. 9th Int. Conf. on Document Analysis and Recognition 1, 367-371, 2007.

[11] Swati Kinhekar , Sharvari S. Govilkar "Comparative Study of Segmentation and Recognition Methods for Handwritten Devnagari Script", *International Journal of Computer Applications (0975 – 8887) Volume 105 – No. 9, November 2014*.

[12] R. Plamondon and S.N. Srihari, "On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63-84, Jan. 2000.

[13] S. Espana-Boquera, M. J. Castro-Bleda, J. Gorbe-Moya, and F. Zamora-Martinez, "Improving offline handwritten text recognition with hybrid HMM/ANN models," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, pp. 767–779, 2011.

[14] A. Graves, S. Ferna´ndez, and J. Schmidhuber, "Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition," Proc. Int'l Conf. Artificial Neural Networks, pp. 799- 804, 2005.

[15] A. Graves, S. Fernandez, M. Liwicki, H. Bunke, and J. Schmidhuber, "Unconstrained Online Handwriting Recognition with Recurrent Neural Networks," Advances in Neural Information Processing Systems 20, J. Platt, D. Koller, Y. Singer, and S. Roweis, eds., 2008.

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.* 9, pp. 1735–1780, Nov. 1997.