

Review on Offline Handwritten Character Recognition using Feed Forward Neural Network and SURF Feature

Reetika Verma¹, Mrs. Rupinder Kaur²

M.Tech Student, Department of Computer Science & Engineering, RIMT-IET (MandiGobindgarh, Punjab), India¹

Assistant Professor, Department of Computer Science & Engineering, RIMT-IET (MandiGobindgarh, Punjab), India²

Abstract: Image pre-processing is the name for operations on images at the lowest level of abstraction whose aim is an improvement of the image data that suppress undesired distortions or enhances some image features important for further processing. It does not increase image information content. Image pre-processing tool, created in Matlab, realizes many brightness transformations and local pre-processing methods. The proposed solutions focus on applying Neural Network Algorithm model for character recognition. The primary function of which is to retrieve in a character stored in memory, when an incomplete or noisy version of that character is presented. The idea is to create a theoretical and practical basis of preprocessing for character recognition using forward-feed neural networks. The Feed Forward Algorithm gives insight into the enter workings of a neural network; followed by the Back Propagation Algorithm which compromises Training and Testing.

Keywords: Character Recognition, Feed Forward Neural Network, Multi layer Perceptron, Supervised learning, SURF Feature.

I. INTRODUCTION

A Neural Network is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of the NN paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. A NN is configured for a specific application, such as character recognition or data classification, through a learning process [1]. Neural network adopt various learning mechanism of which supervised learning and unsupervised learning methods have turned out to be very popular. In supervised learning, a teacher is assumed to be present during the learning process, i.e. the network aims to minimize the error between target (desired) output presented by the teacher and the computed output to achieve better performance. However, in unsupervised learning, there is no teacher present to hand over the desired output and the network therefore tries to learn by itself, organizing the input instances of the problem. NN Architecture has been broadly classified as single layer feed forward networks, multilayer feed forward networks and recurrent networks, over the year several other NN [3]. The training patterns are applied in some random order one by one, and the weights are adjusted using the back propagation learning law. Each application of the training set patterns is called a cycle. The patterns have to be applied for several training cycles to obtain the output error to an acceptable low value. Once the network is trained, it can be used to recall the appropriate pattern for a new input pattern. The computation for recall is straightforward, in the sense that the weights and the output functions of the units in different layers are used to compute the activation values and the output signals [4].

The signals from the output layer correspond to the output. Artificial neural network have been developed as generalization of mathematical model of human cognition or neural biology based on the assumption:

- Information processing occurs at a simple element x called neuron.
- Signals are passed between the neurons over connected link. Each connection link has associated weight which in typical neural net multiply the signals transmitted.
- Each neuron applies to its activation function to its net input to determine the output signal [2].

Many reports of character recognition in English have been published but till high recognition accuracy and minimum training time of handwritten English characters using neural network is an open problem. Therefore, it is a great important to develop an automatic handwritten character recognition system for English language [3].

In this paper, efforts have been made to develop automatic handwritten character recognition system for English language with high recognition accuracy and minimum training and classification time.

Experimental result shows that the approach used in this paper for English character recognition is giving high recognition accuracy and minimum training time [4].

In contrast to limitations of classical computing, Artificial Neural Networks (ANNs), that were first developed in the mid 1900's serve for the emulation of human thinking in computation to a meager, yet appreciable extent [4].

i) Steps For Character Recognition

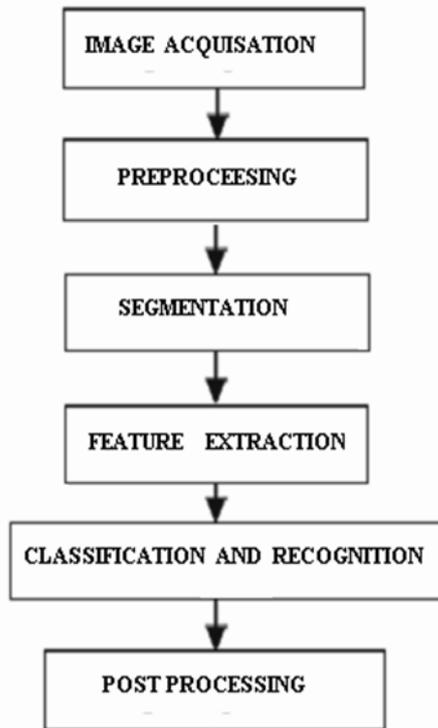


Fig 1: Generic Character Recognition Model

A. Image Acquisition

In Image acquisition, the recognition system acquires a scanned image as an input image. The image should have a specific format such as JPEG, BMT, etc. This image is acquired through a scanner, digital camera or any other suitable digital input device [6, 7].

B. Pre-Processing

The pre-processing is a series of operations performed on scanned input image. It essentially enhances the image rendering it suitable for segmentation. The role of pre-processing is to segment the interesting pattern from the background. Generally, noise filtering, smoothing and normalization should be done in this step. The pre-processing also defines a compact representation of the pattern. Binarization process converts a gray scale image into a binary image. Dilation of edges in the binarized image is done using sobel technique [10].

C. Segmentation

In the segmentation stage, an image of sequence of characters is decomposed into sub-images of individual character [9]. In the proposed system, the pre-processed input image is segmented into isolated characters by assigning a number to each character using a labelling process. This labelling provides information about number of characters in the image. Each individual character is uniformly resized into pixels.

D. Feature Extraction

In this stage, the features of the characters that are crucial for classifying them at recognition stage are extracted. This is an important stage as its effective functioning

improves the recognition rate and reduces the misclassification. Diagonal feature extraction scheme for recognizing off-line handwritten characters is proposed in this work [8, 9]. Every character image is divided into equal zones, each of size 10x10 pixels. The features are extracted from each zone pixels by moving along the diagonals of its respective 10x10 pixels.

E. Classification and Recognition

The classification stage is the decision making part of the recognition system [3]. A feed forward back propagation neural network is used in this work for classifying and recognizing the handwritten characters. The pixels derived from the resized character in the segmentation stage form the input to the classifier. The neural classifier consists of two hidden layers besides an input layer and an output layer. The total number of neurons in the output layer is 26 as the proposed system is designed to recognize English alphabets [6].

F. Post-processing

Post-processing stage is the final stage of the proposed recognition system. It prints the corresponding recognized characters in the structured text form by calculating equivalent ASCII value using recognition index of the test samples.

II. LITERATURE REVIEW

S.P. Kosbatwa, S.K. Pathan [3]: In this paper, use of artificial neural network in applications can dramatically simplify the code and improve quality of recognition while achieving good performance. Another benefit of using neural network in application is extensibility of the system – ability to recognize more character sets than initially defined. Pattern association using back propagation algorithm is essential and helpful to optimize the association of input pattern to output pattern in the neural network. This approach can be used for pattern classification.

Kauleshwar Prasad [5]: In this paper, focus on recognition of English alphabet in a given scanned text document with the help of Neural Networks. This paper carries out a study of various feature based classification techniques for offline handwritten character recognition. The feature extraction step of optical character recognition is the most import. At the current stage of development, the software does perform well either in terms of speed or accuracy but not better. It is unlikely to replace existing OCR methods, especially for English text. A simplistic approach for recognition of Optical characters using artificial neural networks has been described.

Sunith Bandaru [11]: Develop a multi-layered neural network based algorithm using which a computer can learn to identify handwritten characters. A GUI is designed in MATLAB which enables the user to either train or test the network on a ‘one character at a time’ basis. We have used a six element feature vector which is found to be sufficient for reliably identifying all characters that can be entered using a standard US English QWERTY keyboard.

Jingxin Hong, Wu Lin, HaoZhang [12]: In the paper, an image mosaic algorithm based on SURF feature matching is proposed. The algorithm uses SURF operator which has strong robustness and superior performance to extract features instead of conventional SIFT operator. The extracted features are matched by a novel matching scheme-fast bidirectional matching.

Ivan Kastelan, Sandra Kukolj [13]: In this paper, proposed a system for text extraction based on the open-source OCR algorithm. Preparation steps for OCR were developed which detected text regions in the image, and OCR was run on detected regions to read the text. The system is used for functional verification of TV sets. Text on the grabbed image is read in order to verify whether the TV responds to remote control commands successfully.

Anita Pal [16]: Neural network is playing an important role in handwritten character recognition. Many reports of character recognition in English have been published but still high recognition accuracy and minimum training time of handwritten English characters using neural network is an open problem. Therefore, it is a great important to develop an automatic handwritten character recognition system for English language. In this paper, efforts have been made to develop automatic handwritten character recognition system for English language with high recognition accuracy and minimum training and classification time.

Sumedha B. Hallale [17]: In this paper, Optical character recognition is a typical field of application of automatic classification methods. In this paper, we have introduced a whole new idea of recognition of isolated handwritten digits which is known to be a difficult task and still lacks a satisfactory technical solution. Handwritten character recognition is a field of image processing as well as pattern recognition. There are two approaches for the pattern recognition such as statistical and structural. In statistical approach, the series of characteristic measurements of the input data is generated on the statistical basis and is assigned to one of the n classes. The structural description of the object is based on the interconnections and interrelationships of features of input data. In general, both approaches are widely used in the pattern recognition.

Rakesh Kumar Mandal [18]: In this paper, to improve the performance of the previously applied methods. The input image matrix is compressed into a lower dimension matrix in order to reduce non -significant elements of the image matrix. Compared to the previously tried methods of segmentation Column-wise segmentation is a better approach than the row-wise segmentation as segmentation of the matrix column-wise produces more variation in the patterns and helps to obtain a sharp eye in the identification of the pattern.

Muhammad Naeem Ayyaz1 [19]: Proposed a handwritten character recognition system based on a hybrid feature extraction technique has been presented. The system comprised three main stages, i.e. pre-processing, feature extraction technique, and SVM based

training/classification. The proposed hybrid feature extraction technique, as experiments revealed, proved to capture local and global variations in handwritten character styles. The extracted feature vector was a combination of correlation function based features and some statistical/structural features.

III. CONCLUSION

In this paper, we propose the neural networks and Surf Feature Extraction technique for Character Recognition and enhancement. BPNN and Surf Feature technique will be used. The evaluation will be based on the PSNR, MSE. The proposed approach i.e., improved technique for Character Recognition using Neural Network will exhibit outcomes of noise reduction and image quality improvements, with different noise levels, which will qualify it to be suitable for image processing and Pattern matching.

ACKNOWLEDGEMENT

I would like to thank my guide Er. Rupinder Kaur, Assistant Professor in Department of Computer Science and Engineering at RIMT-IET (Mandi Gobindgarh), Punjab, India for motivating me to do research work on the topic Review on Offline Handwritten Character Recognition using Feed Forward Neural Network and SURF Feature. I would also like to thank my mother, my father and my siblings for their continuous support, cooperation, and guidance throughout my M.Tech work. Moreover I would also like to thank my Professors who were always there at the need of the hour and provided with all the help and facilities, which I required, for my thesis work.

REFERENCES

- [1] http://www.myreaders.info/03_Back_Propagation_Network.
- [2] Handwritten Optical character recognition http://en.wikipedia.org/wiki/Optical_character_recognition.
- [3] Prof. S.P.Kosbatwar, Prof.S.K.Pathan, "Pattern Association for character recognition by Back-Propagation algorithm using Neural Network approach" International Journal of Computer Science & Engineering Survey (IJCES) Vol.3, No.1, February 2012.
- [4] Anita Pal & Dayashankar Singh, "Handwritten English Character Recognition Using Neural Network" International Journal of Computer Science & Communication Vol. 1, No. 2, July-December 2010, pp. 141-144.
- [5] Kauleshwar Prasad, Devvrat C. Nigam, Ashmika Lakhotiya: Character Recognition Using Matlab's Neural Network Toolbox, International Journal of u- and e- Service, Science and Technology Vol. 6, No. 1, February, 2013.
- [6] M. Hanmandlu, K. R. M. Mohan and H. Kumar, "Neural-based Handwritten character recognition", in Proceedings of Fifth IEEE International Conference on Document Analysis and Recognition, ICDAR'99, Bangalore, India, (1999), pp. 241-244.
- [7] T. V. Ashwin and P. S. Sastry, "A font and size-independent OCR system for printed Kannada documents using support vector machines", in Sadhana, vol. 27, Part 1, (2002) February, pp. 35-58.
- [8] S. V. Rajashekararadhya and P. Vanajaranjan, "Efficient zone based feature extraction algorithm for handwritten numeral recognition of four popular south-Indian scripts", Journal of Theoretical and Applied Information Technology, JATIT, vol. 4, no. 12, (2008), pp. 1171-1181.
- [9] R. G. Casey and E. Lecolinet, "A Survey of Methods and Strategies in Character Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 7, (1996) July, pp. 690-706.
- [10] R. C. Gonzalez, R. E. Woods and S. L. Eddins, "Digital Image Processing using MATLAB", Pearson Education, Dorling Kindersley, South Asia, (2004).

- [11] SunitBandaru “Handwritten character recognition using neural network”.
- [12] Jingxin Hong, Wu Li “Image Mosaic Based on SURF Feature Matching” The 1st International Conference on Information Science and Engineering (ICISE2009).
- [13] Ivan Kastelan*, Sandra Kukolj*, VukotaPekovic: “ Extraction of Text on TV Screen using Optical Character Recognition, SISY 2012 2012 IEEE 10th Jubilee International Symposium on Intelligent Systems and Informatics • September 20-22, 2012, Subotica, Serbia.
- [14] A. Rajavelu, M. T. Musavi and M. V. Shirvaikar, “A Neural Network Approach to Character Recognition”, Neural Networks, vol. 2, (1989), pp. 387-393.
- [15] Patel, C. I., Patel, R. and Patel, P. 2011. Handwritten Character Recognition using Neural Network. International Journal of Scientific & Engineering Research, vol. 2 – no. 5, pp. 1-6.
- [16] Anita Pal & Dayashankar Singh, “Handwritten English Character Recognition Using Neural Network” International Journal of Computer Science & Communication Vol. 1, No. 2, July-December 2010, pp. 141-144.
- [17] Sumedha B. Hallale¹, Geeta D. Salunke: offline handwritten digit recognition using neural network, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 2, Issue 9, September 2013
- [18] RAKESH KUMAR MANDAL N R MANNA: Hand Written English Character Recognition using Column-wise Segmentation of Image Matrix (CSIM), WSEAS TRANSACTIONS on COMPUTERS: Issue 5, Volume 11, May 2012.
- [19] Muhammad Naeem Ayyaz, Imran Javed and Waqar Mahmood 1 2: Handwritten Character Recognition Using Multiclass SVM Classification with Hybrid Feature Extraction, Pak. J. Engg. & Appl. Sci. Vol. 10, Jan., 2012 (p. 57-67).
- [20] Bansal, R., Sehgal, P. and Bedi, P. 2010. Effective Morphological Extraction of True Fingerprint Minutiae based on the Hit or Miss Transform. International Journal of Biometrics and Bioinformatics, vol. 4 – no.2, pp. 71-85.
- [21] Gonzalez, R.C., and Woods, R.E., 2004. Digital Image Processing (2nd edition), Pearson Education.
- [22] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z. H., Steinbach, M., Hand, D. J. and Steinberg, D. 2007. Top 10 algorithms in datamining. Knowledge and Information Systems.
- [23] Charles, P. K., Harish, V. Swathi, M. and Deepthi, CH. 2012. A Review on the Various Techniques used for Optical Character Recognition. International Journal of Engineering Research and Applications (IJERA), vol. 2 – no. 1, pp. 659-662.