

## **A Comparative Analysis Of Back Propagation And Random Forest Algorithm For Character Recognition From Handwritten Document**

**Prof. (Dr.) Amit Verma<sup>1</sup>, Gagandeep kaur<sup>2</sup>**

*<sup>1</sup>Professor and Head, CSE Department*

*Chandigarh University Gharuan, (Mohali) Punjab*

*<sup>2</sup>Research scholar, CSE Department*

*Chandigarh University Gharuan, (Mohali) Punjab*

### **Abstract**

Handwritten character recognition is one of the most fascinating and challenging research areas in the field of image processing. In handwriting text there is no constraints on the writing style. Handwritten characters are difficult to recognize due to the different human handwriting styles, variation in angle, size and shape of letters. Various algorithms and approaches have been used for character recognition from digital documents. This paper describes, the recognition and classification of handwritten characters are done by using back propagation neural networks. Recognition rate of the system is evaluated and also compare with random forest algorithm of machine learning. The recognition accuracy of the proposed work is found to be high and satisfactory.<sup>1</sup>

---

This work was supported in part by the CSE Department of Computer science and engineering, Head and faculties.

Prof. (Dr.) Amit verma is professor and head department of Computer Science and Engineering, University institute of Engineering, Chandigarh University Gharuan. (E-mail: [amit.verma@cumail.in](mailto:amit.verma@cumail.in)).

Er. Gagandeep kaur is ME Student, department of Computer Science and Engineering, University institute of Engineering Chandigarh University Gharuan (E-mail: [preetg595@gmail.com](mailto:preetg595@gmail.com)).

**Keywords** — Character recognition, Handwritten Document, Neural networks, canny edge detection, 2D-gabor filter, Random forest algorithm, back propagation algorithm.

### 1. INTRODUCTION

In the recent years, character recognition from handwritten document has been one of the most important and interesting research areas in field of image processing and pattern recognition. It gives massively to the improvement of an automation process and also improves the interface between machine and human. Generally, handwriting recognition is categorized into two types such as off-line and on-line handwriting recognition systems. In the off-line recognition, the writing is usually captured optically by a scanner and the completed writing is available as an image.

On the other hand, in on-line handwriting recognition system the two dimensional coordinates of successive points are represented as a function of time. The on-line methods have been shown to be superior to their off-line counterparts in recognizing handwritten characters due to the temporal information available with the former. But, in the off-line systems, neural networks have been successfully used to produce comparably high recognition accuracy levels, Recognition of handwritten character is a problem since there is a variation in same character due to different types of noises or font size. Character recognition is one of the mostly used for verification of somebody as well as text.

### 3. RELATED WORK

Random forest algorithm is also used for character recognition is discussed in related work. Random Forest is used for UCI machine learning repository namely Heart-h, Sonar, Heart-c and Colic etc. But never used for handwritten character databases. Random Forest is a forest of unpruned trees and each tree s built using random sampling of training data which named as bootstrap. Then from these random samples some random features are selected and Gini index is computed for deciding best split. It uses the multiple random trees classification for the given set of inputs to vote on an overall classification. The drawback of Random forest algorithm is to over fit for some datasets with noisy classification/regression tasks. For data including categorical variables with different number of levels, random forests are biased in favor of those attributes with more levels. Therefore, the variable importance scores from random forest are not reliable for this type of data.

**Random forest algorithm:**

Algo_Random forest
<p>BEGIN</p> <p>Calculate the prediction error for all trees</p> <p>Choose training data subset</p> <p>Then</p> <p>Stop condition holds at each node</p> <p>If</p> <p>(condition true)</p> <p>Repeat the process for other nodes</p> <p>Else</p> <p>(condition false)</p> <p>build the next split</p> <p>Choose the variable from subset</p> <p>Compute Gini index at each split point for chosen variable.</p> <p>choose the best split</p> <p>END</p>

**4. PROPOSED ALGORITHM**

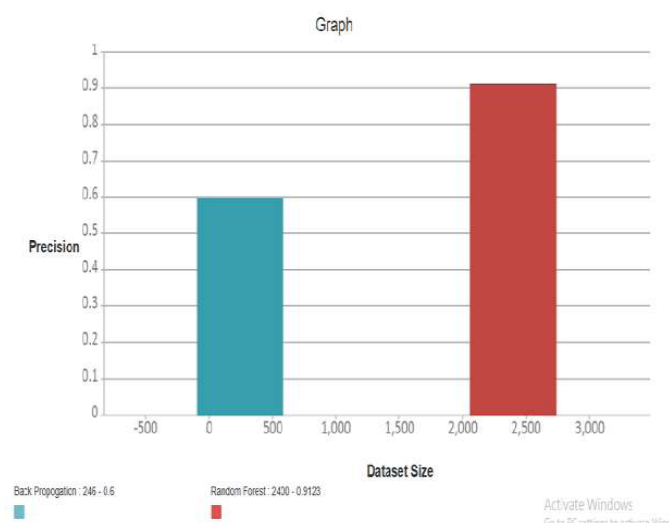
The back propagation is a proposed algorithm to recognize the handwritten characters.

**Back propagation algorithm:**

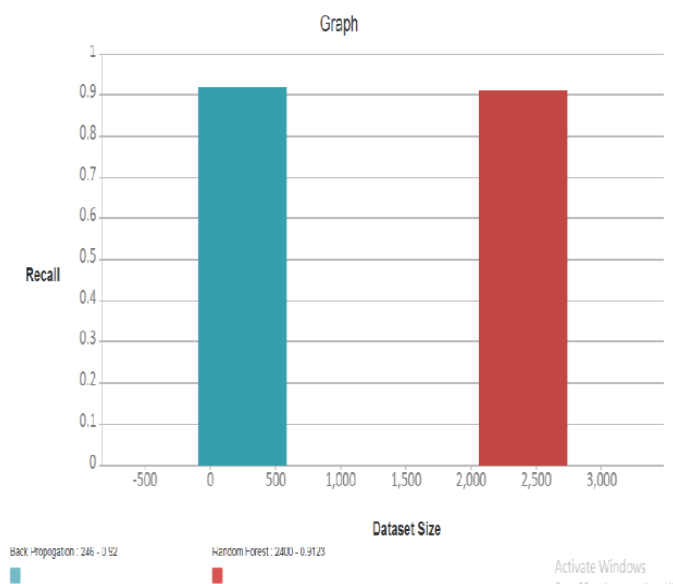
Algo_Back propagation
<p>BEGIN</p> <p>Input: ProblemSize, inputPatterns, Iterationsmax, learn rate</p> <p>Output: Network</p> <p>Network <math>\leftarrow</math> constructNetworkLayer( )</p> <p>Networkweight <math>\leftarrow</math> initializeWeights(network, ProblemSize)</p> <p>For (i=1 to iterationsMax)</p> <p>Pattern<sub>i</sub> = selectInputPattern( InputPatterns)</p> <p>Output<sub>i</sub> = ForwardPattern( Pattern<sub>i</sub>, Network)</p> <p>BackPropagateError(Pattern<sub>i</sub>, Output<sub>i</sub>, Network)</p> <p>UpdateWeights(Pattern<sub>i</sub>, Output<sub>i</sub>, Network, Learnrate)</p> <p>End</p> <p>Return(Network)</p>

## 5. EXPERIMENTAL RESULTS

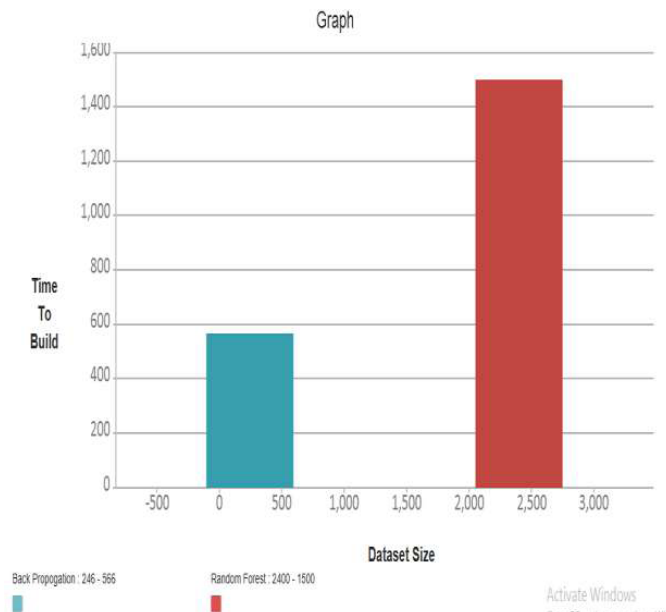
The experiment was carried out on 246 handwritten characters. Different users were asked to write the characters in their own handwriting for each sample. It was observed that every character was different from one another in size, style and shape even if it was written by the same user. From each character, region based features were extracted. The feed forward back propagation neural network algorithm used for recognition and classification and also calculate the precision, recall and time to build factors. It was found that the recognition rate of back propagation & Random Forest is 98.83% and 96% respectively.



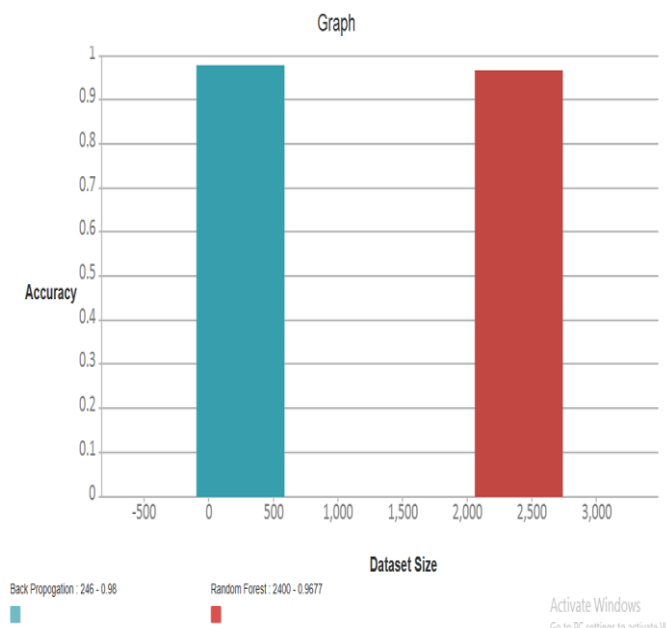
**Fig.2 graph of precision factor**



**Fig.3 graph of recall factor**



**Fig.4 graph of time to build factor**



**Fig.1 comparison between back propagation & random forest algorithm**

## 6. CONCLUSION

In this paper, for the recognition & classification feed forward back propagation neural network Algorithm used and the recognition accuracy of 98.83% was achieved by considering 245 handwritten characters. This accuracy can probably be increased by

taking into account a large data set for the classification and also compare with random forest algorithm of machine learning. The back propagation algorithm is better than random forest algorithm for handwritten character recognition.

## **7. REFERENCES**

- [1] Ramandeep Kaur, Shruti Gujral "Recognition of Similar Shaped Isolated Handwritten Gurumukhi Characters Using Machine Learning," 5th International Conference- Confluence the Next Generation Information Technology Summit (Confluence), Page no.251-256, 2014 IEEE.
- [2] Amrita Hirwani, Neelmani Verma, Sandeep Gonnade "Efficient Handwritten Alphabet Recognition Using LBP based Feature Extraction and Nearest Neighbor Classifier," International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 4, Issue 11, pp.549-553, November 2014.
- [3] Magesh Kasthuri, V.Shanthi Research Scholar "Noise Reduction and Pre-processing techniques in Handwritten Character Recognition using Neural Networks," International Journal of Computing Science and Communication Technologies, VOL.6 NO. 2, ISSN 0974-3375, January 2014
- [4] D.kavitha, P.Shamini "Handwritten Document into Digitized Text Using Segmentation Algorithm," Department of Computer Applications, Easwari Engineering College, Chennai, Tamil nadu, Special Issue, 4th National Conference on Advanced Computing, Applications & Technologies, May 2014.
- [5] Magesh Kasthuri, Dr. V.Shanthi "Pre-processing and Self training techniques in Handwritten Character Recognition," Indian journal of applied research, Volume: 4, Issue: 4, ISSN - 2249-555X, Apr 2014.
- [6] Reetika Verma, Mrs.Ruinder Kaur "Enhanced Character Recognition Using Surf Feature and Neural Network Technique," International Journal of Computer Science and Information Technologies, Vol. 5 (4), ISSN. 0975-9646, pp. 5565-5570, 2014.
- [7] Faisal Mohammad, Jyoti Anarase, Milan Shingote, Pratik Ghanwat "Optical Character Recognition Implementation Using Pattern Matching," International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 5 (2), pp. 2088-2090, ISSN. 0975-9646, 2014.
- [8] Amit Choudhary "A Review of Various Character Segmentation Techniques for Cursive Handwritten Words Recognition," International Journal of Information & Computation Technology, ISSN 0974-2239, Volume 4, Number 6, pp. 559-564, 2014.
- [9] Er. Neetu Bhatia "Optical Character Recognition Techniques: A Review," International Journal of Advanced Research in Computer Science and

- Software Engineering, ISSN: 2277 128X, Volume 4, Issue 5, pp. 1219-1223 May 2014.
- [10] Nisha Sharma, Tushar Patnaik, Bhupendra Kumar “Recognition for Handwritten English Letters,” International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 7, January 2013.
- [11] Rohini B. Kharate, Dr.S.M.Jagade, Sushilkumar N. Holambe “A Brief Review and Survey of Segmentation for Character Recognition,” International Journal of Engineering Sciences, ISSN: 2306-6474, Pages: 14-17, January 2013.
- [12] Sandeep Saha, Nabarag Paul, Sayam Kumar Das, Sandip Kundu, “Optical Character Recognition using 40-point Feature Extraction and Artificial Neural Network,” International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 4, ISSN: 2277 128X, April 2013.
- [13] Parikh Nirav Tushar, Dr. Saurabh Upadhyay “Chain Code Based Handwritten Cursive Character Recognition System with Better Segmentation Using Neural Network,” International Journal of Computational Engineering Research, Vol, 03, Issue, 5, May 2013.
- [14] Gaurav Kumar, Pradeep Kumar Bhatia, Indu “Analytical Review of Preprocessing Techniques for Offline Handwritten Character Recognition,” International Journal of Advances in Engineering Sciences Vol.3 (3), ISSN: 2231-0347 Print-ISSN: 2231-2013, July, 2013.
- [15] Munish Kumar, M. K. Jindal, and R. K. Sharma “MDP Feature Extraction Technique for Offline Handwritten Gurmukhi Character Recognition,” Smart Computing Review, vol. 3, no. 6, December 2013.
- [16] Chirag Patel, Atul Patel, and Dipti Shah “A Review of Character Segmentation Methods,” International Journal of Current Engineering and Technology, ISSN 2277 – 4106, Vol.3, No.5, (December 2013).
- [17] Vipin, Rajeshwar Dass, Rajni “Character Recognition using Neural Network,” International Journal of Advanced Trends in Computer Science and Engineering, Volume 2, No.3, ISSN No. 2278-3091, pp. 62-67, May - June 2013.
- [18] Ankit Sharma, Dipti R Chaudhary “Character Recognition Using Neural Network,” International Journal of Engineering Trends and Technology (IJETT), Volume4, Issue4, and ISSN: 2231-5381, pp. 662-667, April 2013.
- [19] Nisha Vasudeva, Hem Jyotsana Parashar, Singh Vijendra “Offline Character Recognition System Using Artificial Neural Network,” International Journal of Machine Learning and Computing, Vol. 2, No. 4, August 2012.
- [20] Rajbala Tokas, Aruna Bhadu, “A COMPARATIVE ANALYSIS OF FEATURE EXTRACTION TECHNIQUES FOR HANDWRITTEN CHARACTER RECOGNITION,” International Journal of Advanced Technology & Engineering Research (IJATER) ISSN No: 2250-3536 Volume 2, Issue 4, July 2012.
- [21] Rajiv Kumar Nath, Mayuri Rastogi “Improving Various Off-line Techniques used for Handwritten Character Recognition: a Review,” International Journal

of Computer Applications (0975 – 8887) Volume 49– No.18, pp. 11-17, July 2012.