Volume 2, Issue 4, April 2017 ISSN (online): 2456-0006

International Journal of Science Technology Management and Research

Available online at: www.ijstmr.com

Online Detection of Multi-Language Handwritten Document

Neeta R. Panagale
Dept. of Computer Engg.
Shatabdi Institute of Engineering & Research
MH, INDIA

Sonali A. Boraste
Dept.of Computer Engg.
Shatabdi Instituteof Engineering & Research
MH, INDIA

Anuradha K. Chaudhari
Dept.of Computer Engg.
Shatabdi Institute of Engineering & Research
MH, INDIA

Mahesh T. Dhande
Dept. of Computer Engg.
Shatabdi Institute of Engineering & Research
MH, INDIA

***Abstract*:** There are numerous things we people have in similar. Be that as it may, there are different things that are exceptionally interesting of each person - DNA, fingerprints, and so on. Handwriting is one other such thing that is one of a kind to each person, which the current reviews on handwriting analysis have as of now proved to be a good recognition system. In our proposed system, we have implemented to recognize the writer by giving input as their handwriting. This technique proved to be more accurate by implementing neural network. The intention of using neural network is to solve problems in the same way that the human brain would. Neural network has proved to be one of the best approaches in feature extraction for character recognition which shows high accuracy results when compared to other techniques. Experimental results proved that our proposed system has shown better results in terms of accuracy and speed when compared with other techniques.

Keywords: ***Handwriting recognition, neural network, feature extraction, character recognition.***

I. INTRODUCTION

Correct and efficient recognition of handwritten text of many languages is a challenging problem due to the cursive nature of the script. Also for millions of manuscripts and loose pages, the writer of the manuscripts may be in doubt or completely unknown. Optical character recognition is the past when in 1929 Gustav Tauschek got a patent on OCR in Germany took after by Handel who got a US Patent on OCR in USA in 1933. From that point number of character acknowledgment frameworks have been created also, are being used for even business purposes too. Yet at the same time there is a want to assemble some more smart manually written character acknowledgment framework since hand composing vary from one individual to other. His written work style, state of letters in order and their sizes has the effect and multifaceted nature to perceive the characters. We will be developing a system which will examine document and generate an algorithmic analysis of human handwritten document by comparing different samples of handwriting through image processing algorithms, this was stated by I. Guyon, L. Schomaker in 1994[11].

Handwriting recognition is without a doubt one of the most difficult regions of pattern recognition. It is to a great degree helpful in an extensive range of real world issues, documentation checking, bank cheque handling, signature confirmation, and numerous others. A few example recognition approaches have been connected to both on- line and off - line handwriting recognition, counting measurable techniques, auxiliary and syntactic strategies as well as neural systems. Some perusing frameworks distinguish strokes; others attempt to recognize characters, bunches of characters, or whole words. Neural systems are made out of straightforward components working in parallel. These components are enlivened by natural sensory systems. As in nature, the system work is resolved to a great extent by the associations between components. We can prepare a neural system to play out a specific capacity by changing the values of the

associations (weights) between components. Ordinarily neural systems are balanced, or prepared, so that specific information prompts a particular target yield. There, the set is balanced, in view of an examination of the yield and the objective, until the system yield matches the objective.

II. RELATED WORK

In general, handwriting recognition is classified into two types as offline and on-line handwriting recognition methods. In the offline recognition, the writing is usually captured optically by a scanner and the completed writing is available as an image. But in the on-line system the two dimensional coordinates of successive points are represented as a function of time and the order of strokes made by the writer are also available. The on-line methods have been shown to be superior to their offline counterparts in recognizing handwritten characters due to the temporal information available with the former R. Plamondon and S. N. Srihari in 2000[1] [2].As a result, the offline handwriting recognition continues to be an active area forresearch towards exploring the newer techniques that would improve recognition Classification techniques have been applied to handwritten character recognition since the 1990s. These methods include statistical methods based on Bayes decision rule, ANNs [7]), Kernel Methods including Support Vector Machines (SVM) and multiple classifier combination this was stated by C. L. Liu, H. Fujisawa. [4], [5]. U. Pal et al,have proposed a modified quadratic classifier based multiple classifier combination. This was stated by C. L. Liu, H. Fujisawa. In 2005[4], [5]. U. Pal et al, have proposed modified quadratic classifier based scheme to recognize the offline handwritten numerals of six popular Indian scripts [3].Multilayer perception has been used for recognizing Handwritten English characters. This was stated by Anita Pal Dayashankar Singh in 2010[6].The features are extracted from Boundary tracing and their Fourier Descriptors. Character is identified by analyzing its shape and comparing its features that distinguish each character.

III. EXISTING SYSTEM

In existing system, there is a possibility of recognizing handwriting which is written in a single language [6]. Google Translate is not perfect, especially with the morphologically rich languages, like Finnish and Hungarian. Use to recognize character E.g. (arrows, equation signs, star, heart shapes etc.) Googles online handwriting recognition system that currently supports 22 scripts and 97 languages [5]. We are implementing neural network algorithm for proposed methodology, which shows higher efficiency as compared to the existing system (i.e. Google translator)

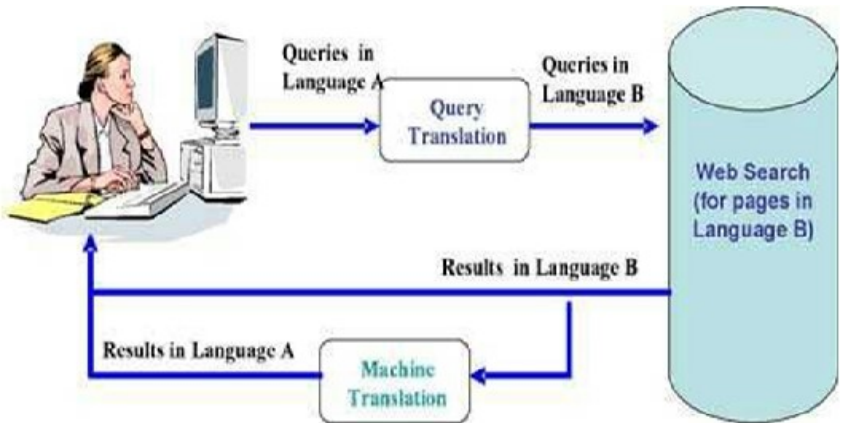

Figure 1: Existing System

IV. PRAPOSED SYSTEM

In our multi-language handwriting recognition system, we have proposed two modules-Training and Testing. In the training module, first of all an image will be uploaded for training. Image processing will be carried out and its features will be extracted. The patterns of the particular handwriting will be identified and stored in a le. Next, for testing our system we will upload another image in which its and writing has to be identified. Again by performing some image pre-processing steps, we will be extracting its features. These extracted features will be compared with the data stored in training set and if it matches, it shows the writer of the handwritten text else not.

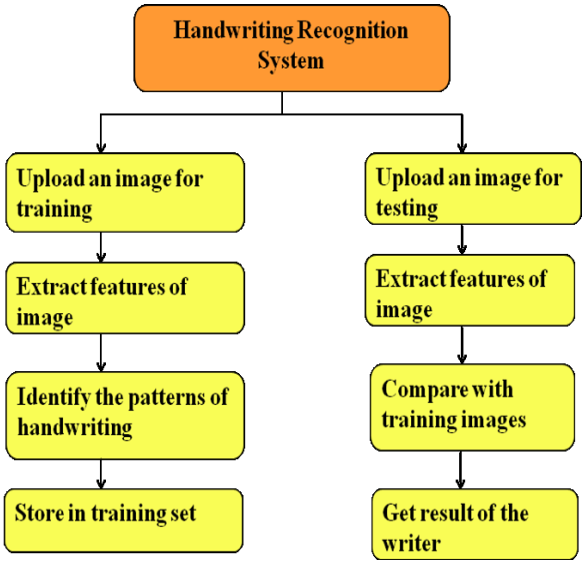

Figure 2: Block Diagram Of Propose System

V. PROPOSED WORK FLOW

In our handwriting recognition system, there are two modules for testing and training. In training module, the user uploads an image which had the handwriting of the person. The features of that image are extracted. The patterns of that particular image are identified and stored in the training set. In the testing module, again the image is uploaded for testing. Similar to training module, here feature extraction is done of the image. It is compared with the training images which are stored in the database and then finally the result is achieved.Our proposed system works on neural network because it gives result better than other algorithms. For the last ten years neural networks have attracted a great deal of attention. They offer an alternative approach to computing and understand the human brain. This approach is not something new. The first artificial neuron was produced in 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pits. At the time sixties, for reasons that are out of the scope of this article, people turned away from neural networks and understood in the symbolic side of Artificial Intelligence. Only in the eighties scientists saw the real potential of neural networks. Neural networks take a different approach to problem solving than that of conventional systems. Conventional systems use an algorithmic approach i.e. the system follows a set of instructions in order to solve a problem. Unless the specific steps that the system needs to follow are known the system cannot solve the problem. That restricts the problem solving capability of conventional systems to problems that we already understand and know how to solve [13]. But systems would be so much more useful if they could do things that we don't actualy know how to do. Neural networks working similar way the human brain does. The network is composed of a huge number of highly interconnected processing elements (neurons) working in parallel to solve a particular problem. They cannot be programmed to perform a particular task. The examples must be selected carefully otherwise useful time is wasted or even worse the network might be functioning not correctly. The problem is that there is no way of knowing if the system is faulty or not, when an error occurs. The building block of a neural net is the neuron. An artificial neuron works much similar way the biological one does. It takes many inputs having different weightings and has one output which depends on the inputs. A biological neuron can either 'fire' or not 'fire'. In an artificial neuron 'firing' is normally specified by a logical one and nor 'firing' by a logical zero.

VI. IMPLEMENTATION DETAILS

- **Working of neural network**

A huge ANN might have hundreds or thousands of processor units, whereas a human brain has billions of neurons with a corresponding increase in magnitude of their complete interaction and emergent behaviour. ANN researchers are generally not concerned with whether their networks accurately resemble biological systems. For example, researchers have correctly simulated the function of retina and modelled the eye rather well. Although the mathematics involved with neural networking, a user easily gain at least an

operational understanding of their structure and function. Neural networks are organized in layers [12]. Layers are made up of a number of interconnected 'nodes' which contain an 'activation function'.

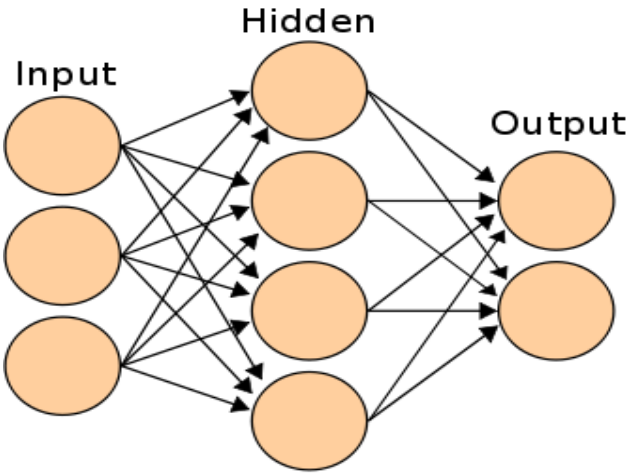

Figure 3: Working of neural network

Patterns are presented to the network by the 'input layer', which communicates to one or more 'hidden layers' where the exact processing is done via a system of weighted 'connections'. The hidden layers then link to an 'output layer' where the result is output. Most ANNs contain some form of 'learning rule' although there are many different kinds of learning rules used by neural networks, this experiment is concerned only with one; the delta rule. The delta rule is often utilized via the most common class of ANNs called 'back propagation neural networks' (BPNNs) [11]. Back propagation is an abbreviation for the backwards propagation of fault. When a neural network is presented with a pattern it makes a random 'guess' as to what it might be. It then sees how far its answer was from the exact one and makes an appropriate adjustment to its connection weights.

Neural network Algorithm steps

1. Take a random training sample for any character (it is better to go ahead in sequence i.e., 0-9 and A-Z) and then generate 12×8 matrix.
2. Generate the corresponding 12×8 initial weight_matrix by taking 3 for a 1 and -3 for 0 of the input matrix.
3. Calculate the weighted sum Oi(net activation) as follows
4. Oi= $\sum_{j=1}^{96} wjixj$

   Where, i=0, 1, 2……..35.
5. Calculate the sum of positive weight, Pwi of the weight_matrix.
6. Calculate Yi=f (Oi) =Oi/Pwi.

   Where,

   Oi →Net activation for each character i (e.g. 0, 1, 2 …9, A, B……Z)

   Pwi→ the sum of positive weight of the weight_matrix for the each character.
7. Pick the maximum Yi.
8. Check if the corresponding neuron fires. If neuron fires then save the weight matrix into a file.
9. Check with other training samples of the same character and update the weight described as above.
10. Fixed the weight matrix and save into a file after the final training samples for that character.
11. Repeat steps 1 to 9 for any other character.
12. If the training is complete then the network is established.

VII. INTERMEDIATE RESULT

Input Image

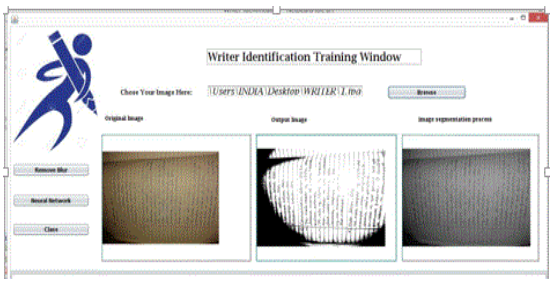

Figure 4: Writer identification training window

In this window we browse the image for train the system . and also perform the different operation on that image .

Output

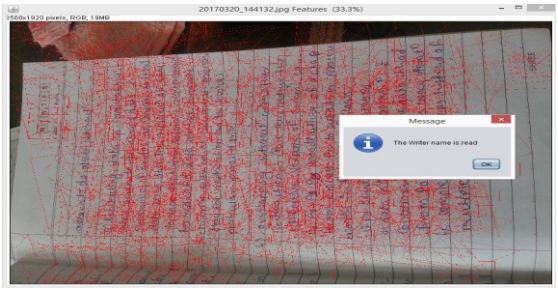

Figure 5: features extracted image with writer

This window displays the result of features extracted image with its identified writer name.

Image Name	Size	Time Taken for Training in ms	Features Extracted	Time Taken for Testing in ms
1.png	19 MB	6234	506	5023
2.png	19 MB	4750	459	4868
3.png	19 MB	4776	441	4950
4.png	19 MB	4359	279	4344
5.png	19 MB	4407	369	4343
6.png	19 MB	5735	494	4407

Table 1: Time taken for training and testing

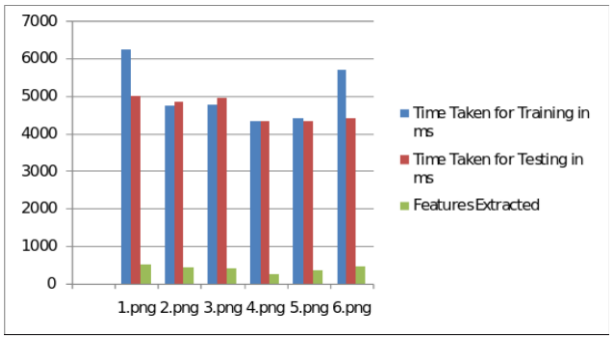

Figure 6: Training and testing time graph

CONCLUSION AND FUTURE WORK

Our work speaks to the distinctive system are accessible for recognizing the manually written scripts. This additionally concentrates on that in today's world handwriting recognition is extremely troublesome yet important. There are numerous applications where we require hand recognition systems like bank cheque, postal addresses, and form reports. In every one of the procedures fundamental stage is feature extraction. This paper speaks to the comparison between all strategies. Neural networks may have any number of outputs when compared to other techniques like SVM etc. Moreover, neural networks provide us the first step towards AI by generating a model based on how our own human body learns. It is a nonlinear model that is easy to use and understand compared to statistical methods and is non-parametric model while most of statistical methods are parametric model that need higher background of statistic.

In future we are expecting to explore the Back-propagation algorithm to make recognition of digits more fast and efficient and improve the overall performance.

REFERENCES

1. R. Plamondon and S. N. Srihari, On-line and offline handwritten character recognition: A comprehensive survey, IEEE. Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, 2000.
2. N. Arica and F. Yarman-Vural, an Overview of Character Recognition Focused on Offline Handwriting, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2001, 31(2)
3. U. Pal, N. Sharma, T. Wakabayashi and F. Kimura, "Handwritten character recognition of popular south Indian scripts", Proceeding SACH'06 Proceedings of the 2006 conference on Arabic and Chinese handwriting recognition Springer-Verlag Berlin, Heidelberg
4. U. Pal, T.Wakabayashi and F. Kimura, "Handwritten numeral recognition of six popular scripts", Ninth International conference on Document Analysis and Recognition ICDAR 07, Vol.2, 2007.

5. Pooja Yadav, Nidhika Yadav, "Handwriting Recognition System - A Review", International Journal of Computer Applications, March 2015
6. P. Shankar Rao, J. Aditya, "Handwriting Recognition – "Offline" Approach", IEEE
7. J. Pradeep, E. Srinivasan and S. Himavathi, "Diagonal Based Feature Extraction for Handwritten Alphabets Recognition System Using Neural Network", International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb 2011
8. C. L. Liu, H. Fujisawa, Classification and Learning for Character Recognition: Comparison of Methods and Remaining Problems, Int. Workshop on Neural Networks and Learning in Document Analysis and Recognition, Seoul, 2005.
9. F. Bortolozzi, A. S. Brito, Luiz S. Oliveira and M. Morita, Recent Advances in Handwritten Recognition, Document Analysis, Umapada Pal, Swapan K. Parui,Bidyut B. Chaudhuri,.
10. Anita Pal Dayashankar Singh, Handwritten English Character Recognition Using Neural, Network International Journal of Computer Science Communication.vol. 1, No. 2, July-December 2010.
11. Alex Graves, Jurgen Schmidhuber, "Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks", IEEE,
12. Jonathan J. Hull, Alan Commike and Tin-Kam Ho, "Multiple Algorithms for Handwritten Character Recognition", Int. Workshop on Frontiers in Handwriting Recognition, Montreal, Canada, April 2-3.
13. E. Kavallieratou, N. Fakotakis, G. Kokkinakis, "An unconstrained handwriting recognition system", IJDAR, 2002.
14. I.Guyon,L.Schomaker,R.Plamondon,M.Liberman,andS.Janet, UNIPEN project of on-line data exchange and recognizer bench- marks, in ICPR, 1994.

AUTHOR PROFILE

Sonali A. Boraste
Dept.of Computer Engg.
Shatabdi Instituteof Engineering
& Research MH, INDIA

Neeta R. Panagale
Dept.of Computer Engg.
Shatabdi Instituteof Engineering
& Research MH, INDIA

Anuradha K. Chaudhari
Dept.of Computer Engg.
Shatabdi Instituteof Engineering
& Research MH, INDIA