

# Neural Network Based Integrated Approach Towards Recognition Of English Handwritten Alphanumeric Characters.

Sneha U. Bohra <sup>1</sup>, Dr. P V. Ingole <sup>2</sup>

*M. E Scholar* <sup>1</sup>, *Guide* <sup>2</sup>

*G.H.Raisoni College of Engineering & Management, Amravati*

## Abstract

*Handwritten Character recognition is one of the active and challenging areas of research in the intense field of Image Processing and Pattern recognition. It has large number of applications which includes, processing of bank cheques, reading aid for blind, conversion of any hand written document into digital form and many more. The main reason behind the activeness of the field are the challenges like the variation and distortion of handwritten characters, different style of writing and also the variety of scripts available across the entire world.*

*This paper proposes system which implements integrated methodology to recognise handwritten English characters. This integration of methods includes Template Matching and Block Segmentation Technique.*

*Keywords: - Handwritten Character recognition, Image Processing, Pattern recognition, Neural Network*

## 1. Introduction

Handwritten Character Recognition has gained immense popularity as it has wide application areas which reduce the task of data entry and save the time in case of Form filling, Postal Automation, and Banking etc. Development of a system that recognizes handwritten character is a challenging job due to the varying shape and size of the character, different writing styles of the character and many other factors. Researchers are now focusing on new techniques and methods that would reduce the processing time and at the same time providing higher recognition accuracy. Character recognition plays vital role where there is large quantity of text based information. This area

is mainly the electronic translation of handwritten, typewritten or printed text into machine editable form. Character recognition is also referred as optical character recognition (OCR). From the given image of characters as input, OCR deals in recognizing the character. In general the character recognition is basically classified into two types: Offline handwritten text recognition, online handwritten text recognition.

The process of Character Recognition is depicted in the figure 1. These phases can be applied to scanned document for text recognition.



Figure 1: - Process of Character Recognition

### 1. Pre-processing Phase

This phase is applied on the image so as to acquire the image suitable for the segmentation phase. It involves various steps like removal of noise from the image, resizing the image, binarization of the image or removal of slant angle from the document etc. Various type of filtering methods like Gaussian filtering method, Min-max filtering method etc are applied for noise removal. The process of binarization is the conversion of colored image into its black and white equivalent image. In order to speed up the process of recognition, the images are often resized.

### 2. Segmentation Phase

This phase is considered to be the most important phase as the final output of the entire process depends upon the output of this phase. This step decomposes the

image in hierarchical way. It first segments the image into rows by applying row histogram, and then identifies words by applying column histogram and then finally the characters are recognized from the words.

### 3. Feature Extraction Phase

This step is the key phase for any pattern recognition application. This step intends in extracting the important features of the objects which can be used to distinguish it from the other objects. The aim of this phase is to minimize the within class variation and maximize the between class variation so that the object can be uniquely recognized. Various Feature extraction techniques like Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), Independent Component Analysis (ICA), Chain Code (CC), Scale Invariant Feature Extraction (SIFT), zoning, Gradient based features, Histogram might be applied to extract the features of individual characters. The feature vector so obtained as an output of this phase is used to train the system.

### 4. Classification

The feature vector is given as an input to this important phase. This step classifies the vector on the basis of various classifiers. Most commonly used trained classifiers are Artificial Neural Network or Support Vector Machine. The classifiers compare the given vector with the stored pattern and give the best match as an output.

## 2. Literature Review

In 2013, Sandeep Saha and et.al presented a system for English handwriting recognition based on 40-point feature extraction of the character. Character recognition system used multilayer feed forward neural network. This method firstly extracts features using 40-point feature extraction and then the data is used to train the artificial neural network. In the end, they tested the artificial neural network and concluded that this method had a good performance at handwritten character recognition [1]. In another paper from the year 2013, authors provided review of the major approaches used in the field of handwritten character recognition during the last decade. Different pre-processing, segmentation, feature extraction, classification techniques were also discussed. They concluded

that the existing OCR for handwritten has very low accuracy and there is need an efficient solution to solve this problem so that overall performance can be increased [2]. A paper from 2013, presented a new approach, named, Block wise Segmentation Technique (BST) to prepare the input character matrix to train a neural network model. BST is the representation of input in terms of a number of equal sized vectors derived from the input matrix. On testing, the model produced a high level of accuracy to identify deviated characters. Although the proposed BST technique is not a complete one, it is found quite satisfactory dealing with different writing styles. They concluded that BST may be an important element for character recognition system when combined with some other techniques of character recognition [3]. Om Prakash Sharma and et.al in their 2013 paper presented recent trends and tools used for feature extraction that helps in efficient classification of the handwritten alphabets. Numerous models of feature extraction have been defined by different researchers in their respective dissertation. It was found that the use of Euler Number in addition to zoning increases the speed and the accuracy of the classifier as it reduces the search space by dividing the character set into three groups [4].

The work presented by Ketan S. Machhale and et.al focused on development of an Offline Hand Written English Character Recognition algorithm based on Artificial Neural Network (ANN). The ANN implemented in this work had single output neuron which showed whether the tested character belongs to a particular cluster or not. The implementation was completely in C language. Ten sets of English alphabets (small-26, capital-26) were used to train the ANN and 5 sets of English alphabets were used to test the network. The characters were collected from different persons over duration of about 25 days. The algorithm was tested with 5 capital letters and 5 small letter sets. The result showed that the algorithm recognized English alphabet patterns with maximum accuracy of 92.59% and False Rejection Rate (FRR) of 0% [5].

In 2012, a paper introduced a new kind of algorithm, based on that chain code along with another significant characteristic, number of vertices (end or junction point) and edges (chain code, from one vertex to another) existing in each alphabet has been used to identify English Capital Characters. They have built a technique that is independent of size and many styles of input characters. Experimental results showed relatively high accuracy [6]. Debi Prasad Bhattacharya in his paper tackled the problem of handwritten character recognition with multiresolution technique using Discrete wavelet transform (DWT) and Euclidean

distance metric (EDM). The technique was tested and found to be more accurate and faster. Characters are classified into 26 pattern classes based on appropriate properties. Features of the handwritten character images are extracted by DWT used with appropriate level of multiresolution technique, and then each pattern class is characterized by a mean vector. Distances from input pattern vector to all the mean vectors were computed by EDM. Minimum distance determines the class membership of input pattern vector. The proposed method provided good recognition accuracy of 90% for handwritten characters even with fewer samples [7].

The paper from the year 2012 identified the most suitable NN for the design of hand written English character recognition system. Different Neural Network (NN) topologies namely, back propagation neural network, nearest neighbour network and radial basis function network were built to classify the characters. All the NN based Recognition systems used the same training data set and are trained for the same target mean square error. Two hundred different character data sets for each of the 26 English characters were used to train the networks. The performances of the recognition systems were found to be satisfactory [8]. The 2012 paper demonstrated the nature of handwritten characters, conversion of handwritten data into electronic data, and the neural network approach to make machine capable of recognizing hand written characters [9].

### 3. Analysis of Problem

In this era of modernization where we carry our entire world in our pockets, it sometimes becomes very difficult to maintain the old data which is available in the form of hard copy. Therefore, there is a need to propose such methodology that aims at converting the hard data into digital (soft) form. Various ways of doing this task have been introduced and Character Recognition is one of them.

### 4. Objective and Proposed Work

In the development and implementation of this project, the aim is to be built such a system that will determines the speed, accuracy, ease of use, and user satisfaction in maintaining and archiving of data. The proposed system would make use of scanned image of a Single Page Feedback form, process it by applying various techniques to extract isolated characters and then recognize the handwritten English character and numerals using neural network.

The main objective in the development of the system is to provide efficient output within reasonable computation time. To achieve these objectives, instead of the implementation of any traditional method of character recognition, we aim to integrate different methodologies for recognition. The proposed system will make use of Template Matching along with Block Segmentation Technique for character matching purpose and the obtained result will be given to the Neural Network for further classification. This integration of these two methods for finding the match will discard the irrelevant match obtained during the mid steps of the process and thus pushing forward the obtained result for recognition process.

The following figure depicts the major working of the proposed system.

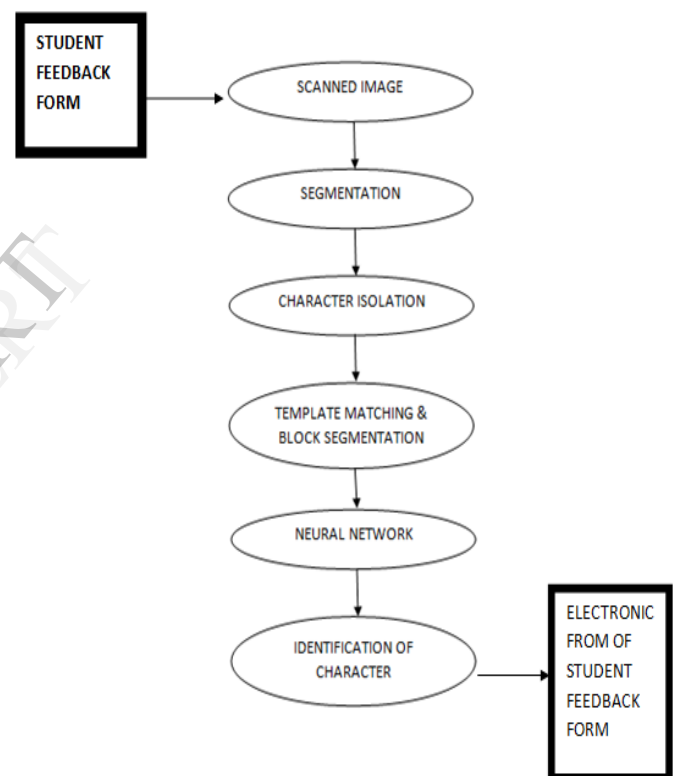


Figure 2: - Flow of Proposed System

The first Step is to obtain the scanned image of the document. Before forwarding the image towards the actual identification process, the system transforms the image into binary and also performs some pre-processing task on the image. Then, the segmentation of the lines, words and characters will take place. When the isolated character will be obtained, the obtained character will then be fed to the integrated module proposed for the recognition. This result of this module will then be fed to the Neural Network for further recognition.

## 5. Conclusion and Future Scope

Our effort is to obtain efficiency both in terms of speed and time. A worth analysis of designing principle has been followed in the development of this project. In future some new and enhanced features will be added to the above mentioned system .For example, it can enhanced to process various forms and process them.

*Devanagari Character Recognition”, 12th International Conference on Frontiers in Handwriting Recognition, 2010.*

## 6. References

- [1] Sandeep Saha, Sayam Kumar Das Nabarag Paul, 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, April 2013.*
- [2] 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.*
- [3] Apash Roy, N. R. Manna,”*Handwritten Character Recognition Using Block wise Segmentation Technique (BST) in Neural Network*”, *International Conference on Intelligent Infrastructure January 2013.*
- [4] Om Prakash Sharma, M. K. Ghose, Krishna Bikram Shah, Benoy Thakur,”*Recent Trends and Tools for Feature Extraction in OCR Technology*”, *International Journal of Soft Computing and Engineering (IJSCE), Volume-2, Issue-6, January 2013.*
- [5]Ketan S.Machhale, Pradnya P. Zode, Pravin P. Zode, ”*Implementation of Number Recognition using Adaptive Template Matching and Feature Extraction Method*”, *International Conference on Communication Systems and Network Technologies (CSNT), 2012.*
- [6] Tirtharaj Dash Tanistha Nayak,”*English Character Recognition using Artificial Neural Network*”, *Proceedings of National Conference on AIRES-2012, Andhra University.*
- [7] Debi Prasad Bhattacharya & Susmita Koner,”*ENGLISH Alphabet Recognition Using Chain Code And LCS*”, *Indian Journal of Computer Science and Engineering (IJCSE), Volume 3, Issue No 2, Apr-May 2012.*
- [8] Dileep Kumar Patel, Tanmoy Som, Sushil Kumar Yadav, Manoj Kumar Singh, “*Handwritten Character Recognition Using Multiresolution Technique and Euclidean Distance Metric*”, *Journal of Signal and Information Processing, March 2012, 208-214.*
- [9] J.Pradeepa, E. Srinivasana, S. Himavathi, “*Neural Network Based Recognition System Integrating Feature Extraction and Classification for English Handwritten*”, *IJE Transactions & Applications, Volume 25, Issue No 2, May 2012.*
- [10] Yusuf Perwej, Ashish Chaturvedi,”*Machine Recognition of Hand Written Characters using Neural Networks*”, *International Journal of Computer Applications, Volume 14, Issue No.2, January 2011, 0975 – 8887.*
- [11]J.Pradeepa, E. Srinivasana, S. Himavathi, ”*Diagonal Based Feature Extraction For Handwritten Character Recognition System Using Neural Network*”, *3<sup>rd</sup> International Conference on Electronics Computer Technology (ICECT), Volume 4, 2011.*
- [12]Sushama Shelke, Shaila Apte,”*A Novel Multi-feature Multi-Classifier Scheme for Unconstrained Handwritten*