

Recognizing Cursive Numberal Handwritten Character Paper

Neetu Bhatia Dept of computer science Kurukshetra Institute of Technology& Management Kurukshetra, India **Er. Darshan Kumar** Assistance Professor in Dept. Computer Science Kurukshetra Institute of Technology& Management Kurukshetra, India

Abstract— Today, Recognizing cursive handwritten character is most problematic, because it requires more trained nets with supervised training. Most of hybrid algorithms are used to recognize the character. By using neural network or other techniques we cannot find correct number or character. To tackle this problem we had tested and found the correct number by using hybrid algorithm BFONN with feature extraction like centroid, bounded box, area To recognize the character, these phases like preprocessing, segmentation and feature extraction are use. This proposed work is present the more accuracy to recognize cursive handwritten character or child's handwriting. This algorithm can be use in global optimizing.

Keywords- optical character recog-nition, feature extraction, bfo with neural network

I. INTRODUCTION

Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic conversion of scanned or photographed images of typewritten or printed text into machine-encoded/computer-readable text. It is widely used as a form of data entry from some sort of original paper data source, whether passport documents, invoices, bank statement, receipts, business card, mail, or any number of printed records The objective of an OCR system is to recognize alphabetic letters, numbers, or other characters, which are in the form of digital images, without any human intervention.

Two different types are in the general term of character recognition:

- On-line character recognition
- Off-line character recognition

On-line character recognition deals with a data stream which comes from a transducer while the user is writing. The typical hardware to collect data is a digitizing tablet which is electromagnetic or pressure sensitive. When the user writes on the tablet, the successive movements of the pen are transformed to a series of electronic signal which is memorized and analyzed by the computer.Off-line character recognition is performed after the writing is finished. The major difference between on-line and off-line character recognition is that on-line character recognition has time-sequence contextual information but off-line data does not.

There are two main areas in character recognition:

- Printed Character Recognition.
- Handwritten Character Recognition.

Printed Character Recognition

The printed texts include all the printed materials such as books, newspapers, magazine documents which are the outputs of typewriters, printers or plotters.

Handwritten Character Recognition

Handwritten character recognition based on the form of written communication can be divided into two categories_ cursive script and hand printed characters. Handwriting is an art of drawing pictures for expressing the human thoughts in a condensed and systematic fashion. As a result there are infinite variations of hand writing. Before developing an effective algorithm for handwritten cursive script, it is necessary to categorize the variations of handwriting for individuals. Figure 1.3 shows an example

This picture will be used for results evaluation also. In this circular shapes have been tried to distort in every row. Like numeric character 5, 8, 9, 4 are changed. Segmentation of these letters also imposes a problem.

Over the past decade, a number of different techniques for recognize character have been developed. Recognize the handwritten cursive characters are visual challenges



Fig ,1.

II. RELATED WORK

[Vijay Laxmi Sahu⁴]This paper presents detailed review in the field of Off-line Handwritten Character Recognition. Various methods are analyzed that have been proposed to realize the core of character recognition in an optical character recognition system. The recognition of handwriting can, however, still is considered an open research problem due to its substantial variation in appearance. Even though, sufficient studies have performed from history to this era, paper describes the techniques for converting textual content from a paper document into machine readable form. Offline handwritten character recognition is a process where the computer understands automatically the image of handwritten script. This material serves as a guide and update for readers working in the Character Recognition area. Selection of a relevant feature extraction method is probably the single most important factor in achieving high recognition performance with much better accuracy in character recognition systems. This paper describe about all phase which are used to recognize character. These phases are : Image aquisition, Pre processing, Segmentation,Feature extraction.

[Dileep Kumar Patel⁷]In the present paper, the problem of handwritten character recognition has been tackled with multiresolution technique using Discrete wavelet transform (DWT) and Euclidean distance metric (EDM). The technique has been tested and found to be more accurate and faster. Characters is 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 tech-nique, and then each pattern class is characterized by a mean vector. Distances from input pattern vector to all the mean vectors are computed by EDM. Minimum distance determines the class membership of input pattern vector. The pro- posed method provides good recognition accuracy of 90% for handwritten characters even with fewer samples.

[Chirag I Patel ¹⁰]This paper is discuss to recognize the characters in a given scanned documents and study the effects of changing the Models of ANN. Today Neural Networks are mostly used for Pattern Recognition task. The paper describes the behaviors of different Models of Neural Network used in OCR. OCR is widespread use of Neural Network. We have considered parameters like number of Hidden Layer, size of Hidden Layer and epochs. We have used Multilayer Feed Forward network with Back propagation. In Preprocessing we have applied some basic algorithms for segmentation of characters, normalizing of characters and De-skewing. We have used different Models of Neural Network and applied the test set on each to find the accuracy of the respective Neural Network.

[Handley, Trevor¹¹]This paper examines the use of Matlab neural networks to accomplish optical character recognition. The NN was built using a dataset of 20000 entries consisting of 16 attributes that define a character. An attempt was made to extract similar features from images of characters. The NN achieved an 80% success rate on the test dataset taken from the 20000 original entries. Applying the NN to manually extracted features resulted in extremely erratic output. Possible reasons for these results are discussed here, as well as ideas for achieving better results.

[Majida Ali Abed⁸]This manuscript considers a new approach to Simplifying Handwritten Characters Recognition based on simulation of the behaviour of schools of fish and flocks of birds, called the Particle Swarm Optimization Approach (PSOA). We present an overview of the proposed approaches to be optimized and tested on a number of handwritten characters in the experiments. Our experimental results demonstrate the higher degree of performance of the proposed approaches. It is noted that the PSOA in general generates an optimized comparison between the input samples and database samples which improves the final recognition rate. Experimental results show that the PSOA is convergent and more accurate in solutions that minimize the error recognition rate.



Fig 2: Plot of error per iteration

[Argha Roy, Diptam Dutta⁹] In this paper, the adaptation of network weights using Particle Swarm Optimization (PSO) was proposed as a mechanism to improve the performance of Artificial Neural Network (ANN) in classification of IRIS dataset. Classification is a machine learning technique used to predict group membership for data instances. To simplify the problem. of classification neural networks are being introduced. This paper focuses on IRIS plant classification using Neural Network. The problem concerns the identification of IRIS plant species on the basis of plant attribute measurements. Classification of IRIS data set would be discovering patterns from examining petal and sepal size of the IRIS plant and how the prediction was made from analyzing the pattern to form the class of IRIS plant. By using this pattern and classification, in future upcoming

years the unknown data can be predicted more precisely. Artificial neural networks have been successfully applied to problems in pattern classification, function approximations, optimization, and associative memories. In this work, Multilayer feed- forward networks are trained using back propagation learning algorithm

In this plot at the Epoch 46 the validation performance returns less Mean square Error. Mean square error is the average square between output & target. The projected result for 54 Epoch we get the test data matrix with the accuracy rate of classified pattern of 97.3%.

III. **PROPOSE WORK**

In our work offline optical character recognition is implemented with the help of neural network. Although neural network is very good in pattern matching yet it lacks in global optimization case. It is stronger in case of local optimization. In computer science, local search is a met heuristic method for solving computationally hard optimization problems. Local search can be used on problems that can be formulated as finding a solution maximizing a criterion among a number of candidate solutions. Local search algorithms move from solution to solution in the space of candidate solutions (the search space) by applying local changes, until a solution deemed optimal is found or a time bound is elapsed. Global optimization is a branch of applied mathematics and numerical analysis that deals with the optimization of a function or a set of functions according to some criteria. Typically, a set of bound and more general constraints is also present, and the decision variables are optimized considering also the constraints.

So if some constraints are subjected to neural network then its searching may terminate prior to solution. BFO is good in global search and weak in local search. So, neural network in association with BFO is used in my work.

A. Algorithm

- 1. Compute histogram and probabilities of each intensity level
- Set up initial $\omega_i(0)_{and} \mu_i(0)$ 2.
- Step through all possible thresholds $t=1\ldots$ maximum intensity 3. 1. Update ω_i and μ_i 2. Compute $\sigma_b^2(t)$
- Desired threshold corresponds to the maximum $\sigma_b^2(t)$ 4.
- You can compute two maxima (and two corresponding thresholds). $\sigma_{b1}^2(t)_{is}$ the greater max and $\sigma_{b2}^2(t)_{is}$ the 5. greater or equal maximum

Desired threshold = 6

 $\frac{\text{threshold}_1 + \text{threshold}_2}{2}$

After generating binary image, white space is to be removed so that selected image is normalised to its corners only. The method is quoted in above flow chart diagram. Below is the way, how we use these coordinates to extract the corners of image in MATLAB.

B. Image features Extraction

Step 1: Compute the input image centroid

Step 2: Divide the input image into 25 equal zones.

Step 3: Compute the distance between the Image centroid to the pixel present in the zone column.

Step 4: Repeat step 3 for the entire pixel present in the zone column.

Step 5: Com pute the average pixel distance in the zone column. (One feature)

Step 6: Repeat steps 3 to 5 for the entire zone column present in the zone (Ten features).

Step 7: Compute the distance between the image centroid to the pixel present in the zone row.

Step 8: Repeat step 7 for the entire pixel present in the zone row.

Step 9: Compute the average pixel distance in the zone row. (One feature)

Step 10: Repeat steps 7 to 9 for the entire zone row present in the zone (Ten features).

Step 11: Repeat the steps 3-10 sequentially for the entire zone present in the image.

Step 12: Finally, make a vector of 26 rows and 1 column.

C. Pattern Matching

For pattern matching neural network and bacterial foraging optimisation is used to match the input with database. But before that database generation is also an important task. In our case a database of standard numeral characters and alphabets after training with neural network is generated. Thus in our database a neural network trained network is generated. This network consists of features of standard characters and a target vector is simulated with neural network. As said earlier neural network is strong in local optimization but weak in global optimization whereas with BFO, this is reverse. So, neural network is trained with BFO to overcome the limitations of network.

Algorithm of pattern Matching

- 1 Setting the basic parameters
- 2 Initialise BP neural network
- 3 Inputting training samples
- 4 Forward output of input vectors
- IF mean square error is less than Error 5
 - Optimise weights and biases with BFO
- GO to step 3 6
- 7 ELSE
 - Saving the network and carry out simulation

8 **END**

IV. **RESULT AND DISCUSSION**

To implement my proposed work MATLAB's image processing and neural network toolbox is used. We use bacterial foraging optimization with neural network to recognize the character for global optimization. This table is present all the parameter which we used for searching matching

dimension of search space	360
number of bacteria	10
Number of chemo tactic steps	10
length of a swim	4
number of elimination-dispersal events	2



The above figure describe the positions attained by bacteria in BFO The difference between neural network output and BFO optimized output for same characters is shown in below figure



Recognize 7 number is incorrect by NEURAL NETWORK



Recognize the same number is correct by BFONN

V. CONCLUSION

There are various ways to recognize the characters that can be handwritten character of any printed character. The main challenge posed in character recognition is by handwritten character recognition. In our work this challenge is tried to tackle by BFONN and succeeded upto a great extent. First of all preprocessing is done over the scanned input document which is used in recognition. RGB to binary conversion is done and then image is cropped to edges. For extraction of features coordinates of bounding box are taken out and then these features are simulated with neural network database. But some characters are not recognized by neural network because of their shapes, so BFO trained neural network is used. To train it by BFO input weights and biases of neural network are optimized to minimum mean square error and then the simulation is done by thus created network. It has been observed that BFONN read all characters correctly and even those which were not recognized by neural network. A GUY is constructed for fast processing speed and to make each function operable independent of each other. It also gives a human machine interaction to user so that operation of OCR engine is easy.

References

- [1]. Kartar Singh Siddharth, Mahesh Jangid, Renu Dhir, Rajneesh Rani," Handwritten Gurmukhi Character Recognition Using Statistical and Background Directional Distribution Features" IJCSE Vol. 3 No. 6 June 2011, pp 2332-2345
- [2]. Sukhpreet Singh, Ashutosh Aggarwal, Renu Dhir," Use of Gabor Filters for Recognition of Handwritten Gurmukhi Character" IJARCSSE Volume 2, Issue 5, May 2012, pp 234-240
- [3]. 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, pp 397-404

- [4]. Liu, Cheng-Lin, and Hiromichi 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. 2005.
- [5]. Vijay Laxmi Sahu, Babita Kubde," Offline Handwritten Character Recognition Techniques using Neural Network: A Review" IJSR Volume 2 Issue 1, January 2013 pp 87-94
- [6]. Gurpreet Singh Chandan Jyoti Kumar Rajneesh Rani Dr. Renu Dhir," Feature Extraction of Gurmukhi Script and Numerals: A Review of Offline Techniques" IJARCSSE Volume 3, Issue 1, January 2013 pp 257-263
- [7]. Dileep Kumar Patel, Tanmoy Som, Sushil Kumar Yadav, Manoj Kumar Singh," Handwritten Character Recognition Using Multiresolution Technique and Euclidean Distance Metric" JSIP 2012, 208-214
- [8]. Majida Ali Abed, Hamid Ali Abed Alasadi," Simplifying Handwritten Characters Recognition Using a Particle Swarm Optimization Approach" European Academic Research, Vol. I, Issue 5/ August 2013 pp-532-552
- [9]. Argha Roy, Diptam Dutta, Kaustav Choudhury," Training Artificial Neural Network using Particle Swarm Optimization Algorithm" IJARCSSE Volume 3, Issue 3, March 2013 pp 43—434
- [10]. Chirag I Patel, Ripal Patel, Palak Patel," Handwritten Character Recognition using Neural Network" IJSER Volume 2, Issue 5, May-2011 ISSN 2229-5518.
- [11]. Handley, Trevor. "Optical Character Recognition with Neural Networks."