

HCR(English) using Neural Network

Mahajan Mangesh A.¹, Navale Ganesh D.², Sharma Hemant P.³, Nikam Krushna V.⁴, Shaikh Riyaj R.⁵

^{1, 2, 3, 4, 5} Students,

Prof. H. Mohapatra.

Department of Computer Engineering,
SRES COE, Kopargaon,
Maharashtra, India

ABSTRACT

Recognition is the conversion of handwritten text into machine encoded text. Character recognition one of the emerging concept in field of Image Processing and Image workouts. Today many people all over the world have interest in digital content whatever it may. Due involvement of new electronic device and apps which has been specified with ability to scan, process, update, upload this technique of CR can be also utilize. Many research works and survey related to HCR has been done but working out with proper system technique and methods can be proved to providing maximum benefits. Different algorithm and methodologies has been working out related to HCR system one of such algorithm is Neural Network. HCR is discussed, this area requires more research and its challenging task for getting better accuracy.

Keyword:-Online Handwritten Recognition, Handwritten Character, Pattern Recognition, Neural Network, Feature Extraction and Segmentation.

1. INTRODUCTION

Written content can be anything a scribble over a page or any name our HCR system should easily recognize it and provide the quick response in Digital form. Text Format of any hand written Data in digital form has become a need in world of digitalization. The English language which has to be workout with proper efficiency.

HCR Can be offline or online. Online deals with captured on special electronic surface and Offline deals with printed or handwritten text format. The difficult task to recognize the offline handwritten characters is handwriting changes as per person, everyone has their different writing style and different handwriting.

They can be used in Banking Systems, Postal works, Printing press or even by Forensic teams. People get attract towards Banners, Information Blinking over LCDs, LEDs and Digital Projections, Wall charts where Scanning written paper work and quick output inform text over Digital devices can be beneficial in railways department and entertainment Fields.

2. LITERATURE SURVEY

Handwritten character recognition is process of converting the hand written work over page to a attractive digital format. HCR is a intelligent work done throw scanning the images will complete the analysis of character with output.

CR require proper handling of complexity of written content, writing environment, materials, etc. HCR techniques are based on extracting various features of handwritten.

Isha Vats, Shamandeep Singh[1] In this paper, system was based on recognition of offline handwriting numerals. The main aim of the proposed work in this paper was to efficiently recognize the offline handwritten digits with a higher accuracy. But a difficult problem in this field was the recognition of completely touching handwritten digits and in this paper the proposed system focused on segmentation for isolating the digits so multiple images can be recognize.

Gunjan Singh, Sushma Lehri [2] Handwritten characters was a difficult task because characters are written in various ways, so they could be of different sizes, orientation, thickness and dimension. An offline HCR(Hindi) system using neural network is presented in this paper. Neural networks were good at recognizing handwritten characters as these networks are insensitive to the missing data. A Backpropagation neural network is used for classification. Experimental result of this system shows that results 93%.

S S Sayyad, Abhay Jadhav, Manoj Jadhav, Pradip Bele, Smita Miraje, Avinash Pandhare [3] In this paper A neural network approach is proposed for automatic offline character recognition system. In this paper, work has been performed to recognize Devanagari characters using multilayer perceptron. Various patterns of characters were created in the matrix with the use of binary form and stored in the file. This system used the back propagation neural network for efficient recognition and neuron values were transmitted by feed forward method in the neural network.

Shabana Mehruz, Member IEEE, 2 Gauri katiyar[4] This paper provides review of existing works in HCR based on soft computing technique during the past decade.

Prof. Swapna Borde, Ms. Vinaya Patil, Ms. Ekta Shah, Ms. Priti Rawat [5] This paper presents a fuzzy approach to recognize characters. Fuzzy sets, fuzzy logic were used as bases for representation of fuzzy character and for recognition. Fuzzy-based algorithm which first segments the character and then using fuzzy system gives the characters that match the given input and then using defuzzification system finally recognizes the character. No training is needed by this system for recognition.

Fatos T. Yarman-Vural and Nafiz Arica [6] The rapidly growing computational power enables the implementation of the present Character Recognition methodologies and creates an increasing demand on many emerging application domains, which require more advanced methodologies. The available Character Recognition techniques with their superiorities and weaknesses are reviewed. The Character Recognition is discussed, and directions for future research are suggested. Special attention is given to the off-line HCR since this area requires more research.

Ms. Seema A. Dongare, Ms. Snehal V. Waghchaure, Prof. Dhananjay B. Kshirsagar [7] proposed system deals with development of grid based method which is combination of image centroid zone and zone centroid zone of individual character or numerical image. Use of feed forward neural network for recognition. Complete process of Devanagari character recognition works in stages as document preprocessing, segmentation, feature extraction, classification using grid based approach followed by recognition using feed forward NN.

Mitrakshi B. Patil, Vaibhav Narawade [8] This paper interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices for recognizing Handwritten Marathi Characters. In this paper, method for recognition of offline handwritten devnagari characters using segmentation and Artificial neural networks.

Mandeep Kaur, Sanjeev Kumar[9] This paper represent Handwritten Gurmukhi Character Recognition system using some statistical features like zone density, projection histograms, 8 directional zone density features in combination with some geometric features like area, perimeter, eccentricity, etc. Techniques like binarization, morphological operations applied to remove noise and then segmented into isolated characters.

Miroslav NOHAJ[10] Created a theoretical and practical basis of preprocessing of printed text for optical character recognition using forward-feed neural networks. Demonstration application was created and its parameters were set according to results of realized experiments.

Nisha Sharma, Tushar Patnaik, Bhupendra Kumar[11] In this paper major steps of an OCR system was discussed like preprocessing, segmentation, feature extraction, classification, postprocessing. This paper gives an overview of research work carried out for recognition of hand written English letters. Hand written letters are difficult to recognize due to diverse human handwriting style, variation in angle, size and shape of letters, also various feature extraction technique and classification method result was discussed.

J.Pradeep, E.Srinivasan and S.Himavathi[12] Off-line handwritten alphabetical character recognition system using multilayer feed forward neural network is given. Method like diagonal based feature extraction is introduced for extracting the features of the handwritten alphabets. Fifty data sets, each containing 26 alphabets written by various

people, are used for training the neural network and 570 different handwritten alphabetical characters are used for testing. This system performs quite well yielding higher levels of recognition accuracy compared to the systems employing the conventional horizontal and vertical methods of feature extraction. Suitable for converting handwritten documents into structural text form and recognizing handwritten names.

3. PROPOSED SYSTEM

This section contains the block diagram and the details about the modules we will be going to use.

General Procedure:

Handwriting recognition is a difficult problem which includes the recognition of input is given in form of image, scan paper. The handwritten character recognition generally involves the following Modules:-

A. Image Acquisition:

In the image acquisition the images for HCR system are acquired by appropriate scanning of handwritten documents, books or by capturing photographs of document. The input image is obtained by camera or through some scanner. The input image may be in gray, color.

B. Preprocessing:

The method of extraction of text from the document is called preprocessing. The preprocessing Consists of a series of operations performed on the scan input image, which include background Noise reduction, image restoration, filtering etc. This system assume that the character segmented is made thin to a unit pixel thickness. Various algorithms may be used for this purpose.

C. Segmentation:

This step deals with breaking of the lines, words for getting all the characters separated. This module involves the identification of the boundaries of the character and separating them for further processing for further steps. In this algorithm we will assume that this step was already done. Hence the input to our system will a single character.

D. Feature Extraction:

To find a set of parameters that uniquely defines the character is called feature extraction. The feature extraction technique should be such that the features of characters should enable clear discrimination of one character from others. To distinguish a class from other class a set of features is extracted for each class. The types of feature may be of statistical, syntactical/structural or hybrid in nature.

E. Classification:

This stage represents the decision making part of a recognition system and it employs the features extracted in the previous stage as inputs to the classifiers. The classifiers compare the input features with the stored features to assign a class for the input. The character recognition task is based on four approaches as template matching; statistical techniques; structural techniques and neural network. Here we are going to use ANN, an Artificial Neural Network usually called "neural network". It is a supervised learning method.

F. Post-processing:

The goal of post processing phase refers to detect and correct linguistic misspellings in the Offline HCR output text after the input image has been completely processed. Post processing steps are used to improve the accuracy of Offline handwritten character recognition system.

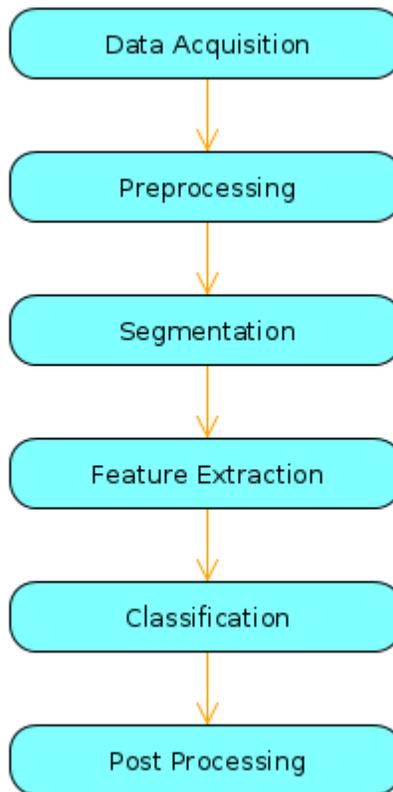


Fig -1: Block diagram of the system

4. NEURAL NETWORK

Neural Network Applying: When the features of the characters in the subword are determined, the next phase is to recognize the characters of the subword. The Neural Network approach will be used for this purpose. Various patterns of characters are created in the matrix ($n*n$) form and are stored in the file. A new pattern of a character is given input to the system. The system matches this pattern with the patterns which are stored in the file already. Neural networks are used for matching the patterns.

An Artificial Neural Network (ANN) consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In most cases an Artificial Neural Network is an adaptive system that changes its structure of external or internal information that flows through the network during the learning or training phase. An Artificial Neural Network usually consists of one input layer, more output layers, except the last output layer all the other intermediate layers are called hidden layer. Output of input layer is provide as input to the hidden layers and output of hidden layers are given as input to the final output layer Hidden layer will be used to integrate, collaborate similar features and if require then adjust the inputs by adding or subtracting weight values, finally one output layer is used to find the overall matching score of the network. If the score will be within the predefined range then the character will recognized else the system should have to train again.

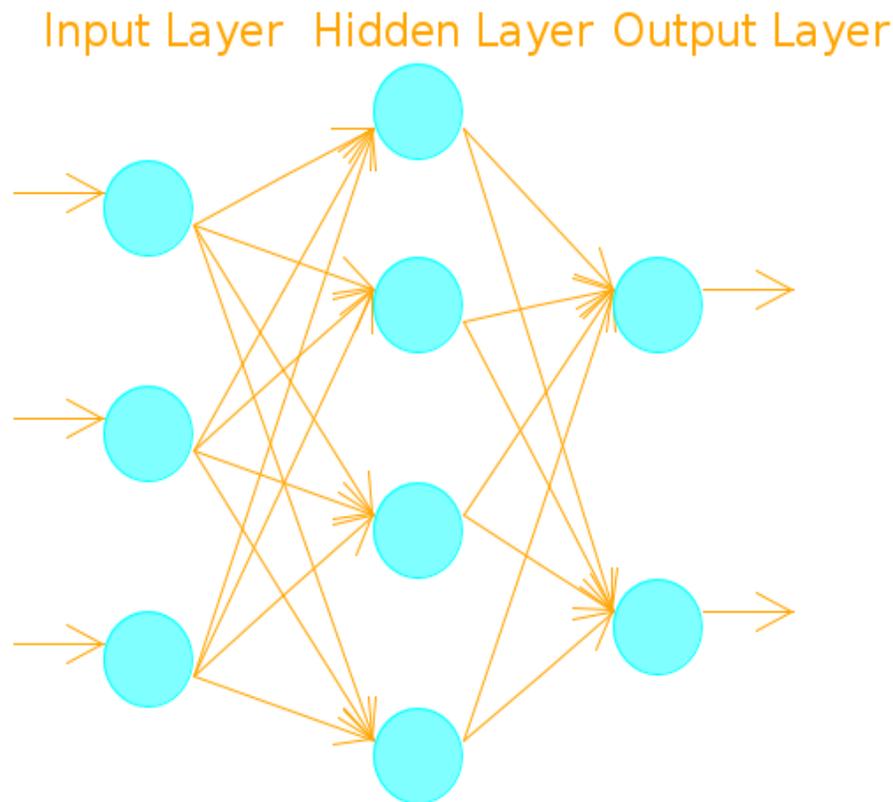


Figure 2: Multilayered Artificial Neural Network

5. CONCLUSIONS

Many regional languages throughout world have different writing styles which can be recognized with HCR systems using proper algorithm and strategies. We have learning for recognition of English characters. It has been found that recognition of handwritten character becomes difficult due to presence of odd characters or similarity in shapes for multiple characters. Scanned image is pre-processed to get a cleaned image and the characters are isolated into individual characters.

Preprocessing work is done in which normalization, filtration is performed using processing steps which produce noise free and clean output. Managing our evolution algorithm with proper training, evaluation other step wise process will lead to successful output of system with better efficiency. Use of some statistical features and geometric features through neural network will provided better recognition result of English characters. The existing handwritten has very low accuracy. We need an efficient solution to solve this problem so that performance can be increased. This work will be helpful to the researchers for the work towards other script.

ACKNOWLEDGEMENT

We dedicate all our project work to our estimated guide Prof. H. Mohapatra, whose interest and guidance help us to complete the work successfully. We also extend our gratitude to Prof.D.B.Kshirsagar(HOD of Computer Dept.) and Dr.A.B.Pawar(Project co-ordinator) who has provided facilities to explore the subject with more enthusiasm.

REFERENCES

- [1] Isha Vats, Shamandeep Singh, "Offline Handwritten English Numerals Recognition using Correlation Method", International Journal of Engineering Research and Technology (IJERT): ISSN: 2278-0181 Vol. 3 Issue 6, June 2014.
- [2] Gunjan Singh, Sushma Lehri, "Recognition of Handwritten Hindi Characters using Back propagation Neural Network", International Journal of Computer Science and Information Technologies ISSN 0975-9646, Vol. 3 (4) , 2012, 4892-4895.
- [3] S S Sayyad, Abhay Jadhav, Manoj Jadhav, Smita Miraje, Pradip Bele, Avinash Pandhare, "Devnagiri Character Recognition Using Neural Networks", International Journal of Engineering and Innovative Technology (IJEIT) Volume 3, Issue 1, July 2013.
- [4] Shabana Mehruz, Member IEEE, Gauri Katiyar, "Intelligent Systems for OffLine Handwritten Character Recognition: A Review", International Journal of Emerging Technology and Advanced Engineering Volume 2, Issue 4, April 2012.
- [5] Prof. Swapna Borde, Ms. Ekta Shah, Ms. Priti Rawat, Ms. Vinaya Patil, "Fuzzy Based Handwritten Character Recognition System", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622, VNCET 30 Mar'12.
- [6] Rahul KALA, Harsh VAZIRANI, Anupam SHUKLA and Ritu TIWARI, "An Overview of Character Recognition Focused on OffLine Handwriting", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS PART APPLICATIONS AND REVIEWS, VOL. 31, NO. 2, MAY 2001.
- [7] Ms. Seema A. Dongare, Prof. Dhananjay B. Kshirsagar, Ms. Snehal V. Waghchaure, "Handwritten Devanagari Character Recognition using Neural Network", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p-ISSN: 2278-8727 Volume 16, Issue 2, Ver. X (Mar-Apr. 2014), PP 74-79.
- [8] Mitrakshi B. Patil, Vaibhav Narawade, "Recognition of Handwritten Devnagari Characters through Segmentation and Artificial neural networks", International Journal of Engineering Research and Technology (IJERT) Vol. 1 Issue 6, August - 2012. ISSN: 2278-0181.
- [9] Mandeep Kaur, Sanjeev Kumar, "A RECOGNITION SYSTEM FOR HANDWRITTEN GURMUKHI CHARACTERS", International Journal of Engineering Research and Technology (IJERT) Vol. 1 Issue 6, August - 2012 ISSN: 2278-0181.
- [10] Miroslav NOHAJ, Rudolf JAK_A A, "Image preprocessing for optical character recognition using neural Networks", Journal of Pattern Recognition Research, 2011.
- [11] Nisha Sharma et al, "Recognition for handwritten English letters: A Review" International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 7, January 2013.
- [12] J. Pradeep et al, "Diagonal based feature extraction for handwritten alphabets recognition System using Neural network", International Journal of Computer Science and Information Technology (IJCSIT), Vol 3, No 1, Feb 2011.

BIOGRAPHIES

	Mahajan Mangesh A. is pursuing B.E. Computer Engg. in SRES COE, Kopargaon. His area of research interest include Image processing.
	Navale Ganesh D. is pursuing B.E. Computer Engg. in SRES COE, Kopargaon. His area of research interest include Image processing.
	Sharma Hemant P. is pursuing B.E. Computer Engg. in SRES COE, Kopargaon. His area of research interest include Image processing.
	Nikam Krushna V. is pursuing B.E. Computer Engg. in SRES COE, Kopargaon. His area of research interest include Image processing.
	Shaikh Riyaj R. is pursuing B.E. Computer Engg. in SRES COE, Kopargaon. His area of research interest include Image processing.