International journal of Emerging Trends in Science and Technology

Handwritten Pattern Recognition-A Survey

Authors
S. Manimurugan, Priya Mohan, Sissy Annamma Johnson

Dept. of Computer Science, MBCCET, Peermade

Abstract: Handwritten pattern recognition is one of the main areas which need an important attention because it can be used in various fields in the society such as banking. A survey on different handwritten pattern recognition systems is done in this paper. The aim is to provide a brief idea about various techniques used for the same. Even though there are number of methods, only the four papers based on Back-Propagation neural network, Text recognition based on neural network, optical character recognition and Comprehensive Devnagari numeral and character database are discussed here.Since the Handwritten pattern recognition is one of the most fascinating area, this paper will have an important role during the study about the handwritten pattern recognition.
Key Terms:Handwritten character recognition, Image processing, Feature extraction, feed forward neural networks.

I. INTRODUCTION

Pattern Recognition-the act of taking in raw data and making an action based on the category of pattern-has been crucial for our survival, and over the past tens of millions of years we have evolved highly sophisticated neural and cognitive systems for such tasks. It is natural that we should seek to design and build machines that can recognize patterns. From automated speech recognition, fingerprint identification and much more, it is clear that reliable, accurate pattern recognition by machine would be immensely useful.

The handwriting recognition provides higher advancement in automation process and also it improved the interaction between user and the system.The latest techniques provide higher accuracy when compare with the older systems. There are off-line recognition and on-line recognition systems. The on-line recognition system is more efficient why because, in the off-line recognition system the image is optically captured by a scanner and the completed writing isgiven as an image. But, in the case of on-line recognition systems the representation of two dimensional co-ordinates of successive points by using function of time and the order of strokes made by the writer [4][5].

The Handwritten recognition system has mainly 3 steps. They are image acquisition preprocessing, and segmentation. Segmentation is the important step in which the input image is separated into individual characters and resized into m x n pixels and given to the training network. The back end of the system is an artificial neural networkwhich is used for classification and recognition [2].

II. RELATED WORKS

In this paper, for making a survey on the topic Handwritten pattern recognition, mainly four papers arediscussed. The papers are: An approach to handwriting recognition using back-propagation neural network by PijushChakraborty andParamitaSarkar, Handwritten text recognition system based on neural network by Neeta Nain SubhashPanwar,OCR-optical character recognition by Line Eikvil, and Development of Comprehensive Devnagari Numeral and Character Database for Offline Handwritten Character Recognition by Vikas.J.Dongre and Vijay.H.Mankar.

This paper contains an overview of different handwritten pattern recognition systems. The methods explained in each paper and the advantages and disadvantages are discussed here.

a)Back-Propagation Neural Network

In computer science field, the handwritten recognition system is an unavoidable topic. A procedure which identify with the help of Back-Propagation neural network is discussed in this paper which is developed by PijushChakraborty and ParamitaSarkar. .The paper also discusses the various steps that are passed by the network for gaining the recognition ability. The steps include from feature extraction to the network training. The following figure gives an idea about the different steps included [5].
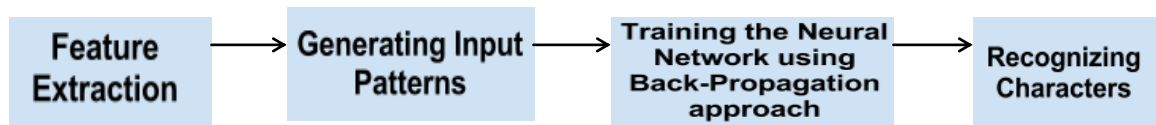

Fig.1:Steps used in the entire process

In Fig:1, initially the input character is given to the user interface. Using these characters the input patterns are generated by the Feature extraction phase. Using these input patterns the neural network is trained by back-propagation algorithm. After the training process the network can recognize the handwritten documents. The network can be updated with new data. The training is done by using supervised learning processes along with the error calculation methods.

b) Text Recognition System Based On Neural Network

In this paper, Neeta Nain SubhashPanwarMalaviya proposed a novel approach for handwriting recognition system. Here segmentation is used for image pre-processing and diagonal based feature extraction is used for character recognition. Paragraphs of running text are used as the input to the system which is pre-processed to segment it into normalized individual words. Further, a diagonal based feature extraction technique is used for extracting the features of handwritten alphabets. The proposed recognition system performs excellently for separate character written documents with 100% accuracy [6].

The author said that using this method, an accuracy of 60% for the cursive handwritten text and a 75% for the readable non-cursive handwritten text documents can be finalize. There will be some problems like Image binarization, Line Segmentation, Skew Correction, Image Scaling, word Segmentation.[9]

c) OCR- Optical Character Recognition

Character recognition systems are mainly associated with the symbolic identity of the image of the character. The widely used character recognition technique is optical character recognition (OCR). In the field of pattern recognition and artificial intelligence, Optical Character Recognition is an important technique.OCR exists for a variety of applications, in many commercial applications, text entry (office automation), data entry, and process automation. Automatic identification is the technique behind each optical character recognition.Optical Character Recognition systems can only recognize optically processed characters. Both the offline and online systems are using OCR technology. But online recognition is more efficient than offline system. The system can recognize both written and printed documents. These paragraphs are pre-processed for segmentation that is to segment each character separately. After segmentation the method diagonal based feature extraction is used for recognizing each character [7].

Fig .2 represents various steps performed in an optical Character Recognition system. In the system, initially the input running paragraphs are given to the optical scanning device. The scanned image is segmented into individual characters after preprocessing. After this stage feature extraction technique is used for recognizing the characters.
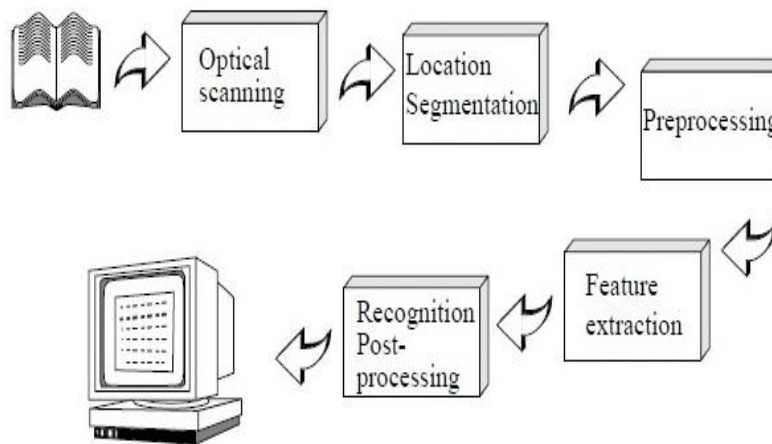

Fig.2: Components of OCR system

Even though the machines are far away from the human reading capabilities, many commercial systems for performing OCR exist for a variety of applications. In the feature extraction phase, the essential characteristics of the symbols are captured and it is generally accepted that this is one of the most difficult problems of pattern recognition. Matching is done based on the similarity measures where the distance between the feature vectors, describing the extracted character and the description of each class is calculated.Different measures may be used, but the common is the Euclidean distance [7].

d)Development of Comprehensive Devnagari Numeral and Character Database for Offline Handwritten Character Recognition

In this paper a separate database is developed for Devnagari numerals and characters. The main drawback of earliest methods is the difficulty in recognizing writings of different persons; such writings will be rejected by the system. For avoiding this problem use a separate database for storing data. In this database asingle character written by different individuals in different way can be updated frequently. The learning process is done with the help of this data in the database.

In the case of Indian languages, recognizing the characters is a difficult problem. By using the database, we can reduce the difficulty up to a limit. Nowadays the same is done for some Indian languages such as Bangla, Kannada, and Devnagari.Since India is a multilingual and multi

script country having more than 1.2 billion population with 22 constitutional languages and 10 different scripts and Devnagari is the most popular script in India, it is important that a study on Devnagari script. Hindi, the national language of India which is spoken by more than 500 million populations worldwide, is written in the Devnagari script [9].
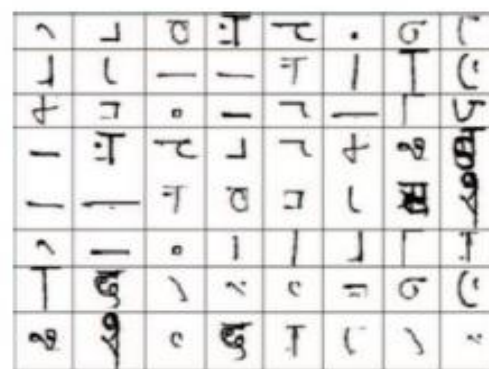
(a) Invalid isolated strokes
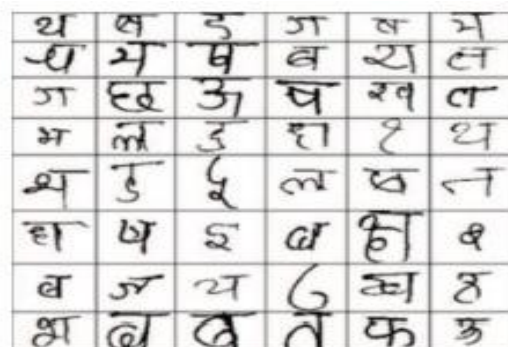
(b) Ambiguous characters

Fig.3: Valid and Invalid characters in Devnagari

The fig.3 shows valid and invalid characters in a particular Indian language. In the figure (a) the letters cannot segment successfully because some parts are not in the correct position. This will lead to a problem that, while matching process is done, it may not recognize the correct character[10].Thus when the data is collected, it should retain the proper features. This can be considered as a problem in this area.For this we are collecting data that is different documents that are written by various people in India and train the network using this data. But the problems associated with this method are invalid isolated strokes, Ambiguous characters etc.

III. CONCLUSION

Various handwritten pattern recognition systems are discussed in this paper .Even though there are number of techniques are available, only four of them are discussed here. The papers are Back-propagtion neural network, Text recognition based on neural network, optical character recognition and Comprehensive Devnagari numeral and character database. The each papers discuss different methods effectively, each of them have disadvantages also. Now feature extraction is generally used for recognizing the characters effectively.OCR systems are used in both offline and online systems.in case of online and offline systems, online systems are more efficient because advanced techniques are used for CR.

IV. REFERENCES

J. Pradeep, E. Srinivasan and S. Himavathi .,"Diagonal Based Feature Extraction for Handwritten Alphabets Recognition System using Neural Networks .,vol 3.,no 1.,Feb 2011

Anil.K.J ,Jain and TorfinnTaxt, Pattern recognition.,"Feature extraction methods for character recognition" – A Survey by, vol.29, no.4, pp.641-662,1996.

R.Plamondon and S.N.Srihari.,"Online and offline handwritten character recognition" Transactions on pattern analysis and machine intelligence, vol.22, no.1,pp.63-84, 2000.

N. Arica and F. Yarman- Vural.,"An overview of character recognition focused on off-line Handwriting"-IEEE Transactions on systems, man and cybernetics, Part-C: Applications and Reviews,2001, 31(2), pp.216-233.

Pijush Chakra borty Student (B.Tech, CSE) Calcutta Institute of Engineering and Management ParamitaSarkar Asst. Professor, CSE dept. Calcutta Institute of Engineering and Management 1993."An Approach to Handwriting Recognition using Back-Propagation Neural Network".

Neeta Nain SubhashPanwarMalaviya National Institute of Technology Jaipur JLN Marg Jaipur 1993."Handwritten Text Recognition System Based On Neural Network".

OCR-Optical Character Recognition Line Eikvil December 1995.

R. Plamondon and S. N. Srihari, "On-line and off- line handwritten character recognition: A comprehensive survey",IEEE. Transactions on Pattern Analysis Recent Advances in Handwriting Recognition FlávioBortolozzi, Alceu de Souza Britto Jr., Luiz S. Oliveira and Marisa Morita Pontifical Catholic University of Paraná (PUCPR) RuaImaculadaConceicao, 1155 – CEP 80215-901 Curitiba - Paraná –Brazil 1996

Vikas J. Dongre and Vijay H.Mankar.,"Development of Comprehensive Devnagari Numeral and Character Database for Offline Handwritten Character Recognition".1998.

NallapareddyPriyanka, Srikanta Pal, RanjuMandal, (2010) "Line and Word SegmentationApproach for Printed Documents", IJCA Special Issue on Recent Trends in Image Processing andPattern Recognition-RTIPPR, pp. 30-36.