Offline Handwritten Character Recognition Techniques using Neural Network: A Review

Vijay Laxmi Sahu[1], Babita Kubde[2]

[1]Rungta College of Engineering & Technology
Bhilai, Chhattisgarh, India - 490021
vijaylaxmibit1987@gmail.com

[2]Rungta College of Engineering & Technology
Bhilai, Chhattisgarh, India - 490021
babita_g7@*rediffmail.com*

Abstract: ***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.***

Keywords: Neural Network, Feature extraction, Segmentation and Training, Classification.

1. Introduction

Handwriting recognition has been a subject of research for several decades. Pattern recognition has three main steps: observation, pattern segmentation, and pattern classification. Recognition of handwritten character is one of the most interesting topics in pattern recognition. Optical Character Recognition (OCR) systems aim at transforming large amount of documents, either printed or handwritten into machine encoded text.

In general, handwriting recognition is classified into two types as off-line and on-line handwriting recognition methods. Off-line handwriting recognition involves automatic conversion of text into an image into letter codes which are usable within computer and text-processing applications. The data obtained by this form is regarded as a static representation of handwriting. Off-line handwriting recognition is comparatively difficult, as different people have different handwriting styles. 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 [1][2].

Optical Character Recognition (OCR) is a field of research in pattern recognition, artificial intelligence and machine vision. Optical character recognition (OCR) is usually referred to as an off-line character recognition process to mean that the system scans and recognizes static images of the characters. It refers to the mechanical or electronic translation of images of handwritten, typewritten or printed text into machine- editable text. OCR consists of many phases such as Pre-processing, Segmentation, Feature Extraction, Classifications and Recognition. The input of one step is the output of next step. The task of preprocessing relates to the removal of noise and variation in handwritten word patterns. Preprocessing may itself be broken down into smaller tasks such as noise removal, Binarization, Thinning, Edge Detection, slant estimation and correction, skew detection, resizing etc to enhance the quality of images and to correct distortion[3][4].

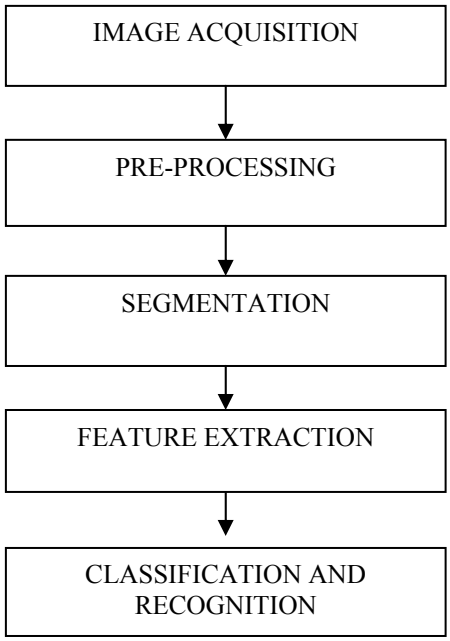

Figure 1: Stages in OCR

Several applications including mail sorting, bank processing, document reading and postal address recognition require off line handwriting recognition systems. As a result, the off-line handwriting recognition continues to be an active area of research towards exploring the newer techniques that would

improve recognition accuracy[5]. Selection of relevant feature extraction plays important role in performance of character recognition.

2. Phases of General Character Recognition System

The Research in this field basically involves the following activities:-

2.1 Image Acquisition

In this phase the input image taken through camera or some scanner. The image should have a specific format such as JPEG; BMT etc. The input captured may be in gray, color or binary from scanner or digital camera.

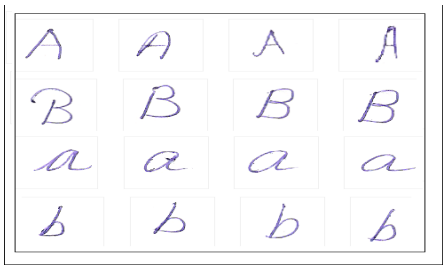

Figure 2: Sample Dataset

2.2 Pre-processing

The pre-processing is a series of operations performed on the scanned input image. It essentially enhances the image rendering it suitable for segmentation [5]. The various tasks performed on the image in pre-processing stage are shown in Fig.1.

2.2.1 Noise Reduction

When the document is scanned, the scanned images might be contaminated by additive noise and these low quality images will affect the next step of document processing. Therefore, a pre-processing step is required to improve the quality of images before sending them to subsequent stages of document processing. Due to the noise there can be the disconnected line segment , large gaps between the lines etc. so it is very essential to remove all of these errors so that's the information can be retrieved in the best way.

There are many kinds of noise in images. One additive noise called "Salt and Pepper Noise", the black points and white points sprinkled all over an image, typically looks like salt and pepper, which can be found in almost all documents. Noise reduction techniques can be categorized in two major groups as filtering, morphological operations.

(a) Filtering

It aims to remove noise and diminish spurious points, usually introduced by uneven writing surface and/or poor sampling rate of the data acquisition device. Various spatial and frequency domain filters can be designed for this purpose [10].

(b) Morphological Operations

Morphological operations are commonly used as a tool in image processing for extracting image components that are useful in the representation and description of region shape. Morphological operations can be successfully used to remove the noise on the document images due to low quality of paper and ink, as well as erratic hand movement.

2.2.2 Binarization

Binarization of gray-scale character images is a crucial step in offline character recognition. Good binarization facilitates segmentation and recognition of characters. Binarization process converts a gray scale image into a binary image. In this paper [14], has described new methods for the binarization of noisy gray-scale character images obtained in an industrial setting. Our methods are specially designed to binarize gray-scale character images more effectively by using the fact that characters are usually composed of thin lines of uniform width. Experimental results show that these methods give the best binarization results.

2.2.3 Edge Detection

Edges characterize object boundaries and are therefore useful for segmentation, registration, and identification of objects. Edge detecting an image significantly reduces the amount of data and filters out useless information, while preserving the important structural properties in an image. There are many ways to perform edge detection. However, the majority of different methods may be grouped into two categories, gradient and Laplacian. The gradient method detects the edges by looking for the maximum and minimum in the first derivative of the image. The Laplacian method searches for zero crossings in the second derivative of the image to find edges [12].

2.2.4 Thresholding

In order to reduce storage requirements and to increase processing speed, it is often desirable to represent grey scale or color images as binary images by picking some threshold value for everything above that value is set to 1 and everything below is set to 0.

Two categories of thresholding exist: Global and Adaptive. Global thresholding picks one threshold value for the entire document image, often based on an estimation of the background level from the intensity histogram of the image. Adaptive thresholding is a method used for images in which different regions of the image may require different threshold values [8]. In [21], a comparison of many common thresholding techniques is given by using an evaluation criterion that is goal-directed in the sense that the accuracies of a character recognition system using different techniques were compared. On those Tested, Niblack's method [22] produced the best result.

2.2.5 Skew Detection

For a document scanning process, there can be the skewness.

There are several commonly used methods for detecting skew in a page; some rely on detecting connected components and finding the average angles connecting their centroids. The skewness should be removed because it reduces the accuracy of the document. The skew angle is calculated and with the help of skew angle, the skewed lines are made horizontal [11].

2.2.6 Slant Estimation and Normalization

Handwritten text is usually characterized by slanted characters. In particular, the slanted characters slope either from right to left or vice versa. Moreover, different deviations may appear not only within a text but also within a single word. The slant correction does not affect the connectivity of the word and the resulting words are natural. Slant normalization is used to normalize all characters to a standard form. The most common method for slant estimation is the calculation of the average angle of near-vertical elements.

In this research paper [13], a slant removal algorithm is presented based on the use of the vertical projection profile of word images and the Wigner-Ville distribution. In this paper [15], slant detection is performed by dividing the image into vertical and horizontal windows. The slant is estimated based on the center of gravity of the upper and lower half of each window averaged over all the windows. Another study in paper [16], in this paper several methods have been proposed for average slant estimation and correction. However, average slant estimation has the problem such that local slant will be overestimated or underestimated when the slant in a word varies from character to character.

To solve the problem, this paper proposes three methods for local slant estimation, which are simple iterative method, high speed iterative method and 8-directional chain code method. The experimental results show that the proposed methods can estimate and correct local slant more accurately than the average slant correction. Lastly, in [19] a variant of Hough transform is used by scanning left to right across the image and calculating projections in the direction of 21 different slants. The top three projections for any slant are added and the slant with the largest count is taken as the slant value.

2.3 Segmentation

In Character Recognition techniques, the Segmentation is the most important process. Segmentation is done to make the separation between the individual characters of an image. Segmentation of unconstrained handwritten word into different zones (upper middle and lower) and characters is more difficult than that of printed documents. This is mainly because of variability in inter-character distance, skew, slant, size and curved like handwriting. Sometimes components of two consecutive characters may be touched or overlapped and this situation complicates the segmentation task greatly. In Indian languages such touching or overlapping occurs frequently because of modified characters of upper-zone and lower-zone[22].Segmentation is an important stage, because the extent one can reach in separation of words, lines or characters directly affects the recognition rate of the script. There are two types of segmentation:

2.3.1 External Segmentation

External segmentation decomposes the page layout into its logical units. External segmentation is the isolation of various writing units, such as paragraphs, sentences or words. It is the most critical part of document analysis. Document Analysis and Recognition (DAR) aims at the automatic extraction of information presented on paper and initially addressed to human comprehension. Segmenting the document image into text and non-text regions is an integral part of the OCR software. Therefore, one who works in the CR field should have a general overview for document analysis techniques. Page segmentation is one important step in layout analysis and is particularly difficult when dealing with complex layouts. Page layout analysis is accomplished in two stages: The first stage is the structural analysis, which is concerned with the segmentation of the image into blocks of document components (paragraph, row, word, etc). The second one is the functional analysis, which uses location, size and various layout rules to label the functional content of document components [23][24]. Page segmentation is then implemented by finding textured regions in gray-scale or color images.

For example, in paper [25] a method for automatically evaluating the quality of document page segmentation algorithms is introduced. They have proposed a bitmap-level automatic scheme to benchmark page segmentation algorithms on mixed text/halftone documents. It provides an accurate qualitative diagnosis of segmentation techniques, from which, a quantitative evaluation is derived.

2.3.2 Internal Segmentation

Internal Segmentation is an operation that seeks to decompose an image of a sequence of characters into sub images of individual symbols. Although, the methods have developed remarkably in the last decade and a variety of techniques have emerged, segmentation of cursive script into letters is still an unsolved problem. Character segmentation strategies are divided into three categories [26].

2.4 Feature Extraction

Feature extraction is the process to retrieve the most important data from the raw data. The most important data means that's on the basis of that's the characters can be represented accurately. The major goal of feature extraction is to extract a set of features, which maximizes the recognition rate with the least amount of elements. In feature extraction stage each character is represented as a feature vector, which becomes its identity. Due to the nature of handwriting with its high degree of variability and imprecision obtaining these features, is a difficult task. Feature extraction methods are based on 3 types of features-Statistical, Structural, Global transformations and moments [6]. Structural and statistical features appear to be complementary in that they highlight different properties of the characters. The widely used feature extraction methods are Template matching, Deformable templates, Unitary Image transforms, Graph description, Projection Histograms,

Contour profiles, Zoning, Geometric moment invariants, Zernike Moments, Spline curve approximation, Fourier descriptors, Gradient feature and Gabor features.

2.4.1 Statistical Features

These features are derived from the statistical distribution of points. They provide high speed and low complexity and take care of style variations to some extent [7] [8]. The followings are the major statistical features:

(a) Zoning

The character image is divided into NxM zones. From each zone features are extracted to form the feature vector. The goal of zoning is to obtain the local characteristics instead of global characteristics [6].

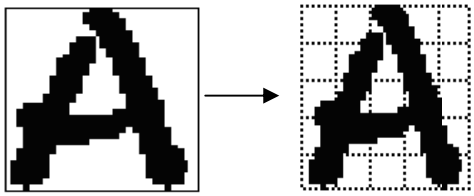

Figure 3: Zoning

In this paper [5], Diagonal feature extraction scheme for recognizing off-line handwritten characters is proposed. In this every character image of size 90x 60 pixels is divided into 54 equal zones, each of size 10x10 pixels. The features are extracted from each zone pixels by moving along the diagonals of its respective 10x10 pixels. Each zone has 19 diagonal lines and the foreground pixels present along each diagonal line is summed to get a single sub-feature. Thus 19 sub features are obtained from the each zone. These 19 sub features values are averaged to form a single feature value and placed in the corresponding zone.

(b) Characteristics Loci

For every white point in the background of the character, vertical and horizontal vectors are generated. The number of times that the line segments intersected by these vectors are used as features [8].

In this paper [27], work is concerned with handwritten and printed numeral recognition based on an improved version of the loci characteristic method (CL) for extracting the numeral features. After a preprocessing of the numeral image, the method divides the image into four equal parts and applies the traditional CL to each of the parts. The recognition rate obtained by this method is improved indicating that the numeral features extracted contain more details. In this paper [28], Glucksman's "characteristic loci" were utilized in experiments with the well-known Highleyman data, as well as samples generated at Stanford Research Institute and Honeywell. Two recognition algorithms were tested. Results on numeric samples compare favorably with those of other investigators despite the small dimensionality of the feature vector. On the constrained Honeywell samples, recognition rates exceeding 98 percent were achieved using the simpler algorithm.

(c) Crossing and Distance

Crossings count the number of transitions from background to foreground pixels along vertical and horizontal lines through the character image and Distances calculate the distances of the first image pixel detected from the upper and lower boundaries, of the image, along vertical lines and from the left and right boundaries along horizontal lines. Another study [46] encodes the location and number of transitions from background to foreground pixels along vertical lines through the word. Also, the distance of line segments from a given boundary, such as the upper and lower portion of the frame, can be used as statistical features [45].

2.4.2 Structural Features

Characters can be represented by structural features with high tolerance to distortions and style variations. This type of representation may also encode some knowledge about the structure of the object or may provide some knowledge as to what sort of components make up that object. Structural features are based on topological and geometrical properties of the character, such as aspect ratio, cross points, loops, branch points, strokes and their directions, inflection between two points, horizontal curves at top or bottom, etc [6].

2.4.3 Global Transformation and Series Expansion

It includes Fourier Transform, Gabor Transforms, wavelets, moments and Karhuen-Loeve Expansion**.** A continuous signal generally contains more information than needs to be represented for the purpose of classification. This may be true for discrete approximations of continuous signals as well. One way to represent a signal is by a linear combination of a series of simpler well-defined functions. The coefficients of the linear combination provide a compact encoding known as transformation or/and series expansion Deformations like translation and rotations are invariant under global transformation and series expansion. Common transform and series expansion methods used in the CR field are:

Gabor Transform: The Gabor transform is the one type of short time Fourier transform. The use of Gabor transform is to find the sinusoidal frequency. The Gabor transform is also used to find the phase content of local sections of a signal as it changes over time. The function which is to be transformed first of all is multiplied by a Gaussian function, and the result is known as a window function. The window function is then transformed with a Fourier transform which derives the time frequency analysis. The window function means that the signal near the time being analyzed will have higher weight [11].

Fourier Transforms*:* The general procedure is to choose magnitude spectrum of the measurement vector as the features in an n-dimensional Euclidean space. One of the most attractive properties of the Fourier transform is the ability to recognize the position-shifted characters, when it observes the magnitude spectrum and ignores the phase. Fourier transforms have been applied to CR in many ways [17][18].

Wavelets: Wavelet transformation is a series expansion technique that allows us to represent the signal at different levels of resolution. In OCR area, it is our advantage to handle each resolution separately [20].

2.5 Classification and Recognition

The classification stage is the decision making part of a recognition system and it uses the features extracted in the previous stage. We summarize the classification methods in categories of statistical methods, artificial neural networks (ANNs), kernel methods, and multiple classifier combination. Character classifier can be Baye's classifier, nearest neighbor classifier, Radial basis function, Support Vector Machine, Neural Network etc. Numerous techniques for CR can be investigated in four general approaches of Pattern Recognition, as suggested in: Template Matching; Statistical Techniques; Structural Techniques; Neural Networks.

2.5.1 Template Matching

Optical Character Recognition by using Template Matching is a system prototype that useful to recognize the character or alphabet by comparing two images of the alphabet. Template matching is the process of finding the location of a sub image called a template inside an image. Once a number of corresponding templates is found their centers are used as corresponding points to determine the registration parameters. Template matching involves determining similarities between a given template and windows of the same size in an image and identifying the window that produces the highest similarity measure [42]. Matching techniques can be studied in two classes.

Deformable Templates and Elastic Matching: Deformable templates have been used extensively in several object recognition applications. An alternative method is the use of deformable templates, where an image deformation is used to match an unknown image against a database of known images. Two characters are matched by deforming the shape of one, to fit the edge power of the other [48]. The basic idea of elastic matching is to optimally match the unknown symbol against all possible elastic stretching and compression of each prototype A dissimilarity measure is derived from the amount of bend needed, the decency of fit of the edges and the interior overlap between the distorted shapes (see figure 4). Recently Del Bimbo et al.[44] proposed to use deformable templates for character recognition in gray scale images of credit card slips with poor print quality. The templates used were character skeletons .It is not clear how the initial positions in the image were to be tried, then the computational time would be prohibitive.

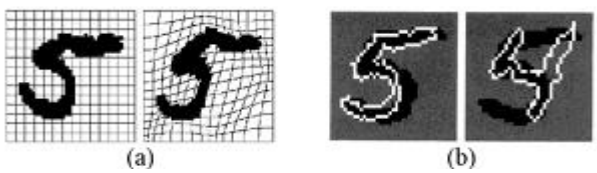

Figure 4 (a): Deformations of a sample digit, **(b)** Deformed Template superimposed on target image, with dissimilarity measures [47]

Direct Matching: A gray-level or binary input character is directly compared to a standard set of stored prototypes. According to a similarity measure (e.g.:Euclidean ,Mahalanobis, Jaccard or Yule similarity measures etc.), a prototype matching is done for recognition. The matching techniques can be as simple as one-to-one comparison or as complex as decision tree analysis in which only selected pixels are tested. Although direct matching method is intuitive and has a solid mathematical background, the recognition rate of this method is very sensitive to noise [2].

In [50] Srihari et al. propose a parallel architecture for offline cursive script word recognition, where they combine three algorithms; template matching, mixed statistical-structural classifier and structural classifier. The results derived from three algorithms are combined in a logical way. Significant increase in the recognition rate is reported.

2.5.2 Statistical methods

Statistical classifiers are rooted in the Bayes decision rule, and can be divided into parametric ones and non-parametric ones [30] [31]. Non-parametric methods, such as Parzen window and k-NN rule, are not practical for real-time applications since all training samples are stored and compared. The major statistical approaches, applied in the CR field are the followings:

a) **Non-parametric Recognition**

The finest known method of non-parametric categorization is the Nearest Neighbor (NN) and is widely used in CR. An incoming pattern is classified using the cluster, whose center is the minimum distance from the pattern over all the clusters. It does not involve a priori information about the data [51].

b) Parametric Recognition

Since a priori information is available about the characters in the training data, it is possible to obtain a parametric model for each character [52]. Once the consideration of the model, which is based on some probabilities, is obtained, the characters are classify according to some decision rules such as Baye's method or maximum Likelihood.

In this paper [29], a novel character recognition system is proposed in this paper. By using the virtual reconfigurable architecture-based evolvable hardware, a series of recognition systems are evolved. To improve the recognition accuracy of the proposed systems, a statistical pattern recognition-inspired methodology is introduced. The performance of the proposed method is evaluated on the recognition of characters with different levels of noise. The experimental results show that the proposed statistical pattern recognition-based scheme significantly outperforms the traditional approach in terms of character recognition accuracy. For 1-bit noise, the recognition accuracy is increased from 84.8% to 96.7%.In this paper [33] a handwritten Kannada and English Character recognition system based on spatial features is presented. Directional spatial features via stroke length, stroke density and the number of stokes are employed as potential & relevant features to characterize the handwritten Kannada

numerals/vowels and English uppercase alphabets. KNN classifier is used to classify the characters based on these features with four fold cross validation. The proposed system achieves the recognition accuracy as 96.2%, 90.1% and 91.04% for handwritten Kannada numerals, vowels and English uppercase alphabets respectively.

2.5.3Structural Techniques

Within the area of structural recognition, syntactic methods are among the most prevalent approaches. These patterns are used to describe and classify the characters in the CR systems.

(a) Syntactic methods

Measures of similarity based on relationships between structural components may be formulated by using grammatical concepts. The idea is that each class has its own grammar defining the composition of the character. A grammar may be represented as strings or trees, and the structural component extracted from an unknown character is matched against the grammars of each class. Suppose that we have two different character classes which can be generated by the two grammars G1 and G2, respectively. Given an unknown character, we say that it is more similar to the first class if it may be generated by the grammar G1, but not by G2.

2.5.4 Neural network

An Artificial Neural Network as the backend is used for performing classification and Recognition tasks. In offline character recognition systems, the Neural Network has emerged as the fast and reliable tools for classification towards achieving high recognition. Neural network architectures can be classified into two major sets specifically; feed-forward and feedback (recurrent) networks and the majority common ANN used in the CR systems are the multilayer perceptron of the feed forward networks and the Kohonens Self Organizing Map (SOM) of the feedback networks, use Feed Forward Neural Network. In a feed-forward neural network, nodes are organized into layers; each "stacked" on one another. The neural network consists of an input layer of nodes, one or more hidden layers, and an output layer . Each node in the layer has one corresponding node in the next layer, thus creating the stacking effect. Back propagation is a learning rule for the training of multi-layer feed-forward neural network. Back propagation derives its name from the technique of propagating the error in the network backward from the output layer. To train a Back propagation neural network, it must be exposed to a training data set and the answers or correct interpretations of the set [32].

The RBF network can yield competitive accuracy with the MLP when training all parameters by error minimization [35]. Vector quantization (VQ) networks and auto-association networks, with the sub-net of each class trained independently in unsupervised learning, are also useful for classification. The learning vector quantization

(LVQ) of Kohonen [36] is a supervised learning method and can give higher classification accuracy than VQ.

J.Pradeep et al [5] applied an offline handwritten alphabetic character recognition system using multilayer feed forward network. Diagonal based feature extraction is introduced in this method. So dataset each containing 26 alphabets written by various people is used for training the neural network & 570 different alphabets are used for training.

T.P. Singh et al [37] presented an effort to compare the performance for pattern recognition with conventional hebbian learning rule and with evolutionary algorithm in Hopfield model of feed forward network. The storing of the object has been performed using hebbian rule and recalling of these stored pattern on presentation of proto-type input pattern has been used by using both convolution hebbian rule and evolutionary algorithm.

The feed forward NN approach to the machine-printed CR problem is proven to be successful in [38], where the NN is trained with a database of 94 characters and tested in 300 000 characters generated by a postscript laser printer, with 12 common fonts in varying size. No errors were detected. In this study, Garland *et al.* propose a two-layer NN, trained by a centroid dithering process.

The modular NN architecture is used for unconstrained handwritten numeral recognition in [39]. The whole classifier is composed of sub networks. A sub network, which contains three layers, is responsible for a class among ten classes.

A recent study proposed by Maragos and Pessoa incorporates the properties of multilayer perceptron and morphological rank NNs for handwritten CR. They claim that this unified approach gives higher recognition rates than a multilayer perceptron with smaller processing time [40].

In Multiple classifier combination, combining multiple classifiers has been long pursued for improving the accuracy of single classifiers. Parallel (horizontal) combination is more often adopted for high accuracy, while sequential (cascaded, vertical) combination is mainly used for accelerating large category set classification.

In this paper [41], the paper describes the process of character recognition using the Multi Class SVM classifier. This paper presents a system of English handwritten character recognition. Recognition results with statistical feature are 98% which is better than that of recognition results with structural features that is 97%. By combining both feature sets that is statistical and structural the highest recognition rates are possible, which is 99.9%.

Kernel methods give a systematic and principled approach to training learning machines and the good generalization Kernel methods, including support vector machines (SVMs) primarily and kernel principal component analysis (KPCA), kernel Fisher discriminant analysis (KFDA), etc. are receiving increasing attention and have shown superior performance in pattern recognition. An SVM is a binary classifier with discriminant function being the weighted combination of kernel functions over all training samples. Kernel based Radial Basis Function (RBF) networks have been widely studied because they exhibit good generalization and universal approximation through use of RBF nodes in the hidden layer.

In this paper [34], a recognition model for English handwritten character recognition has proposed that uses Freeman chain code (FCC) as the representation technique of an image character. FCC is generated from the characters that used as the features for classification. The main problem in representing the characters using FCC is the length of the FCC that depends on the starting points. Then classification using the features generated from FCC is performed by SVM. Our recognition model was built from SVM classifiers. Our test results shows that applying the proposed model, we reached a relatively high accuracy for the problem of English handwritten recognition.

In [49], Xu et al. studied the methods of combining multiple classifiers and their application to handwritten recognition. They proposed a serial combination of structural classification and relaxation matching algorithm for the recognition of handwritten zip codes. It is reported that the algorithm has very low error rate and high computational cost.

3. Conclusion

It is hoped that this detailed discussion will be beneficial insight into various concepts involved, and boost further advances in the area. The accurate recognition is directly depending on the nature of the material to be read and by its quality. Current research is not directly concern to the characters, but also words and phrases, and even the complete documents. From various studies we have seen that selection of relevant feature extraction and classification technique plays an important role in performance of character recognition rate. This review establishes a complete system that converts scanned images of handwritten characters to text documents. This material serves as a guide and update for readers working in the Character Recognition area.

4. Future Work

A lot of Research is still needed for exploiting new features to improve the current performance. We can use some features specific to the mostly confusing characters, to increase the recognition rate. To recognize strings in the form of words or sentences segmentation phase play a major role for segmentation at character level and modifier level. So, there is still a need to do the research in the area character recognition.

5. References

[1] Anita Jindal, Renu Dhir, Rajneesh Rani "Diagonal Features and SVM Classifier for Handwritten Gurumukhi Character Recognition," Volume 2, Issue 5, May 2012 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering.

[2] 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, vol.31 no.2, pp. 216 - 233. 2001.

[3] Gita Sinha, Anita Rani, Prof. Renu Dhir, Mrs. Rajneesh Rani "Zone-Based Feature Extraction Techniques and SVM for Handwritten Gurmukhi Character Recognition," Volume2, Issue 6, June 2012 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering.

[4] S.V. Rajashekararadhya, Dr P. Vanaja Ranjan, 2008 "efficient zone based feature extraction algorithm for handwritten numeral recognition of four popular south indian" journal of theoretical and applied information technology.

[5] J.Pradeep, E.Srinivasan, S.Himavathi "Diagonal Based Feature Extraction for Handwritten Character Recognition System Using Neural Network".

[6] Giorgos Vamvakas" Optical Character Recognition for Handwritten Characters" National Center for Scientific Research "Demokritos" Athens – Greece Institute of Informatics and Telecommunications Computational Intelligence Laboratory (CIL).

[7] C.Y. Suen, M. Berthod and S. Mori, Automatic Recogniti"on of Handprinted-Characters _ the State of the Art in Proceedings of the IEEE, Vol: 68, No: 4, 1980.

[8] Nariz Arica" An Offline Character Recognition System for Free Style Handwritting" 1998.

[9] Nafiz Arica, Fatos T. Yarman-Vural," An Overview Of Character Recognition Focused On Off-line Handwriting".

[10] S. Mo, V. J. Mathews, Adaptive, Quadratic Preprocessing of Document Images for Binarization, IEEE Trans. Image Processing 7(7), 992-999, 1998.

[11] Neeraj Pratap1 and Dr. Shwetank Arya "A Review of Devnagari Character Recognition from Past to Future" International Journal of Computer Science and Telecommunications [Volume 3, Issue 6, June 2012].

[12] Bill GREEN Edge Detection Tutorial.

[13] E. Kavallieratou, N. Fakotakis, G. Kokkinakis" Slant estimation algorithm for OCR systems" Pattern Recognition 34 (2001) 2515}2522.

[14] Jeong-Hun Jang, Ki-Sang Hong" Binarization of noisy gray-scale character images by thin line modeling," Pattern Recognition 32 (1999) 743-752.

[15] R. M. Bozinovic and S. N. Srihari, "Off-line cursive script word recognition,"IEEE Trans. Pattern Anal. Machine Intell., vol. 11, pp. 68–83, Jan. 1989.

[16] Ding Y, Wakabayashi Tetsushi, Kimura Fumitaka, Miyake Yasuji," Local Slant Estimation and Correction for Handwritten English Word.

[17] S. S.Wang, P. C. Chen, and W. G. Lin, "Invariant pattern recognition by moment Fourier descriptor," Pattern Recognit., vol. 27, pp. 1735–1742, 1994.

[18] X. Zhu, Y. Shi, and S. Wang, "A new algorithm of Connected character image based on Fourier transform," in Proc. 5th Int. Conf. Document Anal. Recognition. Bangalore, India, 1999, pp. 788–791.

[19] S. Connell, "A Comparison of Hidden Markov Model Features for the Recognition of Cursive Handwriting, Master Thesis, Michigan State University 1996.

[20] S. W. Lee and Y.J. Kim, Multi resolutional Recognition of Handwritten Numerals with Wavelet Transform and Multilayer Cluster Neural Network, 3rd International Conference on Document Analysis and Recognition (ICDAR), Canada, 1995.

[21] O.D.Trier and A.K. Jain, Goal Directed Evaluation of Binarization Methods_ IEEE Trans, Pattern recognition and Machine Intelligence vol 17, pp.1191-1201, 1995.

[22] W. Niblack, An Introduction to Digital Image Processing, Prentice Hall, Engle- wood Cliffs, NJ, 1986.
[23] Roy, K. "Word & Character Segmentation for Bangla Handwriting Analysis & Recognition".
[24] Simone Marinai," Introduction to Document Analysis and Recognition " University of Florence Dipartimento di Sistemie Informatica (DSI) Via S. Marta, 3, I-50139, Firenze, Italy
[25] L. O Gorman, "The Document Spectrum for Page Layout Analysis", IEEE Trans. Pattern Analysis and Machine Intelligence, vol.15, pp.162-173, 1993.
[26] S. Randriamasy, L. Vincent "Benchmarking Page Segmentation Algorithms" Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Seattle WA, June 1994.
[27] R. G. Casey, E. Lecolinet, "A Survey of Methods and Strategies in Character Segmentation", IEEE Trans. Pattern Analysis and Machine Intelligence, vol.18, no.7, pp.690-706, 1996.
[28] Ouafae EL Melhaoui Mohamed El Hitmy Fairouz Lekhal "Arabic Numerals Recognition based on an Improved Version of the Loci Characteristic"
[29] A.L Knoll, Experiments with "Characteristics Loci" for Recognition of Hand printed characters.
[30] Wang Jin, Tang Bin-bin, piao Chang-hao, Lei Gai-hui "Statistical method-based evolvable character recognition system" Key Lab. of Network control & Intell. Instrum., Chongqing Univ. of Posts & Commun., Chongqing, China.
[31] K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd edition, Academic Press, 1990.
[32] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, second edition, Wiley Interscience, 2001.
[33] Manish Mangal, Manu Pratap Singh,"Handwritten English Vowels Recognition Using Hybrid Evolutionary Feed-Forward Neural Network".
[34] Velappa Ganapathy, and Kok Leong Liew ,Handwritten Character Recognition Using Multiscale Neural Network Training Technique, World Academy of Science, Engineering and Technology 39 2008.
[35] Dewi Nasien, Habibollah Haron, Siti Sophiayati Yuhaniz," Support Vector Machine (Svm) For English Handwritten Character Recognition" 2010 Second International Conference on Computer Engineering and Applications.
[36] C.M. Bishop, Neural Networks for Pattern Recognition, Claderon Press, Oxford, 1995.
[37] T. Kohonen, The self-organizing map, Proc. IEEE, 78(9): 1464-1480, 1990.
[38] T P Singh, Dr. M P Singh, Somesh Kumar," Performance Analysis of Hopfield Model of Neural Network with Evolutionary Approach for Pattern Recalling".
[39] H. I. Avi-Itzhak, T. A. Diep, and H. Gartland, "High accuracy optical character recognition using neural network with centroid dithering," IEEE Trans. Pattern Anal. Machine Intell., vol. 17, pp. 218–228, Feb. 1995.
[40] I. S. Oh et al., "Class-expert approach to unconstrained handwritten numeral recognition," in Proc. 5th Int. Workshop Frontiers Handwriting Recogniion., Essex, U.K., 1996, pp. 95–102.
[41] L. F. C. Pessoa and P. Maragos, "Neural networks with hybrid morphological/rank/linear nodes: A unifying framework with applications to handwritten character recognition," Pattern Recognit., vol. 33, pp. 945–960, 2000.
[42] Shubhangi D.C, Dr. P .S. Hiremath," Handwritten English character recognition by combining SVM classifier," International Journal of Computer Science and Applications Vol. 2, No. 2, November / December 2009.
[43] Nadira Muda, Nik Kamariah Nik Ismail, Siti Azami Abu Bakar, Jasni Mohamad Zain Fakulti Sistem Komputer & Kejuruteraan Perisian," Optical Character Recognition By Using Template Matching(Alphabet)".
[44] A.D. Bimbo, S. Santin, and J. Sanz, "OCR from poor quality images by deformation of elastic templates," in proceedings of 12th IAPR Int. Conf. pattern Recognition, vol.2, pp.433-435,1994.
[45] M. A. Mohamed, P. Gader, "Handwritten Word Recognition Using Segmentation-Free Hidden Markov Modeling and Segmentation Based Dynamic Programming Techniques", IEEE Trans. Pattern Analysis and Machine Intelligence, vol.18, no.5, pp.548-554, 1996.
[46] M. K. Brown, S. Ganapathy, "Preprocessing Techniques for Cursive Script Word Recognition", Pattern Recognition, vol.16, no.5, 1983.
[47] Mohamed Cheriet, Nawwaf Kharma, Cheng-Lin Liu, Ching Y. Suen, Character Recognition Systems: A Guide for students and Practitioners, (John Wiley & Sons, Inc., Hoboken, New Jersey, 2007).
[48] Dr. Yadana Thein , San Su Su Yee, High Accuracy Myanmar Handwritten Character Recognition using Hybrid approach through MICR and Neural Network ,IJCSI International Journal of Computer Science Issues, 7(6), November 2010.
[49] L. Xu, A. Krzyzak, C.Y. Suen, "Methods of combining Multiple classifiers and their Application to Handwritten Recognition", IEEE Trans. Systems Man and Cybernetics, vol 22, no 3, pp418-435, 1992.
[50] R. M. Bozinovic, S. N. Srihari, "Off-line Cursive Script Word Recognition", IEEE Trans. Pattern Analysis and Machine Intelligence,vol.11, no.1, pp.68-83, 1989.
[51] Rajiv Kumar Nath, Mayuri Rastogi," Improving Various Off-line Techniques used for Handwritten Character Recognition: a Review," International Journal of Computer Applications (0975 – 8887) Volume 49– No.18, July 2012.
[52] S. O. Belkasim, M. Shridhar, M. Ahmadi, "Pattern Recognition with Moment Invariants: A comparative Survey", Pattern Recognition, vol.24, no.12, pp.1117-1138, 1991.

Author Profile

Vijay Laxmi Sahu received the B.E degree in Information Technology From Bhilai Institute of Technology, Durg (C.G) in 2011 and Now she is pursuing M.Tech in Computer Science Engineering from Rungta College of Engineering and Technology, Bhilai (C.G) respectively.

Mrs. Babita Kubde working as Reader in Department of Computer Science & Engineering at Rungta College of Engineering and Technology, Bhilai (C.G).