ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 7, July 2016

Support Vector Machine and Multiclass LDA Based Hand Written Character Recognition System

M.Poongothai,
Department of Information Technology,
Institute of Road and Transport Technology,
Erode, India.
poongothaiirtt@gmail.com

Abstract: The proposed method for hand written character recognition is associated with the reduction of feature vector by Multiclass LDA and recognition using the Support Vector Machine (SVM) classifier. The inherent characteristics or features found within the digital image is extracted and formed as a feature vector. The extracted feature vector is reduced by the dimensionality reduction technique called Multiclass Linear Discriminant Analysis. The reduced feature vector is used to train the SVM classifier. Later test images are given as input and tested the performance of the Classifier. The efficiency of hand written character recognition using SVM Classifier with Multiclass LDA is compared with Back Propagation based Neural Network (BPNN) Classifier. The recognition rates achieved using the proposed method were very good, in comparison with BPNN system.

Keywords: Support Vector Machine, Character Recognition, Multiclass LDA, Back Propagation Neural Network.

I. INTRODUCTION

Character recognition is the technology of using machines to automatically identify human readable symbols, most often alphanumeric characters and then to express their identities in machine readable codes. This is the subclass of object recognition. Character Recognition machines are used mainly for input data to computer systems. A great attempt of research worker in machine learning and data mining has been contrived to achieve efficient approaches for approximation of recognition from data.

One of the challenges in handwritten characters recognition wholly lies in the variation and distortion of handwritten character set because distinct community may use diverse style of handwriting, and control to draw the similar pattern of the characters of their recognized script. Identification of digit from where best discriminating features can be extracted is one of the major tasks in the area of digit recognition system. To locate such regions different kind of region sampling techniques are used in pattern recognition. The challenge in handwritten character recognition is mainly caused by the large variation of individual writing styles. Hence, the robust feature extraction is very important to improve the performance of a handwritten character recognition system. Nowadays handwritten digit recognition has obtained lot of concentration in the area of pattern recognition system sowing to its application in diverse fields.

The main goal in digit recognition is feature extraction is to remove the redundancy from the data and gain a more effective embodiment of the word image through a set of numerical attributes. It deals with extracting most of the essential information from image raw data. Furthermore, characters dataset can be drawn in different sizes and the orientation which are always supposed to be written on a guideline in an upright or downright point. Accordingly, an efficient handwritten recognition system can be developed by considering these limitations. It is quiet exhausting that

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 7, July 2016

sometimes to identify hand written characters as it can be seen that most of the human beings can't even recognize their own written scripts. Hand Written Character Recognition has been the focus of considerable research during the last four decades.
[1]utilizes Support Vector Machine for recognition of Persian Font Recognition. [2] uses the support vector machine with the local features for classifying the leaf images.[3] suggests that the support vector machine performs well in identifying micro parts. [4], uses the moment invariants and Gray level co-variance matrix for the war scene classification. [5][Ronald (2006)] uses the support vector machine for automatic identification of impairments on eye diagram.

This paper discusses a formulation of an Hand Written Character Recognition model in recognizing an object using Support Vector Machine. The features of an image are extracted and the feature vector is constructed. Since the high dimensional data leads to huge processing time, the dimensionality reduction process Multiclass LDA [6] is applied. The reduced vector is applied as an input to the SVM classifier to test image in recognizing the hand written character.
The rest of the paper is organized as follows. Section 2 gives an overview of Multiclass LDA. A summary of Classifiers like Support Vector Machine, and Back Propagation Neural Network are given in Section 3. Section 4 gives an outline of the proposed system. Section 5 illustrates the Experimental results and section 6 for the conclusion.

II. MULTICLASS LDA

Linear discriminant analysis (LDA), normal discriminant analysis (NDA),or discriminant function analysis is a generalization of Fisher's linear discriminant, a method used in statistics and other fields, to find a linear combination of features that characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification.

LDA is closely related to analysis of variance (ANOVA) and regression analysis, which also attempt to express one dependent variable as a linear combination of other features or measurements.[1][2] However, ANOVA uses categorical independent variables and a continuous dependent variable, whereas discriminant analysis has continuous independent variables and a categorical dependent variable (i.e. the class label). Logistic regression and probit regression are more similar to LDA than ANOVA is, as they also explain a categorical variable by the values of continuous independent variables. These other methods are preferable in applications where it is not reasonable to assume that the independent variables are normally distributed, which is a fundamental assumption of the LDA method.

LDA is also closely related to principal component analysis (PCA) and factor analysis in that they both look for linear combinations of variables which best explain the data. LDA explicitly attempts to model the difference between the classes of data. PCA, in contrast, does not take into account any difference in class, and factor analysis builds the feature combinations based on differences rather than similarities. Discriminant analysis is also different from factor analysis in that it is not an interdependence technique: a distinction between independent variables and dependent variables (also called criterion variables) must be made.

LDA works when the measurements made on independent variables for each observation are continuous quantities. When dealing with categorical independent variables, the equivalent technique is discriminant correspondence analysis.

Discriminant analysis is used when groups are known a priori (unlike in cluster analysis). Each case must have a score on one or more quantitative predictor measures, and a score on a group measure. In simple terms, discriminant function analysis is classification - the act of distributing things into groups, classes or categories of the same type.In the case where there are

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 7, July 2016

more than two classes, the analysis used in the derivation of the Fisher discriminant can be extended to find a subspace which appears to contain all of the class variability[7]. This generalization is due to C. R. Rao[8].[1] Suppose that each of C classes has a mean μ_i and the same covariance .

The class separation in a direction $w^{-->}$ the vector is the normal to the discriminant hyperplane) in this case will be given by

$$S = \frac{w^{-->T} \square_b w^{-->}}{w^{-->T} \square w^{-->}} \quad (1)$$

This means that when $w^{-->}$ is an eigenvector of $\square^{\square}\square_b$ the separation will be equal to the corresponding eigenvalue.

If $\square^{\square}\square_b$ is diagonalizable, the variability between features will be contained in the subspace spanned by the eigenvectors corresponding to the $C-1$ largest eigenvalues (since $\square_b$ is of rank $C-1$ at most). These eigenvectors are primarily used in feature reduction, as in PCA. The eigenvectors corresponding to the smaller eigenvalues will tend to be very sensitive to the exact choice of training data, and it is often necessary to use regularisation.

For instance, the classes may be partitioned, and a standard Fisher discriminant or LDA used to classify each partition. A common example of this is "one against the rest" where the points from one class are put in one group, and everything else in the other, and then LDA applied. This will result in C classifiers, whose results are combined. Another common method is pairwise classification, where a new classifier is created for each pair of classes (giving $C(C-1)/2$ classifiers in total), with the individual classifiers combined to produce a final classification.

III. Support Vector Classifier

Support vector machine is a technique often used in pattern recognition and recognition of images. The goal of the support vector machine is to construct a hyperplane as the decision surface in such a way that optimization methods enhance the margin of separation between positive and negative instances. Generally, linear functions in the feature space are used as a separate hyperplane. Several kernel functions, such as polynomial function and radial-bias function are used to achieve better efficiency, and polynomial function is used as kernel function in this article. The scalar product can be implicitly computed in the kernel feature space by using kernel functions.

In this proposed work, the system starts with training sample $\{(x_i, y_i)\}^N$, where the training vector is x_i and its class label is y_i . The proposed method aims to find the optimum weight vector w and the bias b of the separating hyperplane such that [Haykin 999)]

$$y_i(\omega^T\varphi(\xi)+b) \geq 1-\xi_\iota, \quad \forall_\iota \quad (2)$$

$$\xi = \sum_\iota (y_\iota - Y_i) \quad (4)$$

y_i is the prediction value and Y_i is the actual value with $\square$ and x the slack variables(or prediction error) function given below $\square_i$ minimizing the cost

$$\phi(\omega,\xi_\iota) = 1/2(\omega^T.\omega) + C\sum_\iota \xi_\iota \quad (5)$$

Where the slack variables $\square_{\square}$ represent the error measures of data, C is the value assigned to the errors, and $\square\square\square\square\square\square$is a kernel mapping which maps the data into a higher dimensional feature space.

IV. Back Propagation Neural Network

The proper training of a Neural Network is the most important aspect of making a reliable model. *Back-propagation* is the essence of neural net training. It is the practice of fine-tuning the weights of a neural net based on the error rate (i.e. loss) obtained in the previous epoch (i.e. iteration). Proper tuning of the weights ensures lower error rates, making the model reliable by increasing its generalization.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 7, July 2016

In this analysis, the purpose of using the Back-Propagation neural network is to adopt the characteristics of memorizing and referencing properties that recognize the 2D image function of the test. The network input is the function data extracted from the image. And the specified index of the object is the target. When the BPN is educated, The input pattern (x1,x2) is fed to the network while training the BPN, via the hidden layers to the output layer. To find the deviation, the output pattern is compared with the target pattern. These extracted features are continuously fed into the BPN and the network adjusts itself until a set of weights (w_{11}, w_{12}, w_{21}, w_{22}, y_{11} and y_{21}) with an error value has been specified. Then the storage of these weights used for validating new samples.

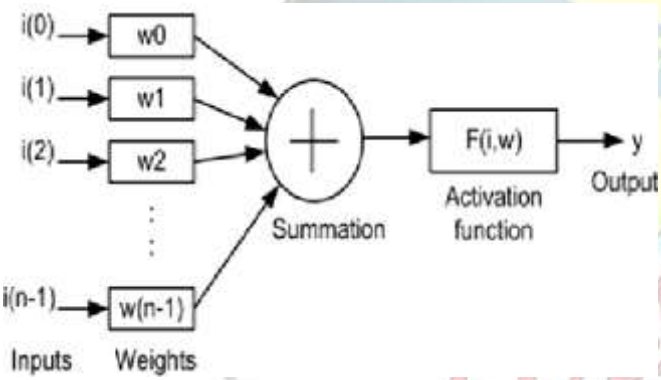

Fig 1. Processing Element in a Neural Network

A nueron output y is ,

$$y = F\left(\sum_i(\omega_i * x_i) + b\right) \quad (6)$$

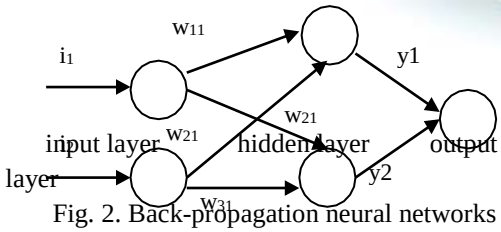

Fig. 2. Back-propagation neural networks

IV. PROPOSED SYSTEM

The proposed system for recognizing the character from an image is given below fig.3. The system is implemented as two stages. During first stage the system is trained with several set of training images. Initially the training images effect of the illumination and lack of contrast on the training images. These are the 4 bit grey scale image. As binary images are easiest to manipulate and occupies less amount of memory space. The indexed image is converted to 2 bit grey scale image where 1's represent white and 0's represent black.

In the Preprocessing, the binary image is subjected to fitted into a matrix of least possible size called framing. Then the framed image is applied to edge detection process. For Edge Detection Sobel Filter is used. After edge detection process, bwskel(Image) function reduces all objects in the 2-D binary image to 1-pixel wide curved lines, without changing the essential structure of the image. This process, called *skeletonization*, extracts the centerline while preserving the topology and Euler number (also known as the Euler characteristic) of the objects.

After preprocessing, the essential features are extracted from the processed image. The choice of features to be derived from the image radically influences the effectiveness of identification.

The set of Features chosen for identification are Short horizontal line(SH), Long horizontal line(LH), Short vertical line(SV), Long vertical line(LV), Left slant(LS), Right slant(RS), Corner(C).

The Feature Vector P = [SH,LH,SV,LV,RS,LS,C] (7)

For instance ,the character A is represented by the feature vector,

P =[1 0 0 0 1 1 2]

The constructed feature vector is high-dimensional. To reduce the high dimension of the data to low dimension without losing the important properties, Multiclass LDA is done. Then Support Vector Machine is used for classification. During testing phase, once the test image is given as input to the proposed system, the pre-processing and feature extraction process are done as

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 7, July 2016

specified in the training phase. The computed feature vector is given as input to the Support Vector Machine Classifier, based on the support vector generated during training phase; the input image is recognized and labeled. To compare the results of Support Vector Classifier, the Back Propagation Neural Network is trained with feature vector and tested.

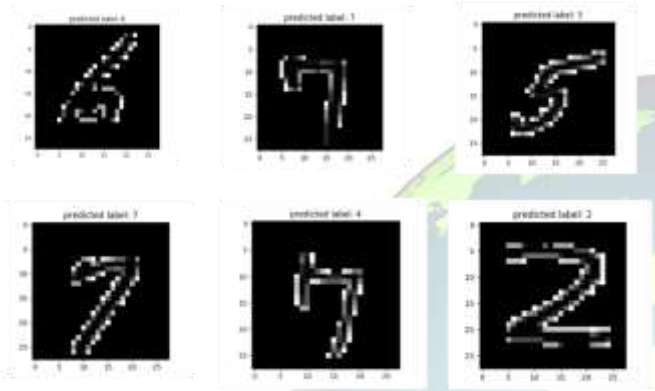

Fig.3. Sample Training Data samples

V. EXPERIMENTAL RESULTS

Machine learning is about learning some properties of a data set and then testing those properties against another data set. A common practice in machine learning is to evaluate an algorithm by splitting a data set into two. We call one of those sets the **training set**, on which we learn some properties; we call the other set the **testing set**, on which we test the learned properties. To experiment the proposed method, sklearn database which is widely used in 2D Character Recognition researches. The data-set is based on gray-scale images of handwritten digits and, each image is 28 pixel in height and 28 pixel in width. Each pixel has a number associated with it, where 0 represents a dark pixel and, 255 represents a white pixel. Both the train and test data-set have 785 columns where, 'label' column represents the handwritten digit and remaining 784 columns represent the (28, 28) pixel values. The train and test data-set contains 60,000 and 10,000 samples respectively. We have used 12000 samples and 5000 samples from the training and test data-sets just to reduce the time of computation and it is recommended to use the full set to obtain a better score and avoid selection bias.. The experiment is conducted to estimate the recognition accuracy of the proposed method.

The proposed system is experimented with each set of training and test images and the results are shown in the table. The proposed method has been implemented using Python. To evaluate the recognition accuracy for the proposed methods and the traditional methods the Prediction ratio (PR) were determined.

The PR is calculated using the formula ,

$$PR = Np/N \quad (8)$$

where Np is the number of characters correctly predicted, N is the total number of images tested.

Table shows the Performance analysis of the proposed and traditional methods in terms of CRP,

Method	PR %
SVM with Multiclass LDA	95
SVM	83
NN(Back propagation algorithm)	81

VI. CONCLUSION

The proposed method for hand written character recognition is associated with the reduction of feature vector by Multiclass LDA and recognition using the Support Vector Machine (SVM) classifier. The reduced feature vector is used to train the SVM classifier. Later test images are given as input and tested the performance of the Classifier. The efficiency of hand written character recognition using SVM Classifier with Multiclass LDA is compared with Back Propagation based Neural Network (BPNN) Classifier. The recognition rates achieved using the proposed method were very good, in comparison with BPNN system.

References:

[1]. Borji, A.and Hamidi, M., "Support Vector Machine for Persian Font Recognition", World Academy of Science, Engineering and

Technology, Vol.28, 2007, pp.10 – 13.

[2]. Krishna Singh, Indra Gupta, and Sangeeta Gupa. "SVM-BDT PNN and Fourier Moment Technique for Classification of Leaf Shape", International Journal of Signal Processing, Image Processing and Pattern Recognition, vol.3,no.4, 2010, pp. 67-78.

[3]. Xin-Han Huang, Xiang-Jin Zeng, Min wang "SVM-Based Identification and Un-calibrated Visual Servoing for Micro- Manipulation", International Journal of Automation and Computing, Vol-7, No-1, 2010, pp.47-54.

[4]. Daniel Madan Raja, S., and Shanmugam, A. "Artificial Neural Network Based War Scene Classification using Invariant Moments and GLCM Features: A Comparative Study", International Journal of Engineering Science and Technology, Vol.3,No.2, 2011, pp. 1189 – 1195.

[5]. Ronald A. Skoog, Thomas C. Banwell, Joel W. Gannett, Sarry F. Habbiby, Marcus Pang, Michael E, Rauch, and Paul Toliver. "Automatic Identification of Impairments Using Support Vector Machine Pattern Classification on Eye Diagrams", IEEE Photonics Technology Letters,Vol-18, No-22, 2006,pp. 2398 – 2400.

[6]. https://en.wikipedia.org/wiki/Linear_discriminant_analysis

[7]. Garson, G. D. (2008). Discriminant function analysis. "Archived copy". *Archived from* the original *on* Rao, R. C. *(1948). "The utilization of multiple measurements in problems of biological classification". Journal of the Royal Statistical Society, Series B.* ***10*** *(2):159–203.* JSTOR 2983775.*2008-03-12*.
Retrieved 2008-03-04.