Offline Signature Recognition and Verification using PCA and Neural Network Approach

Deepti Yadav[1] Dr. Chhavi Saxena[2]
[1]Research Scholar [2]Associate Professor
[1,2]Department of Electronics & Communication Engineering
[1,2]MPCT College, Gwalior, India

Abstract— A Signature of a person is a unique biometric trait which can be used to validate a human identity. As signatures continue to play a vital role in legal, commercial and financial transactions, strictly secured authentication becomes more and more critical. Offline signatures are treated as the most natural means of authenticating a person's identity. A signature by an authorized person is treated as the "seal of approval" and remains the most ideal means of authentication. This paper presents a novel approach for offline signature recognition and verification. Offline signature recognition is implemented using Principal component analysis (PCA) technique while for signature verification efficient back propagation artificial neural network is designed and trained using features extracted from Grey Level Difference Method (GLDM) technique. The system was tested against 100 test signature samples, which comprise genuine and forged signatures of ten individuals giving an average accuracy of 94%.

Key words: Biometrics, Central Moments, GLDM, Neural Network, Offline Signature Recognition and Verification, PCA

I. INTRODUCTION

Biometrics is an automatic scheme of recognizing or identifying a person based on his/her physiological and behavioral characteristics. Various biometric methods have been proposed for personal recognition in the past. Among vision based is iris scanning, face recognition, retina scanning and fingerprint recognition [1] while voice recognition and signature verification comes under non-vision based. Although the selection of an accurate biometric depends to a large extent on a given application, it should be unique, hard to copy, acceptable by the public, and have lower implementation cost.

Signature has been a distinguishing feature for person identification through ages. It is a unique behavioral trait for a person's identity. An important advantage of signature verification compared to other biometric characteristics is its traditional use in many common commercial fields such as legal documents ,E-business, which includes online banking transactions, financial transactions .electronic payments, access control, and so on. Now a day, signatures are preferred over other means of authentication like PIN code, password as they cannot be stolen ,borrowed and even hard to be forgotten. A signature is generally a fusion of exceptional characters and/or the individual's name. It is frequently composed in an extraordinary manner, regularly coming to an incomprehensible state [2]. Systems in this field can be generally classified into either signature verification or recognition systems. These two systems are equally important where authentication is concerned. Verification (Am I whom I claim to be?) involves confirming or rejecting a person's claimed identity. In the Recognition method, one has to establish a person's identity (Who am I?). Each one of these has its own complication and could possibly be solved by a signature authentication system.

Based on the acquisition method used, signature verification systems can be classified into two modes: online (or dynamic) and offline (or static) systems **[3].** In an online mode, users write their signature in digitizing tablet, which acquires the signature in real time and thus get the dynamic information like velocity, acceleration, pressure, position of the signature while in offline mode signatures are scanned from blank paper, and then identifies the signature by analyzing its shape. As in this we get only the 2D image of signature, which make design of its verification system more complex.

The performance of a verification system is evaluated through two important parameters...First one is False Acceptance Ratio (FAR) defined as rate of forgeries accepted as genuine and False Rejection Ratio (FRR) as rate of genuine signatures rejected as forged. For an efficient verification system these errors should be low.

The aim of this paper, is to authenticate a person by correctly recognizing its signature among a list of N signatures present in the database and improving the overall accuracy of classifying a signature as genuine or forged using neural network as a classifier.

II. LITERATURE SURVEY

Harpreet.A et al [3] proposed enhanced signature verification and recognition using matlab. For authentication of signature, the proposed method used geometrical and statistical feature extraction technique which include extracting eccentricity, skewness, kurtosis, mean, standard deviation and many more. These extracted features are then fed to neural network for its training based on which investigation signature is verified.
This technique is found appropriate for diverse applications like passports, bank transactions with good authentication results.

Indrajit.B et al [4] in his paper proposed pixel matching technique for signature verification and recognition. The performance of the projected system has been judge against the existing Artificial Neural Network's (ANN) back-propagation method and Support Vector Machine (SVM) system and outcome shows comparable performance with the advantage of being simple and easy to implement.

Lakshmi.K et al [5] proposed an adaptive machine learning technique known a Multi-Layered Neural Network Model (NN Model) which is trained using certain unique standard statistical features in its feature extraction phase. The performance of the proposed system is then evaluated by calculating the FAR and FRR for a small set of data which come up as 12% and 8% respectively.

Angadi.S et al [6] presented an offline signature recognition system based on local radon features. In this to distinguish different signatures total 16 radon transform based projection characteristics are extracted and then finally back propagation neural network is sketched and trained with 16 extracted characteristics. The trained Neural Network is then further employed for signature recognition and gives an accuracy ranging from 87%-97%.

Nilesh.Y et al [7] proposed an offline signature recognition & verification using back propagation neural network where Invariant Central Moment and Modified Zernike moment methods are used for invariant feature extraction. Preprocessing is done before applying these method for removing unwanted noise present in the signature. The system is firstly trained using record of 56 person's signatures then a mean signature is obtained for each by integrating the features obtained from a set of their genuine sample signatures. For designing this system MATLAB software is used.

Mandeep.K et al [8] proposed signature verification entailing principal component analysis as a feature extractor. In this article, firstly, he reviewed the standard PCA technique and its applications in biometrics data analysis like in signature verification and then introduces several recently proposed PCA-based techniques and concluded that current technique is a step towards improving the current situation in the biometric applications.

Nandi.N et al [9] presented detection and classification of breast cancer using Grey Level Difference Method (GLDM) as a texture feature extracting technique and Support Vector Machine (SVM) as classifier. It was evaluated on 80 images containing malignant and benign masses with different size and shape. Using the SVM classifier and GLDM as feature extractor, he reached upto an accuracy of 92%.

III. PROPOSED METHODOLOGY

The methodology followed for offline signature recognition and verification consists of some fundamental steps which should be tracked with precise rules. The simple block diagram designed for the same is as shown below in Fig. 1 which presents the brief idea about it.

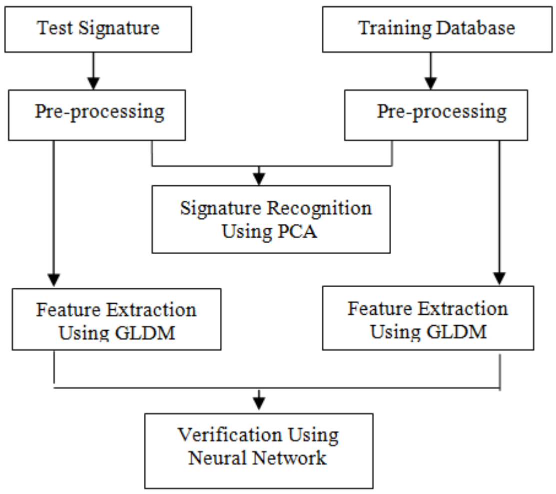

Fig. 1: Block Diagram of Proposed System

The main steps followed are described as follows:

A. *Data Acquisition:*

The SVC2000 handwritten signature database is used in experiments. Some of the signature templates from the database are as depicted in Fig 2.

S. No.	Genuine	Forged
1.		
2.		
3.		
4.		
5.		

Fig. 2: Signature Template

The database consists of 100 signatures of 10 persons i.e. 10 signatures from each user with 5 genuine and 5 forgery signatures respectively.

B. *Pre-Processing:*

This phase is included to eliminate any noise if they get induced during the acquisition stage. This step is applied to both the phases i.e. training as well as testing phase. Following pre-processing steps are followed:

1) *Binarization:*

In this acquired image is converted into a binary image consisting of only 1's and 0's. This allow to process signature based on the matrix having only zeroes and ones.

2) *Resizing:*

This is done so as to bring all signatures present in the database to the same and secure size.

3) *Filling Holes:*

In this closed region of an image are filled with white pixel to get the proper geometric shape of the signature

4) *Thinning:*

Thinning is performed to make extracted features unaffected to image characteristics like quality of pen and paper. It reduces the binary objects to strokes that are single pixel wide.

The result of above mentioned preprocessing steps are depicted in Fig 3.

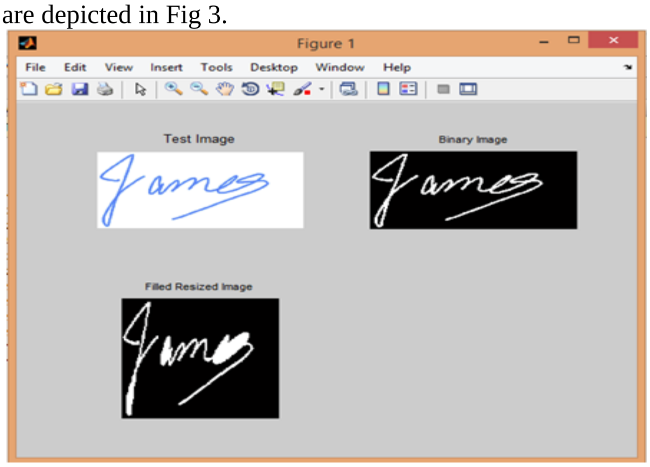

Fig. 3: Various Pre-Processing Steps

C. Feature Extraction:

The performance of signature verification system mainly depends on the selection of resourceful feature extraction technique. In short, the goal of feature extraction is to find preferably petite number of features that are particularly distinguishing for the classification process, and that are invariant to irrelevant transformations of the data. For the same the proposed system has used Gray level Difference Method (GLDM) and invariant central moments.

1) Gray level Difference Method (GLDM):

It is an efficient statistical approach for texture analysis which estimates the Probability Density Functions for a given image. Let I(x, y) be image intensity, then for given displacement vector d=(Δx , Δy) let $I_d(x,y)= |I(x,y)-I(x+\Delta x,y+\Delta y)|$ and f(i|d) be the corresponding probability density function for $I_d(x,y)$. The value of probability density function f(i|d) is estimated from the number of times change in intensity $I_d(x,y)$, occurs for a given displacement d, i.e f(i|d) = P(I_d(x, y) = i). If a texture is directional, the degree of spread of the values in f(i|d) will vary with the direction of d. Thus, texture directionality can be analyzed by comparing spread measures of f(i|d) for various directions of d [9]. In the present study, four possible forms of the vector d were considered :(0, d), (d, 0), (-d, d), and (-d, -d), with d being the inter pixel distance and then the feature vectors are derived from the following five features.

2) Contrast:

Used to compute local gray levels variations present in an image.

$$\text{Contrast} = \sum i^2 P(i)$$

3) Homogeneity:

Used to compute the neighbouring homogeneity of an image. It has opposite behaviour to that of contrast:

$$\text{Homogeneity} = \sum P(i)/(1+i^2)$$

4) Energy:

It reveals pixel-pair repetitions. Very less gray level transitions take place in homogeneous images, which result into higher energy.

$$\text{Energy} = \sum P(i)^2$$

5) Entropy:

The feature entropy is a measure of non-uniformity in the image or region of interest.

$$\text{Entropy} = -\sum P(i)\log(P(i))$$

6) Mean:

The mean is determined by the homogenous brightness or darkness of the image

$$\text{Mean} = \sum i\, P(i)$$

7) Invariant Central Moment:

The geometrical central moments of order equal to (p+q) for a two-dimensional discrete function like image **[7]** is computed by

$$\mu_{pq} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}(x-\bar{x})^p(y-\bar{y})^q f(x,y)$$

Where f(x, y) is image function, M and N are image dimensions and $\bar{x}$ and $\bar{y}$ are center of gravity of image and are calculated as

$$\bar{x}=\frac{m_{10}}{m_{00}},\quad \bar{y}=\frac{m_{01}}{m_{00}}$$

Finally, these translation normalized central moments, are further normalised for the effects of change of scale which makes central moments as translation and scale invariant, by applying

$$\eta_{p,q} = \frac{\mu_{p,q}}{(\mu_{p,q})^k}$$

Where

$$k = 1 + \frac{(p+q)}{2} for\ p+q \geq 2$$

D. Signature Recognition:

For Signature recognition Principal component Analysis (PCA) technique is utilized. It identifies patterns in data, and expresses it in a way to highlight their similarities and differences. In PCA, it is assumed that the information is carried in the variance of the features, which means the features which exhibit higher variance carry more information than others. Hence, PCA employs a linear transformation that is based on preserving the most variance in the data using the least number of dimensions.

The following steps are followed for signature recognition using principal component analysis (PCA) technique.

1) Acquire Data:

Firstly acquire an initial set of N training signature images represented by I1, I2,…., IN . Every signature Ii is represented to as a vector Fi.

2) Subtract the Mean:

For PCA to work properly, we estimate deviation for every signature image from the mean (average) signature vector given by φ. The mean subtracted is average across each dimension. The equation is:

$$\emptyset i = \text{Fi} - \varphi$$

Where $\varphi = \frac{1}{N}\sum_{i=1}^{N} \text{Fi}$ (Mean signature vector)

3) Calculate Covariance Matrix:

A set of deviation vector of N images given by A is first determined

$$A= [\varphi1, \varphi2, \varphi3,\dots\dots \varphi n]$$

Where, φi is the deviation vector for ith image, which will then be used for forming covariance matrix given by D.

$$D = A\, A^T$$

Here D is an MxM matrix and A is an MxN matrix.

4) Calculate the Eigenvectors and Eigenvalues:

Calculate the eigen values and eigen vectors of the covariance matrix. These values provide us with information about the patterns in data. So, by taking the Eigen vectors of the covariance matrix, we are able to extract the lines that characterize data. While computing the Eigenvectors of covariance matrix, it would return N Eigenvectors, each of dimensions Nx1. Once eigenvectors are found, they are ordered according to their eigenvalue, highest to lowest, giving the components in order of their significance. The eigenvector with the highest eigenvalue is the principle component of the data set.

5) Forming a feature vector:

By ignoring the components of lesser significance we form a feature vector with the selected eigenvectors in the columns.

Feature Vector = (eig1, eig2,eig3........eign)

6) Signature Recognition:

For test signature recognition all the above 5 steps are followed for test signature also. Then finally we recognize the test signature by finding Euclidean distances between the projected test image and the projection of all centered training images on the eigen space, with matched signature be the one with minimum euclidean distance as shown in Fig 4.

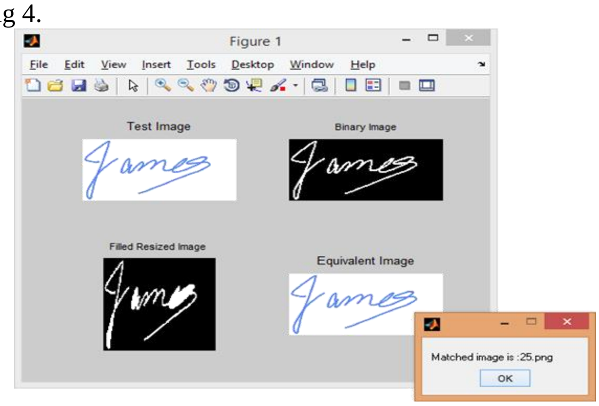

Fig. 4: Matched Signature Image

E. Verification:

For Signature verification, a neural network is designed and implemented using MATLAB 12a with Neural Network Toolbox. Neural Network is a model which is designed by examining neural network processing in biological system. Various algorithms can be used to create neural network, but Back propagation is chosen as it is easiest to implement, while maintaining efficiency of the network.

This network use to categorize signature according to feature vector characteristic. It is formed by generalizing the gradient descent with momentum weight and bias learning rule. Since it is a supervised training neural network method, both input and its expected output is given for training. Neural network is trained until it classify all the described pattern. Weights are changed according to training algorithms which compute the gradient of the performance function to minimize performance. This gradient is resolved using a method called back propagation, which engage performing computations backwards through the network. The network consists of 3 layers i.e. input layer, hidden layer and an output layer.

In the proposed system, neural network is designed as shown in Fig 5 using nntool for simulations using the following specifications.

1) No. of Neural network layers: 3
2) Activation function of hidden layer = Sigmoidal
3) No. of inputs, n = 12
4) No. of neuron in hidden layer, m = 30
5) No. of output = 1(b/w 1 and 0)

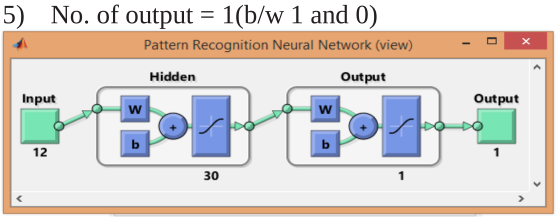

Fig. 5 : Feed Forward Back Propagation NN

IV. RESULT ANALYSIS & COMPARISON

In this study, off-Line Signature Recognition is presented using PCA technique which shows high quality results by correctly recognizing the inputted test signatures form SVC2000 database consisting of 100 signature of 10 person. While for classifying a signature as genuine or forged the proposed system has designed a feed forward back propagation neural network using Scaled Conjugate Gradient (trainscg) as its training algorithm.

The proposed network was trained with feature vector of data cases. As the training process is completed, the last weights of the network were saved to be equipped for the testing procedure. The results for the same are analyzed by different plots as explained below.

A. Receiver Operator Characteristic Measure (ROC) Plot:

It is a plot between the true positive rate (sensitivity) and the false positive rate (1 -specificity) as the threshold is varied. A perfect test result would show points in the upper-left corner, with 100% sensitivity and 100% specificity.The network performs well for this problem as depicted in Fig 6.

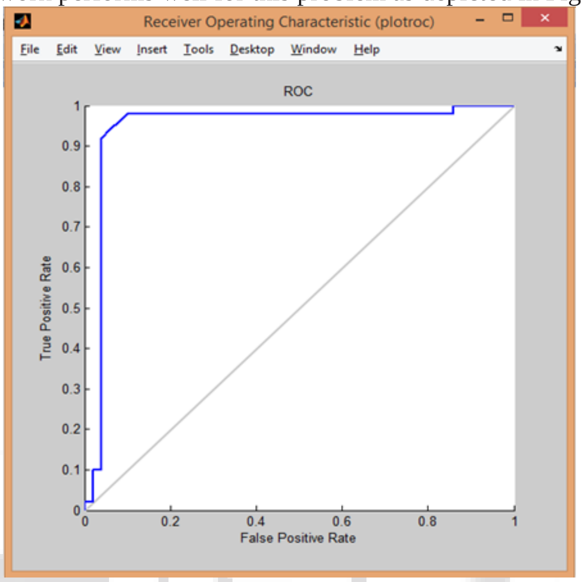

Fig. 6: ROC Curve

B. Training State Plot:

The plot as depicted in Fig 7 represents different training states in training process along with validation check graph.

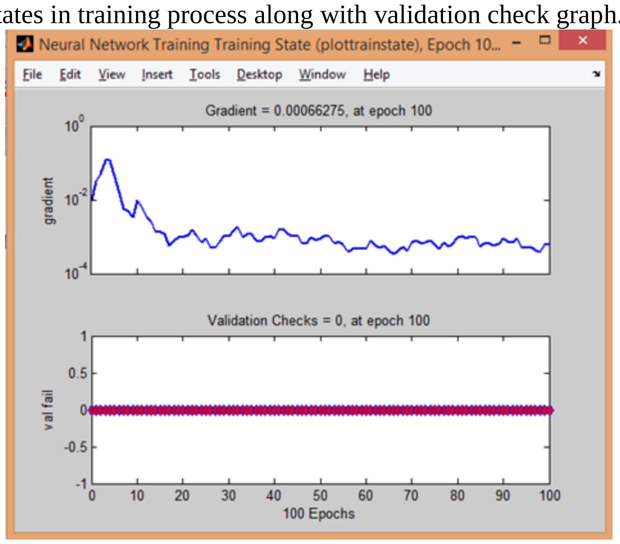

Fig. 7: Training State Plot

C. Confusion Matrix:

A confusion matrix as depicted in Fig 8 contains information about actual and predicted classification done by classifier.

This matrix contains essential information about accepted positive (TP= 46), accepted negative (TN = 48), false positive (FP = 2) and false negative (FN = 4) which will then be used to compute important performance parameters as given below

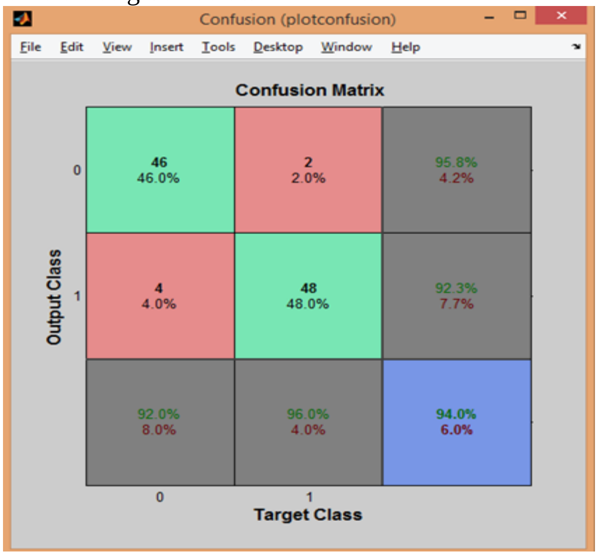

Fig. 8: Confusion Matrix

1) $FAR = FP/(TN+FP) = 0.04$
2) $FRR = FN/(TP + FN) = 0.08$
3) $AER = (FAR + FRR)/2 = 0.06$
4) $ACC = (TP+ TN)/(P+N) = 0.94$

The proposed system performs very well as can be seen from lower values of error rates(FAR & FRR) and with good accuracy rate reaching upto 94% as can also be seen in lower right blue squrare(in green color) of confusion matrix which illustrate the overall accuracy

A result comparison study has been shown in Table 1 in order to prove that our proposed system shows better performance than some currently existing methods.

Feature Extraction Tech.	FAR%	FRR%	AER%	Accuracy
Proposed (GLDM)	4%	8%	6%	94%
Geometric & Statistical Tech[3]	-	-	-	85 %
PMT [4]	6%	12%	9%	-
Unique Statisical Features [5]	12%	8%	10%	-

Table 1: Result Comparison

V. CONCLUSION & FUTURE WORK

Signatures are extensively used as a means of personal verification, which highlights the requirement for an automatic verification system. The proposed system has used PCA technique for signature recognition which shows high-quality results by correctly recognizing inputted test signature to the system and back propagation artificial neural network as classifier to verify a signature as genuine or forged. It has been observed that essential features extracted using GLDM technique along with central moments are found to be efficient for signature verification achieving lower error rates and improved accuracy upto 94% for enrollment of 10 persons.

In this research, we work on offline system, so we work with only static images. As it is difficult to find out dynamic properties from it. But further research work can proceed to find out some dynamic features from these static images. Along with this, the proposed work can be combined with other existing biometric (like face, handwritten geometry) based verification system, for designing more accurate and reliable authentication system.

REFERENCES

[1] Upasana Dewan, Javed Ashraf,''Offline Signature Verification Using Neural Network'' International Journal of Computational Engineering & Management, Vol. 15 Issue 4, July 2012

[2] Mujahed Jarad, Dr. Nijad Al-Najdawi, Dr. Sara Tedmori, "Offline Handwritten Signature Verification System Using a Supervised Neural Network Approach," 2014 IEEE 6th International Conference on CSIT, pp 189-195

[3] Harpreet Anand, D.L Bhombe," Enhanced Signature Verification And Recognition Using Matlab" 2014 International Journal of Innovative Research in Advanced Engineering, pp 88-94

[4] Indrajit Bhattacharya,Prabir Ghosh, Swarup Biswas,'' Offline Signature Verification Using Pixel Matching Technique'' Procedia Technology 10 (2013) , pp 970 – 977

[5] K. V. Lakshmi, Seema Nayak,"Off-line Signature Verification Using Neural Network", 2013 3rd IEEE International Advance Computing Conference(IACC), pp 1065-1069

[6] S. A. Angadi,Smita Gour,Gayatri Bajantri," Offline Signature Recognition System Using Radon Transform",2014 Fifth International Conference on Signals and Image Processing,pg 56-61.

[7] Nilesh Y. Choudhary,Mrs. Rupal Patil, Dr. Umesh Bhadade, Prof. Bhupendra M Chaudhari,'' Signature Recognition & Verification System Using Back Propagation Neural Network'' Vol 2, No. 1,Jan 2013

[8] Er. Mandeep Kaur,Dr. A.K Sharma,''Signature Verification Entailing Principal Component Analysis As A Feature Extractor'', International Journal of Advanced Technology & Engineering Research (IJATER) ISSN No: 2250-3536,pp 28 -32

[9] Nandi Nwe Win, Nang Aye Aye Htwe,''Detection and Classification of Breast Cancer", International Journal of Scietific Engineering and Technology Research Vol.03,Issue.08, May-2014