International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 10 (2018) pp. 252-260

Multiple Classifier System for Writer Independent Offline Handwritten Signature Verification using Hybrid Features

Ashok Kumar[1] and Karamjit Bhatia[2]

[1] Research Scholar, Department of Computer Science, Gurukula Kangri Vishwavidyalaya
Haridwar Roorkee Highway, Haridwar, India
& Associate Professor, Invertis University, Lucknow Delhi Highway, Bareilly, India
Email: ashok_gangwar@rediffmail.com

[2] Professor, Department of Computer Science, Gurukula Kangri Vishwavidyalaya
Haridwar Roorkee Highway, Haridwar, India
Email: karamjitbhatia@yahoo.co.in

***Abstract*: Offline handwritten signature verification is a very challenging area of research as the handwriting of two people may bear similarity whereas handwriting of a person may vary at different times. The accuracy of handwritten signature verification system depends on the classifier system and the way of feature extraction. Keeping this point of view, four types of hybrid feature sets and three types of classifiers specifically support vector machine with polynomial kernel, support vector machine with quadratic kernel and decision tree are investigated for writer-independent offline handwritten signature verification in the present work. To obtain hybrid feature sets, local oriented statistical information booster, discrete wavelet transform, and histogram of oriented gradient feature descriptors are extracted and are coupled with each other. To create multiple classifier system, the training set is partitioned into subsets and these training subsets are used to train the classifiers of multiple classifier system using same training algorithm for all classifiers. The performance analysis is carried out using two scenarios. In the first scenario, genuine and random forgery signatures are used to train the classifiers whereas genuine, random, unskilled and simulated forgery signatures are used to train the classifiers in the second scenario. False rejection rate 8.00 and false acceptance rate 0.00 for all types of forgeries are reported as the best result of the experiments.**

Keywords: Writer-Independent Offline Signature Verification System, Local Oriented Statistical Information Booster Features, Histogram of Oriented Gradient Features, Discrete Wavelet Transform Features, Support Vector Machine, Decision Tree, Multiple Classifier System.

I. Introduction

The handwritten signature of a human being is a biometric characteristic. Handwritten signatures are most widely used for verification of a human being as well as the authenticity of financial, legal and official documents. The Handwritten Signature Verification (HSV) framework validates the signature of an individual as genuine or forges. In preceding few decades, numerous offline as well as online signature verification systems, have been investigated by the researchers. In online approach, an optical pen is utilized by the writer for signature and sensors are utilized to pick dynamic features of handwriting like the speed of writing, the order of strokes, and pressure at various positions of the signature etc. In offline approach, the paper sheet is utilized to collect the signature of writers and optical scanner is used to change over the signatures into digital form [1]. Due to inaccessibility of dynamic features, the development of competent offline HSV systems is a hard task. For developing HSV system the forgeries set is generally divided into three forgery subsets namely- random, unskilled, and simulated. The genuine signature of a different writer is considered as a random forgery for a genuine writer. In unskilled (also called simple) forgery creation process, the genuine writer's name is known to the forger whereas the forger knows the genuine signature of the writer very well and has practiced the signature many times to create simulated (also called skilled) forgery [2]. The genuine signature of a writer and its corresponding random, unskilled and simulated forgeries are illustrated in Figure 1.

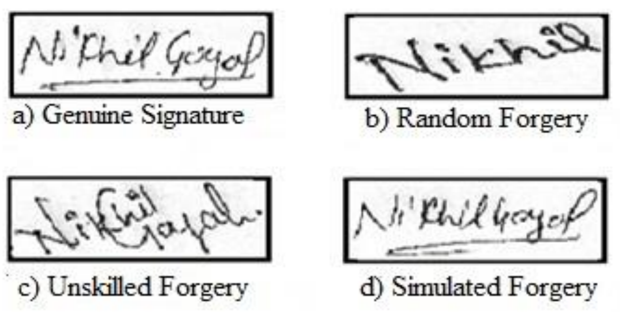

Figure 1. Genuine and forgery signatures

In HSV framework, False Rejection Rate (FRR) and False Acceptance Rate (FAR) are two performance metrics which are generally utilized to assess the HSV system performance. The percentage of genuine signatures of writer acknowledged as forgery signature by the system is known as FRR whereas FAR is calculated as the percentage of forgery signatures of writer acknowledged as a genuine signature [2]. In literature, another term called Average Error

Rate (AER) or Mean Error Rate (MER) which is the average of FRR and FAR is also reported.

Writer-Dependent (WD), as well as Writer-Independent (WI), approaches are used to develop offline HSV framework [3]. In WD approach, a personal model is built for every writer on the basis of two dissimilar classes, Class1 and Class2, where genuine signature samples of a specific writer constitute Class1 whereas Class2 consists of forgery signature samples. The WD approach suffers from two major drawbacks, first, it requires to include a vast number of genuine samples and second, its incapability to absorb a new writer without generating a new personal model for the writer. Then again, WI approach (also called global model) requires a single model to manage all writers and is proficient enough to absorb unknown individual without retraining the model. The improvement of WI approach is that one can build reliable model even when few number of genuine signature samples are available.

In WI approach, the feature vector of the questioned signature is compared with feature vectors of reference signatures to classify the questioned signature. To perform classification process, the dissimilarity between the feature vector of questioned signature Q and the feature vector of reference signatures Ref k (k = 1, 2, …,N) is computed. The dissimilarity representation concept is introduced by Pekalska et al. [4] and is based on the idea that dissimilarities among the same class objects are less as compared to those among objects belonging to different classes. The difference between the feature vector of reference signature and feature vector of the questioned signature is obtained as diff = | fref - fq | to create dissimilarity vector. Dissimilarity feature vector is fed to the classifier to get the partial decision. Finally, fusion strategy is utilized to combine the partial decisions to get the final decision as illustrated in Figure 2.

The present study aims at developing the Multiple Classifier System (MCS) for writer-independent offline HSV system with reduced average error rate using hybrid features. Hybrid feature sets are obtained by combining the extracted feature descriptors of Histogram of Oriented Gradients (HOG), Discrete Wavelet Transform (DWT), and Local Oriented Statistical Information Booster (LOSIB). The signature database of 260 writers is used to develop the system. Support Vector Machine with Polynomial Kernel (SVM – POLY), Support Vector Machine with Quadratic Kernel (SVM – QUAD) and Decision Tree (DT) classifiers are used to generate the MCS. The training set is partitioned into k partitions using k-fold cross-validation technique to create MCS of k diverse classifiers and same training algorithm is utilized to train all classifiers of MCS. The classifiers of MCS are trained by using two scenarios namely-Scenario-I and Scenario-II. In Scenario-I, genuine and random forgery signatures are used to train the classifiers of MCS whereas genuine, random, unskilled and simulated forgery signatures are used to train the classifiers of MCS in Scenario-II. Unskilled and simulated forgeries are used in both scenarios to test the classification accuracy of the developed system. These multiple classifier systems aim at classifying the handwritten signature of writers as genuine or forged.

The rest of the paper is prepared as follows: literature review related to writer-independent offline HSV systems is presented in Section II. The motivation and objectives of the proposed study are described in Section III. Section IV presents different techniques to design multiple classifier system. The research methodology used in this study is presented in Section V. The result of experiments and discussion are given in Section VI and the conclusion of this paper is given in Section VII.

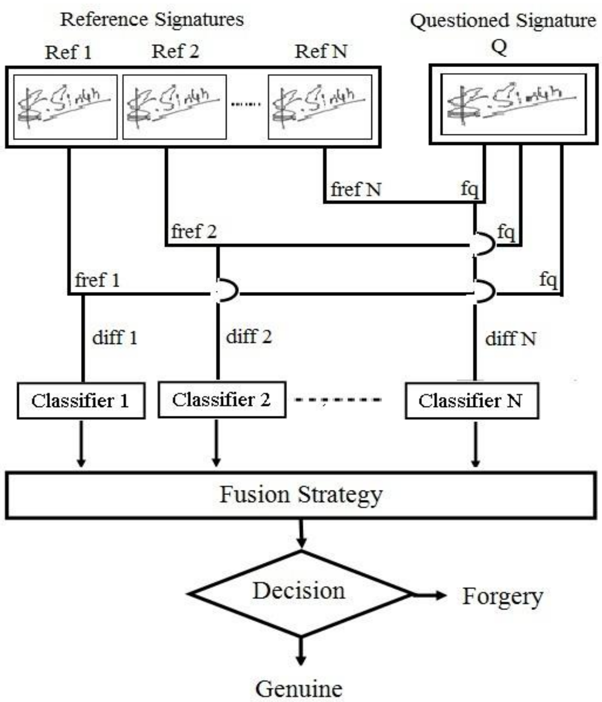

Figure 2. Writer-independent approach for offline signature verification system

II. Literature Review

A HSV system using the writer-independent approach is not a widely addressed research problem as compared to the writer-dependent approach. The writer-independent approach was proposed by C. Santos et al. [4]. The researchers used this approach with Neural Network (NN) and graphometric features to develop WI offline HSV system and claimed AER of 8.02. D. Bertolini et al. [3] improved the classification accuracy of WI offline HSV system through a pool of SVM classifiers and graphometric features. The authors claimed AER of 6.28 through their experiment. D. Rivard et al. [5] utilized two grid based techniques specifically Directional Probability Density Function (DPDF) and Extended Shadow Code (ESC) to extract the features from the signature image and acquired AER of 5.19 through SVM classifier. An approach based on surroundedness belongings governed features and two classifiers, namely- NN and SVM to broaden the WI offline HSV system is proposed by R. Kumar et al. [6]. The authors claimed classification accuracy of 86.24. The spatial distribution and additional orientation of stroke features are utilized by S. Eskander et al. [7] to develop the WI offline HSV system. The authors claimed AER of 5.38 using SVM classifier. To develop the WI offline HSV system, J.

Swanepoel et al. [8] proposed Dynamic Time Wrapping (DTW) and Discrete Radon Transform (DRT) features and claimed AER of 4.93. G. Eskender et al. [9] claimed a more secure, accurate and less complex offline HSV system using the SVM classifier and reported AER of 7.24. Writer-independent system for signature verification with lesser number of references against questioned signature is reported by A. Hamadene et al. [10]. The researchers utilized Contourlet Transform (CT) and Directional Code Co-event Matrix (DCCM) in their approach and acquired AER of 18.42 by utilizing one class SVM classifier. L. Hafeman et al. [11] used SVM – RBF classifier and obtained AER of 3.96 through experiments.

III. Motivation and Objectives

Literature review reveals that various approaches have been proposed by researchers for WI offline HSV but the obtained classification rate of WI offline HSV system is not satisfactory. Further, due to unavailability of dynamic features, the development of competent and consistent offline HSV system is supposed to be a hard task as compared to the development of online HSV system. The complexity of handwritten signature makes it more difficult to achieve the higher classification rate for offline handwritten signatures. Thus, in the framework of WI offline handwritten signature verification, there is still a huge scope of research work to propose techniques for obtaining classification error rate as close as possible to 0.00. Therefore, there is a need for competent techniques to develop WI offline HSV system.

The primary objective of this planned research effort is to offer a competent approach for feature extraction and creation of multiple classifier system to reduce the FAR for unskilled and simulated forgeries. Two primary objectives of this work are: (1) the approach will absorb well the handwritten signature of unknown writers without retraining the model (2) the approach will reduce FAR for unskilled and simulated forgeries and Average Error Rate.

IV. Multiple Classifier System Design Methods

The benefits of multiple classifier system over the single classifier system have been investigated in the studies of several researchers [12]. There are two effective methods to create the multiple classifier systems. In the first method, a different type of classifiers such as SVM – POLY, SVM – QUAD, DT, neural network, etc. are trained using same training set and combined to create the MCS. In the other method, the training set is divided into different subsets by using sampling and these subsets are used to train the classifiers of MCS. Subsample and subspace are two acceptable sampling methods. Bagging [13] and Boosting [14] are two methods to create the MCS belonging to the sub sampling method whereas Random Subspace Method (RSM) [15] belongs to the subspace method.

A. Bagging

In bagging method, classifiers of MCS are produced autonomously. In this method, training sets are formed from the input training set with replacement by the sampling. The dimension of the formed training sets is identical to the dimension of the input training set. Consequently, some feature vectors may not emerge in the replications while other feature vectors may appear more than once. Such training set generating process using replication is known as the bootstrap aggregating. There also exist sampling algorithms without replacement, such as k-fold cross-validation [16]. In k-fold cross-validation, k training sets are obtained by partitioning the input training set into k partitions randomly. In this manner, k diverse classifiers are generated using k training sets obtained from the input training set. At last, all the generated classifiers are pooled to create the MCS.

B. Boosting

In boosting method, classifiers of MCS are produced successively. An arrangement of weights over the training set is kept up in this technique. At first, all weights are same and input training set is utilized to train the first classifier in the sequence. After this, weights are altered based on the performance of present classifier and re-weighted training set is utilized to train the next classifier in the sequence. In this manner, the classifiers of MCS are trained.

C. Random subspace method

Random subspace method is a sampling method which samples the feature space rather than the training set. In RSM, x-dimensional feature subspace is obtained by choosing x features arbitrarily from the y-dimensional feature space ($x < y$) [15]. At that point, the classifiers of MCS are produced by utilizing the x-dimensional feature vectors. At last, all the produced classifiers are joined to make the MCS.

D. Test and select method

In test and select method, diversity is integrated into classifiers of MCS by over-producing classifiers and then selecting some of them to create the MCS. Overproduction of classifiers is the beginning phase of this method [17]. Two test sets, namely- validation and final test sets are required in this method. To select the best classifiers from the pool of overproduced classifiers, the validation test set is used. After this, the selected classifiers of MCS are tested using a final test set and the performance of MCS is reported.

V. Research Methodology

The major steps used to develop the MCS for WI offline HSV system in this work are: creation of a signature database, feature extraction, the creation of a dissimilarity feature vector set, creation of MCS and classification of signature of the writer as genuine or forge through MCS.

A. Creation of signature database

To develop the writer-independent offline HSV system, 260 writers signature database is created in the present work. The signature database incorporates genuine, unskilled forgery and simulated forgery signature samples of writers. In training and testing phase, signatures of 160 writers and 100 writers, respectively are used. To create the signature database, the signatures of undergraduate and postgraduate students of an institute are collected using the A4 white paper sheet. After this, the signature samples are converted into digital form using the scanner at 600 dpi gray levels. Each student signed 20 genuine signatures. For each genuine writer, 4 forgers are chosen for making unskilled and simulated forgery signature samples. Each forger is assigned a task to sign 5 unskilled as well as 5 simulated forgery signatures. In this manner, total 20 unskilled and 20 simulated forgery signature samples per genuine writer are collected. The name of the writer is known to the forgers to create the unskilled forgeries whereas forgers have practiced with genuine signatures of the writer many times to produce the simulated forgeries.

B. *Feature extraction*

In this study, preprocessed genuine and forgery signature images are utilized to extract the feature descriptors. At preprocessing point, the median filter is utilized to eliminate the noise from genuine and forgery signature images. After this, a gray level signature image is changed into a binary image by calculating threshold value using Otsu's method [18]. The signature image is then cropped and resized to the image size 256 x 512.

Hybrid features are utilized to develop the MCS for WI offline HSV system. Reason to do so is that the hybrid features are extracted to improve the system performance as they combine the efficiencies of individual feature sets. The HOG, DWT, and LOSIB feature descriptors are utilized to form hybrid feature sets. The procedure of feature extraction of HOG, DWT, and LOSIB is presented in the Sub-Section 1, 2 and 3, respectively.

1) *Histogram of oriented gradients feature descriptor*

For present work, to extract HOG features, an approach introduced by Navneed and Bill Trigs is adopted [19]. Following steps are utilized to obtain the HOG feature descriptor.

1. *Gradient computation:* To calculate the gradient directions, the signature image is partitioned into cells and then gradient direction for all pixels of each cell of the signature image is computed.
   In gradient computation, one dimensional mask $M_x = [-1 \ \ 0 \ \ 1]$ is used in horizontal direction and $M_y = -M_x'$ is used in the vertical direction. For a given image IM, the derivatives D_x and D_y are obtained by equation (1) and (2), respectively. The magnitude of gradient $|G|$ and orientation of gradient Θ are obtained using equation (3) and (4), respectively.

$$D_x = IM \times M_x \tag{1}$$

$$D_y = IM \times M_y \tag{2}$$

$$|G| = \sqrt{{D_x}^2 + {D_y}^2} \tag{3}$$

$$\Theta = \arctan\left(\frac{D_y}{D_x}\right) \tag{4}$$

2. *Orientation binning:* Magnitude value of each pixel in the cell is assigned into one of nine bins (the range of 20 degrees per bin) equally spread over 0^0 to 180^0. Cell histograms are created in the orientation binning process.
3. *Formation of HOG feature set:* Features are extracted from each cell and all the cell features are combined to obtain the final HOG feature set.

In this manner, the HOG feature set of length 81 is obtained.

2) *Discrete wavelet transform feature descriptor*

In this study, Haar wavelet is used to compute the wavelet coefficients. The approximation coefficients (low pass values), and detailed coefficients (high pass values) of horizontal, diagonal and vertical directions are used to extract the features from the signature image [20]. Following steps are accomplished to find the DWT feature descriptor in the present study:

1.a. The maximum value of approximation coefficients and detailed coefficients is computed. In this manner, 4 features are obtained.

b. After this, approximation coefficients are decomposed into further two levels. The maximum value of approximation coefficients and detailed coefficients for each level is found. In this way, 8 more features (4 features from each level) are obtained from the whole signature image.

Thus, 12 features are obtained from the whole image of the signature in step 1. Three level decomposition of a signature image is illustrated in Figure 3.

2. The signature image is segmented into four equal cells namely - upper left cell, upper right cell, lower left cell, and lower right cell. Step 1 is repeated for each cell to extract the 12 features from each cell.

Thus, 48 features are extracted from the four cells of the whole signature image.

In this manner, discrete wavelet transform feature descriptor of length 60 is obtained.

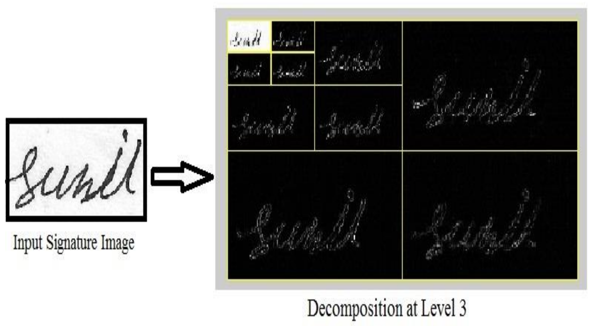

Figure 3. Three level decomposition of signature image

3) *Local oriented statistical information booster feature descriptor*

The LOSIB feature descriptor is extracted from a signature image by using the approach introduced by Oscar Garcıa Olalla et al. [21]. To find the LOSIB feature descriptor, following steps are used:

1. *The location of neighbor pixels of a neighborhood of given radius with respect to the central pixel of the neighborhood is obtained.*

   To do this, let R is the radius of the neighborhood and N is the number of pixels in the neighborhood. Suppose that central pixel q is at position (X_q, Y_q) and r^{th} neighbor pixel of the neighborhood is at position (X_r, Y_r). For a given pixel q, the coordinate (X_r, Y_r) of its r^{th} neighbor is obtained by equation (5).

$$(X_r, Y_r) = \left(X_q + R \times \cos\frac{2\pi r}{N},\ Y_q - R \times \sin\frac{2\pi r}{N}\right) \quad (5)$$

2. *The absolute difference of pixel gray-level value between a central pixel and neighbor pixels of the neighborhood is taken.*

   To do this, suppose that S_q and S_r are the gray levels of q, the central pixel, and r^{th} neighbor pixel, respectively. Then the absolute difference D_r between the pixel gray-level values is obtained by using equation (6).

$$D_r(X_q, Y_q) = |S_q - S_r| \quad (6)$$

3. *The mean of all absolute differences along the same orientation is computed.*

   The mean of all absolute differences along the same orientation is computed by the means of equation (7).

$$MD_r = \frac{\sum_{Xq=1}^{C} \sum_{Yq=1}^{R} Dr(Xq,Yq)}{C*R} \quad (7)$$

   where, C and R are the number of columns and rows of the signature image, respectively.

4. *The mean values for all orientations are used to obtain the LOSIB feature descriptor.*

Thus, by using radius 1, 8 features are obtained from 8 pixels in the neighborhood. Similarly, 16 features are obtained from 16 pixels in the neighborhood by using radius 2 and 24 features are obtained from 24 pixels in the neighborhood by using radius 3. In this manner, the LOSIB feature descriptor of length 48 is obtained.

To form the hybrid feature sets, the feature vector of HOG, DWT, and LOSIB feature sets are joined with each other. Consequently, four hybrid feature sets specifically H1 (HOG feature descriptor plus DWT feature descriptor), H2 (HOG feature descriptor plus LOSIB feature descriptor), H3 (DWT feature descriptor plus LOSIB feature descriptor), and H4 (HOG feature descriptor plus DWT feature descriptor plus LOSIB feature descriptor) are obtained by joining the HOG, DWT, and LOSIB feature vectors with each other. Furthermore, the length of the HOG, DWT, and LOSIB feature vector is 81, 60, and 48, respectively. Therefore, the hybrid feature vectors specifically H1of length 141, H2 of length 129, H3 of length 108, and H4 of length 189 are obtained.

C. Creation of dissimilarity feature vector set

In this study, training and testing of MCS are performed using dissimilarity feature vector set. The dissimilarity feature vector set contains two subsets namely- positive (genuine) feature vector subset and negative (forgery) feature vector subset. Positive and negative feature vector subsets are generated using the feature vectors of genuine signature, unskilled and simulated forgery samples.

In the training phase, two scenarios (Scenario-I and Scenario-II) are considered in this work. In Scenario-I, only the feature vectors of genuine signatures of writer and random forgery (genuine signature of another writer) are used. To generate the positive feature vector subset, dissimilarity among 6 genuine signature feature vectors of the writer is computed. In this way, 15 dissimilarity feature vectors per writer are obtained. This resulted in 2400 positive feature vectors from 160 writers. To create the negative feature vector subset, the dissimilarity between 4 genuine signature feature vectors of the first 5 writers and 4 genuine signature feature vectors of 140 writers from the remaining training set of writers is computed. In this way, 2800 negative feature vectors are obtained. Finally, the dissimilarity feature vector set of 5200 feature vectors (2400 positive feature vectors plus 2800 negative feature vectors) is used to train the classifiers of MCS.

In Scenario-II, feature vectors of genuine signature, random forgery, unskilled forgery and simulated forgery signatures are used to train the classifiers. The positive feature vector subset of 2400 positive feature vectors from 160 writers is obtained by computing the dissimilarity among 6 genuine signature feature vectors of the writer. To create the negative feature vector subset, random forgery, unskilled forgery, and simulated forgery samples are used. To generate the negative feature vectors using random forgery, the dissimilarity between 2 genuine signature feature vectors of the first 5 writers and 2 genuine signature feature vectors of 150 writers from the rest of training set of writers is computed. In this way, 1500 negative feature vectors using random forgery are obtained. To generate the negative feature vectors using unskilled forgery, the dissimilarity between 2 genuine signature feature vectors and 2 unskilled forgery signature feature vectors of the writer is computed. In the same way, the dissimilarity between 2 genuine signature feature vectors and 2 simulated forgery signature feature vectors of the writer is computed to obtain the negative feature vectors using simulated forgery. In this way, 1280 negative feature vectors is obtained using unskilled and simulated forgery signature feature vectors. Finally, the dissimilarity feature vector set of 5180 feature vectors (2400 positive feature vectors plus 2780 negative feature vectors) is used to train the classifiers of MCS.

In the testing phase of classifiers of MCS, dissimilarity feature vectors of unskilled and simulated forgery signatures along with genuine signature and random forgery signatures are used for both scenarios. The number of required genuine signatures, random forgery, unskilled forgery and simulated forgery feature vectors to create the positive and negative feature vector subsets is dependent on the number of references used for the questioned signature.

D. Creation of multiple classifier system

To create the MCS, DT, SVM – POLY and SVM – QUAD classifiers are used. To generate the MCS, k-fold cross-validation method is utilized and the value of k is taken 5 in this study. In this way, three multiple classifier systems, namely- MCS of 5 SVM – QUAD classifiers, MCS of 5 SVM – POLY classifiers and MCS of 5 DT classifiers are created.

E. Classification through multiple classifier system

All three multiple classifier systems are trained using H1, H2, H3, and H4 hybrid feature sets for Scenario-I as well as Scenario-II in this study. The questioned signature samples is classified as genuine or forge using 5, 7, 9, 11, and 13 Reference Signatures (RS) and all trained multiple classifier systems are used to classify the questioned signature samples.

VI. Experimental Results and Discussion

In this work, the performance of three multiple classifier systems e.g. MCS of SVM – POLY classifiers, MCS of SVM – QUAD classifiers and MCS of DT classifiers is evaluated using four Hybrid Feature Sets (HFSs) for Scenario-I and Scenario-II. Thus, total 24 experiments are performed in this study. MATLAB 2013a is used to carry out the experiments using 260 writers database. The performance of multiple classifier systems is evaluated using 5, 7, 9, 11 and 13 reference signatures in terms of FAR, FRR, and AER metrics against questioned signature sample using max fusion strategy. FAR is computed for all types of forgeries of the genuine signatures namely- Random (FARR), Unskilled (FARU) and Simulated (FARS). The performance of MCS of SVM – POLY classifiers, MCS of SVM – QUAD classifiers and MCS of DT classifiers under Scenario-I in terms of FRR, FAR and AER using H1, H2, H3, and H4 hybrid feature sets is presented in Table 1, Table 2, Table 3, and Table 4, respectively whereas Table 5, Table 6, Table 7, and Table 8 present the performance of Scenario-II using H1, H2, H3, and H4 hybrid feature sets, respectively.

Table 9 presents the summary of the best result obtained through various experiments performed in this work. The comparison of the performance between the proposed writer-independent HSV system and the existing writer-independent HSV systems in terms of FRR, FAR and AER is presented in Table 10.

Table 1. Performance of multiple classifier systems under Scenario-I using H1 feature set

MCS	RS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
SVM–POLY	5	7.00	0.00	2.00	8.00	4.25
	7	7.00	0.00	2.00	5.00	3.50
	9	7.00	0.00	1.00	4.00	3.00
	11	8.00	0.00	1.00	4.00	3.25
	13	8.00	0.00	1.00	3.00	3.00
SVM–QUAD	5	13.00	1.00	1.00	5.00	5.00
	7	13.00	0.00	1.00	2.00	4.00
	9	13.00	0.00	0.00	2.00	3.75
	11	13.00	0.00	0.00	2.00	3.75
	13	14.00	0.00	0.00	2.00	4.00
DT	5	25.00	1.00	2.00	6.00	8.50
	7	28.00	1.00	2.00	6.00	9.25
	9	28.00	0.00	2.00	6.00	9.00
	11	28.00	0.00	2.00	6.00	9.00
	13	28.00	0.00	2.00	6.00	9.00

Table 2. Performance of multiple classifier systems under Scenario-I using H2 feature set

MCS	RS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
SVM–POLY	5	7.00	0.00	0.00	5.00	3.00
	7	7.00	0.00	0.00	2.00	2.25
	9	7.00	0.00	0.00	2.00	2.25
	11	8.00	0.00	0.00	2.00	2.50
	13	8.00	0.00	0.00	2.00	2.50
SVM–QUAD	5	8.00	0.00	0.00	0.00	2.00
	7	9.00	0.00	0.00	0.00	2.25
	9	10.00	0.00	0.00	0.00	2.50
	11	11.00	0.00	0.00	0.00	2.75
	13	11.00	0.00	0.00	0.00	2.75
DT	5	10.00	6.00	5.00	12.00	8.25
	7	10.00	6.00	3.00	12.00	7.75
	9	14.00	2.00	3.00	12.00	7.75
	11	15.00	2.00	3.00	12.00	8.00
	13	15.00	1.00	3.00	12.00	7.75

Table 3. Performance of multiple classifier systems under Scenario-I using H3 feature set

MCS	RS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
SVM–POLY	5	16.00	0.00	3.00	2.00	5.25
	7	16.00	0.00	2.00	2.00	5.00
	9	16.00	0.00	2.00	1.00	4.75
	11	18.00	0.00	2.00	1.00	5.25
	13	18.00	0.00	0.00	1.00	4.75
SVM–QUAD	5	24.00	0.00	0.00	1.00	6.25
	7	25.00	0.00	0.00	1.00	6.50
	9	25.00	0.00	0.00	1.00	6.50
	11	27.00	0.00	0.00	1.00	7.00
	13	27.00	0.00	0.00	1.00	7.00
DT	5	9.00	6.00	6.00	13.00	8.50
	7	9.00	6.00	4.00	13.00	8.00
	9	13.00	2.00	4.00	13.00	8.00
	11	14.00	2.00	4.00	13.00	8.25
	13	14.00	1.00	4.00	13.00	8.00

Table 4. Performance of multiple classifier systems under Scenario-I using H4 feature set

MCS	RS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
SVM–POLY	5	6.00	0.00	0.00	4.00	2.50
	7	6.00	0.00	0.00	3.00	2.25
	9	6.00	0.00	0.00	3.00	2.25
	11	7.00	0.00	0.00	3.00	2.50
	13	7.00	0.00	0.00	2.00	2.25
SVM–QUAD	5	9.00	0.00	0.00	3.00	3.00
	7	9.00	0.00	0.00	1.00	2.50
	9	9.00	0.00	0.00	0.00	2.25
	11	9.00	0.00	0.00	0.00	2.25

	13	10.00	0.00	0.00	0.00	2.50
DT	5	9.00	6.00	5.00	13.00	8.25
	7	9.00	6.00	3.00	13.00	7.75
	9	13.00	2.00	3.00	13.00	7.75
	11	14.00	2.00	3.00	13.00	8.00
	13	14.00	1.00	3.00	13.00	7.75

Table 5. Performance of multiple classifier systems under Scenario-II using H1 feature set

MCS	RS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
SVM–POLY	5	8.00	0.00	0.00	5.00	3.25
	7	8.00	0.00	0.00	2.00	2.50
	9	9.00	0.00	0.00	2.00	2.75
	11	9.00	0.00	0.00	2.00	2.75
	13	9.00	0.00	0.00	2.00	2.75
SVM–QUAD	5	13.00	0.00	1.00	4.00	4.50
	7	14.00	0.00	1.00	1.00	4.00
	9	15.00	0.00	0.00	0.00	3.75
	11	16.00	0.00	0.00	0.00	4.00
	13	16.00	0.00	0.00	0.00	4.00
DT	5	19.00	1.00	1.00	6.00	6.75
	7	21.00	1.00	1.00	5.00	7.00
	9	22.00	0.00	1.00	4.00	6.75
	11	24.00	0.00	1.00	4.00	7.25
	13	24.00	0.00	1.00	4.00	7.25

Table 6. Performance of multiple classifier systems under Scenario-II using H2 feature set

MCS	RS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
SVM–POLY	5	11.00	0.00	0.00	3.00	3.50
	7	11.00	0.00	0.00	1.00	3.00
	9	13.00	0.00	0.00	0.00	3.25
	11	14.00	0.00	0.00	0.00	3.50
	13	14.00	0.00	0.00	0.00	3.50
SVM–QUAD	5	14.00	0.00	0.00	0.00	3.50
	7	15.00	0.00	0.00	0.00	3.75
	9	16.00	0.00	0.00	0.00	4.00
	11	17.00	0.00	0.00	0.00	4.25
	13	17.00	0.00	0.00	0.00	4.25
DT	5	21.00	0.00	2.00	4.00	6.75
	7	24.00	0.00	1.00	3.00	7.00
	9	24.00	0.00	1.00	3.00	7.00
	11	24.00	0.00	1.00	3.00	7.00
	13	24.00	0.00	1.00	2.00	6.75

Table 7. Performance of multiple classifier systems under Scenario-II using H3 feature set

MCS	RS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
SVM–POLY	5	17.00	0.00	1.00	0.00	4.50
	7	18.00	0.00	1.00	0.00	4.75
	9	19.00	0.00	0.00	0.00	4.75
	11	22.00	0.00	0.00	0.00	5.50
	13	22.00	0.00	0.00	0.00	5.50
SVM–QUAD	5	27.00	0.00	1.00	0.00	7.00
	7	29.00	0.00	0.00	0.00	7.25
	9	32.00	0.00	0.00	0.00	8.00
	11	34.00	0.00	0.00	0.00	8.50
	13	34.00	0.00	0.00	0.00	8.50
DT	5	22.00	0.00	2.00	3.00	6.75
	7	24.00	0.00	1.00	2.00	6.75
	9	24.00	0.00	1.00	2.00	6.75
	11	24.00	0.00	1.00	2.00	6.75
	13	24.00	0.00	1.00	1.00	6.50

Table 8. Performance of multiple classifier systems under Scenario-II using H4 feature set

MCS	RS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
SVM–POLY	5	11.00	0.00	0.00	0.00	2.75
	7	13.00	0.00	0.00	0.00	3.25
	9	13.00	0.00	0.00	0.00	3.25
	11	13.00	0.00	0.00	0.00	3.25
	13	13.00	0.00	0.00	0.00	3.25
SVM–QUAD	5	13.00	0.00	0.00	0.00	3.25
	7	15.00	0.00	0.00	0.00	3.75
	9	15.00	0.00	0.00	0.00	3.75
	11	16.00	0.00	0.00	0.00	4.00
	13	16.00	0.00	0.00	0.00	4.00
DT	5	13.00	0.00	3.00	5.00	5.25
	7	13.00	0.00	2.00	4.00	4.75
	9	16.00	0.00	2.00	4.00	5.50
	11	18.00	0.00	2.00	4.00	6.00
	13	18.00	0.00	2.00	4.00	6.00

Table 9. Summary of Best Result of Various Experiments

SC	MCS	HFS	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
Scenario-I	POLY	H1	7.00	0.00	1.00	4.00	3.00
	QUAD	H1	13.0	0.00	0.00	2.00	3.75
	DT	H1	25.0	1.00	2.00	6.00	8.50
	POLY	H2	7.00	0.00	0.00	2.00	2.25
	QUAD	H2	8.00	0.00	0.00	0.00	2.00
	DT	H2	14.0	2.00	3.00	12.00	7.75
	POLY	H3	16.0	0.00	2.00	1.00	4.75
	QUAD	H3	24.0	0.00	0.00	1.00	6.25
	DT	H3	13.0	2.00	4.00	13.00	8.00
	POLY	H4	6.00	0.00	0.00	3.00	2.25
	QUAD	H4	9.00	0.00	0.00	0.00	2.25
	DT	H4	9.00	6.00	3.00	13.00	7.75
Scenario-II	POLY	H1	8.00	0.00	0.00	2.00	2.50
	QUAD	H1	15.0	0.00	0.00	0.00	3.75
	DT	H1	19.0	1.00	1.00	6.00	6.75
	POLY	H2	11.0	0.00	0.00	1.00	3.00
	QUAD	H2	14.0	0.00	0.00	0.00	3.50
	DT	H2	21.0	0.00	2.00	4.00	6.75
	POLY	H3	17.0	0.00	1.00	0.00	4.50
	QUAD	H3	27.0	0.00	1.00	0.00	7.00
	DT	H3	24.0	0.00	1.00	1.00	6.50
	POLY	H4	11.0	0.00	0.00	0.00	2.75
	QUAD	H4	13.0	0.00	0.00	0.00	3.25
	DT	H4	13.0	0.00	2.00	4.00	4.75

The performance of MCS of SVM – POLY classifiers is better than MCS of SVM – QUAD classifiers and MCS of DT classifiers in terms of AER for both scenarios and for all hybrid feature sets H1, H2, H3, and H4 in most of the cases. Likewise, MCS of SVM – QUAD reports better results for hybrid feature sets H1, H2, and H4 in terms of FRR, FAR and AER for both scenarios as compared to MCS of DT classifiers. MCS of SVM – QUAD also reports better results in terms FAR for unskilled and simulated forgeries for both scenarios and for hybrid feature sets H1 & H2 as compared to MCS of SVM – POLY classifiers. Similarly, MCS of DT classifiers reports better results using hybrid feature set H3 for both scenarios as compared to MCS of SVM – QUAD in terms of FRR.

The multiple classifier systems of Scenario-I report better results in terms of FRR as compared the multiple classifier systems of Scenario-II whereas the multiple classifier systems of Scenario-II report better results in terms of FAR for unskilled and simulated forgeries as compared to the multiple classifier systems of Scenario-I.

The performance of hybrid feature sets H2 and H4 is better than hybrid feature sets H1 and H3 in terms of AER for Scenario-I whereas, in Scenario-II, the performance obtained through hybrid feature sets H1 and H4 is better than hybrid feature sets H2 and H3 in terms of AER for most of the cases. For Scenario-I, the lowest AER of 2.00 is obtained through the experiment performed using H2 feature set along with MCS of SVM – QUAD for 5 reference signatures whereas, in Scenario-II, the experiment performed using H1 feature set along with MCS of SVM – POLY is reported lowest AER of 2.50 for 7 reference signatures.

Table 10. Comparison of Existing and Proposed Writer Independent Offline HSV Systems

SN	Authors	Classifier	Feature Set(s)	FRR (%)	FARR (%)	FARU (%)	FARS (%)	AER (%)
1	C. Santos et. al. [4] (2004)	Neural Network	Graphometric	10.33	4.41	1.67	15.67	8.02
2	D. Bertolini et. al. [3] (2010)	SVM	Graphometric	11.32	4.32	3.00	6.48	6.28
3	D. Rivard et. al. [5] (2011)	SVM	ESC & DPDF	9.77	0.02	0.32	10.65	5.19
4	R. Kumar et. al. [6] (2012)	NN &SVM- RBF	Surroundedness	13.76	-	-	13.76	13.76
5	G. Eskander et. al. [7] (2012)	SVM	ESC & DPDF	7.73	0.016	0.17	13.50	5.38
6	J.Swanepoel et. al. [8] (2012)	LDF &QDF	DRT & DTW	-	-	-	-	4.93
7	G. Eskander et. al. [9] (2013)	SVM	ESC & DPDF	14.36	0.02	0.35	14.24	7.24
8	A. Hamadene et. al. [10] (2016)	OC – SVM	CT & DCCM	-	-	-	-	18.42
9	L. Hafemann et. al. [11] (2016)	SVM – RBF	CNN	2.17	0.17	0.50	13.00	3.96
10	Proposed Approach (Scenario-II)	SVM – POLY	H1 Feature Set	8.00	0.00	0.00	2.00	2.50
11	**Proposed Approach (Scenario-I)**	**SVM – QUAD**	**H2 Feature Set**	**8.00**	**0.00**	**0.00**	**0.00**	**2.00**

VII. Conclusion

The study aimed at proposing a writer-independent offline HSV system with reduced FAR for unskilled and simulated forgeries as well as reduced AER. The writers involved in the testing process are not included in the training process and MCS used in proposed approach is able to classify the questioned signatures of writers of the testing set very well without retraining. This implies that developed MCS for writer-independent offline HSV is capable of absorbing the signature of an unknown writer without retraining.

It is observed from the experiments, FAR for unskilled and simulated forgeries is high in most of the cases when classifiers of MCS are trained using only genuine signature and random forgery samples whereas FAR for unskilled and simulated forgeries is reduced but FRR is increased when unskilled and simulated forgeries signature samples are involved in the training of classifiers of MCS.

It is also observed from the experiments, the performance of the MCS depends on the classifiers and feature set used in the system. The experiments performed using hybrid feature set H2 (HOG plus LOSIB) along with MCS of SVM – QUAD classifiers for Scenario-I and hybrid feature set H1 (HOG plus DWT) along with MCS of SVM – POLY classifiers for Scenario-II report better performance in terms of the AER as compared to other experiments.

From the comparison between proposed and existing writer-independent offline HSV systems, it is evident that proposed writer-independent offline HSV system using MCS of SVM – QUAD classifiers along with H2 hybrid feature set under Scenario-I and MCS of SVM – POLY classifiers along with H1 hybrid feature set for Scenario-II outperform the existing WI offline HSV systems in terms of FAR for unskilled and simulated forgeries as well as for AER. It is therefore, concluded that a competent MCS for writer-independent offline HSV with reduced FAR for unskilled and simulated forgeries and AER can be developed using H2 feature set along with SVM – QUAD classifiers and H1 feature set along with SVM – POLY classifiers. However, as a future research endeavor deep learning, a new emerging research, can be utilized to improve the performance of WI offline HSV system.

References

[1] N. H. Abbas, K. N. Yasen, K. H. Faraj, L. F. Razak, and F. L. Malallah, “Offline Handwritten Signature Recognition using Histogram Orientation Gradient and Support Vector Machine”, *Journal of Theoretical and Applied Information Technology*, 96(8), pp. 2075-2084, 2018.

[2] G. Hiremath, “Verification of Offline Signature using Local Binary and Directional Pattern”, *International*

Journal of Innovative Science, Engineering and Technology, 3(1), pp. 76-80, 2016.

[3] D. Bertolini, L. Oliveira, E. Justino, and R. Sabourin, "Reducing Forgeries in Writer Independent Offline Signature Verification through Ensemble of Classifiers", *Elsevier Journal of Pattern Recognition,* 3(1), pp. 387 – 396, 2010.

[4] C. Santos, E. Justino, F. Bortolozzi, and R. Sabourin, "An Offline Signature Verification Method Based on the Questioned Document Expert's Approach and a Neural Network Classifier", In *Proceeding of IEEE 9th Int'l Workshop on Frontiers in Handwriting Recognition*, Washington, DC, USA, pp. 498 – 502, 2004.

[5] D. Rivard, E. Granger, and R. Sabourin, "Multi Feature Extraction and Selection in Writer Independent Offline Signature Verification", *International Journal on Document Analysis and Recognition,* 16(1), pp. 1 - 21, 2011.

[6] R. Kumar, J. Sharma, and B. Chanda, "Writer Independent Offline Signature Verification using Surroundedness Feature", *Pattern Recognition Letters*, 33(3), pp. 301–308, 2012.

[7] G. Eskander, R. Sabourin, and E. Granger, "Adaptation of Writer Independent Systems for Offline Signature Verification", In *Proceeding of International Conference on Frontiers in Handwriting Recognition,* Bari, Italy, pp. 434 – 439, 2012.

[8] J. Swanepoel, and J. Coetzer, "Writer Specific Dissimilarity Normalisation for Improved Writer Independent Signature Verification", In *Proceeding of IEEE International Conference on Frontiers in Handwriting Recognition*, Bari, Italy, pp. 391- 396, 2012.

[9] G. Eskander, R. Sabourin, and E. Granger, "Hybrid Writer Independent Writer Dependent Offline Signature Verification System", *IET Biometrics,* 2 (4), pp. 169– 181, 2013.

[10] A. Hamadene and Y. Chiban, "One Class Writer Independent Offline Signature Verification using Feature Dissimilarity Thresholding", *IEEE Transactions on Information Forensics and Security,* 11(6), pp. 1226- 1238, 2016.

[11] L. Hafemann, R. Sabourin, and L. Oliveira, "Writer Independent Feature Learning for Offline Signature Verification using Deep Convolutional Neural Networks", In *Proceeding of International Joint Conference on Neural Networks,* Vancouver, Canada, pp. 2576 – 2583, 2016.

[12] L. Kuncheva, "*Combining Pattern Classifiers: Methods and Algorithm*", Wiley, New York, 2004.

[13] L. Breiman, "Bagging Predictor", *Machine Learning,* 24(2), pp.123 – 140, 1996.

[14] Y. Freund, and R. Schapire, "Experiments with a New Boosting Algorithm", In *Proceeding of Thirteen International Conference on Machine Leaning*, Bari, Italy, pp. 148 – 156, 1996.

[15] Tin. K. Ho., "The Random Subspace Method for Constructing Decision Forest", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 20(8), pp. 832 – 844, 1998.

[16] P. Zhang, "Model Selection via Multifold Cross Validation", *Annals of Statistics,* 21(1), pp. 299 – 313, 1993.

[17] J. Amanda, C. Sharkey, and N. Sharkey, "The Test and Select Approach to Ensemble Combination", In *Proceeding of First International Workshop on Multiple Classifier Systems Springer Verlog*, London, UK, pp. 30 – 44, 2000.

[18] H.Vala, and A. Baxi, "A Review on Otsu Image Segmen-tation Algorithm", *International Journal of Advanced Research in Computer Engineering & Technology*, 2(2), pp. 387 – 389, 2013.

[19] M. Berkay, B. Yanikoglu, and A. Kholmatov, "Offline Signature Verification using Classifier Combination of HOG and LBP Features", In *Proceeding of IEEE International Joint Conference on Biometrics*, Washington, Dc, USA, pp. 1- 8, 2011.

[20] M. Kociołek, A. Materka, M. Strzelecki, and P. Szczypiński, " Discrete Wavelet Transform Derived Features for Digital Image Texture Analysis", In *Proceeding of Proceeding of Interational Conference on Signals and Electronic Systems*, Lodz, Poland, pp. 163-168, 2001,

[21] O. Oscar, E. Alegre, L. Robles, and V. Castro, "Local Oriented Statistics Information Booster for Texture Classification" In *Proceeding of 22nd IEEE International Conference on Pattern Recognition*, Stockholm, Sweden, pp. 1 - 6, 2014.

Author Biographies

Ashok Kumar received the Master's degree in Computer Applications from Dr. Bhim Rao Ambedkar University Agra, India in 2002 He is pursuing Ph. D. in Computer Science from Gurukula Kangri Vishwavidyalaya Haridwar, India. He is an Associate Professor at the Department of Computer Applications, Invertis University Bareilly, India. He has authored 16 research papers in the International/ national Journals and Conferences.

Karamjit Bhatia received the M.Phil. **(**Computer Applications) in 1998 from University of Roorkee, Roorkee (now IIT Roorkee), India and Ph. D. (Computer Science) in 2001 from Gurukula Kangri Vishwavidyalaya, Haridwar, India. Currently, he is working as a Professor at Department of Computer Science, Gurukula Kangri Vishwavidyalaya, Haridwar, India. His research interests include Distributed Systems, Image Processing, and Mobile Ad-hoc Networks. He has published around 40 research papers in International Journals of repute and National/ International Conference proceedings. Six research scholars have been awarded Ph. D. degree under his supervision.