Offline Handwritten Signature Verification Using SVM and LBP

Ashok Kumar\textsuperscript{1} and Karamjit Bhatia\textsuperscript{2}

\textsuperscript{1}Associate Professor, AIT CSE, Chandigarh University, Mohali, Punjab, India.

\textsuperscript{2}Professor, Department of Computer Science, Gurukula Kangri (Deemed to be University), Haridwar, India.

Abstract
Signature verification is a difficult research area since two individual's signatures may seem alike, but an individual's signature might change depending on the situation. The accuracy of the signature verification framework is mostly determined by the classifier and feature extraction scheme employed in the classification process. Keeping this in mind, the purpose of this research is to examine how well a support vector machine with polynomial kernel classifier and a Local Binary Pattern feature set can be coupled to create a writer-independent offline signature verification system. Two signature databases with 100 and 260 writers are employed to assess the system's performance. Genuine signatures, as well as random forgery are taken into account in the development of the desired system, and genuine signatures, as well as simulated forgery, unskilled forgery, and random forgery signatures are used to evaluate the developed system's performance. In a simulation investigation, the false acceptance rate for random, unskilled, and simulated forged signatures is achieved 0.00 \%, 7.00 \%, and 18.00 \%, respectively, but the false rejection rate is achieved 0.00 \% by utilizing the Local Binary Pattern feature set.

Keywords: Writer-independent approach, Offline signature verification system, Local binary pattern features, False acceptance rate and false rejection rates, Support vector machine.
1. INTRODUCTION

An individual’s signature is regarded as a recognized biometric feature. Individual signature can be used to verify and validate a person's personal identification as well as the legitimacy of a document [1]. The HSV (Handwritten Signature Verification) system is used to validate a signature’s authenticity. For the offline HSV system, the forgeries set is divided into three forgery subsets: random forgery, unskilled forgery, and simulated forgery. The signature of a different writer is regarded a random forgery. The forger in the unskilled (also known as simple) forgery creation process only knows the writer's name, whereas the forger in the simulated (also known as skilled) forgery creation process knows the writer's name as well as the genuine signature of the writer and has practiced the genuine signature many times. Fig. 1 depicts a writer's genuine signature and its accompanying random forgery, unskilled forgery, and Simulated forgery.

![Genuine and corresponding forgery signatures](image)

**Fig.1:** Genuine and its corresponding forgery signatures.

False Acceptance Rate (FAR), Average Error Rate (AER), and False Rejection Rate (FRR) performance measures are utilized to assess the performance of an offline HSV system [2]. FRR stands for the percentage of real signatures that are identified as forgery signatures, whereas FAR stands for the percentage rate of forgery signatures that are recognized as genuine signatures. The AER is calculated by taking the mean of the FAR and FRR. Random Forgery (FARR) signatures, Unskilled Forgery (FARU) signatures, and Simulated Forgery (FARS) signatures are used to calculate FAR in this work [15].

To build the offline/online HSV systems, researchers used Writer-Dependent (WD) as well as Writer-Independent (WI) techniques [2]. In WD approach, a personal model (also called specific model) is constructed for every writer. In contrast, using the WI technique, an alternative model (also called global model) is constructed to cope with signatures from all writers, and the resulting model is capable of classifying new writer signatures without retraining [21]. Fig.2 illustrates a Writer-independent offline signature verification mechanism [17].
In the writer-independent technique, the Dissimilarity Feature Vector (DFV) of the questioned-signature ($Q$) is compared with DFVs of the reference signature samples $R_n$ to determine whether it is genuine or forged. Dissimilarity concept was proposed by Pekalska E. [3], which said that dissimilarities should be maximal for the things of different classes and minimal for things of the same class. The difference between $R_n$ and $Q$ is passed to a classifier to make a partial judgement and finally, fusion procedures are employed to arrive at a final judgement from a collection of partial decisions.

WI offline HSV system is created employing Local-Binary-Pattern (LBP) features in this suggested study. LBP features are a type of pseudo dynamic features that extract the signature image's local properties. Local features are better for distinguishing real and simulated forged signatures because they can discern signature images with minor alterations [5]. As a result, LBP features are proposed in this study for constructing an offline HSV system based on the WI technique.
The novelty of this work is two-fold. First, the fitness of LBP features in designing an offline HSV system using WI approach is probed as these features are capable of distinguishing the images with minor variation, and second, the usefulness of Support Vector Machine with Polynomial Kernel (SVM-POLY) classifier in development of a WI approach based offline HSV system using is investigated.

The present work is devoted to device an approach with following objectives:

1. to absorb unfamiliar writers' handwritten signatures without retraining,
2. to minimize the FRR through SVM-POLY classifier system, and
3. to study the efficacy of LBP features and SVM-POLY classifier for developing an offline HSV.

2. LITERATURE REVIEW

Several researchers have contributed to the construction of an offline HSV system based on the WD method using LBP feature set in the literature. Wang et al. [4] presented LBP features for texture categorization at first. Ahlawat et al. [5], Hiremath et al. [6], Madhavi et al. [7], and Mathew et al. [8] employed the LBP feature set in conjunction with the Local-Directional-Pattern feature set to build the WD offline systems. Different classifiers are employed by researchers for classification process in order to construct the LBP features based offline system utilizing the WD technique, such as the k-NN classifier used by Pal et al. [9], Vickram et al. and [10], and Patki et al. [11]. Ramesh et al. [12], Singh et al. [13], and Wajid et al. [14] developed the WD offline system using the Support Vector Machine (SVM) classifier and the LBP feature set. The WD offline system was developed by all of the above-mentioned researchers using a single/multiple classifier system.

Researchers have not addressed the WI method for developing an offline HSV system using LBP features, according to a literature review. For the offline HSV system, Santos et al. [16] presented the WI method. Bertolini et al. [2] built the HSV system using an ensemble of SVM classifiers & a GF set and reported FRR of 11.32 %, FARR of 4.32%, FARU of 3.00 %, and FARS of 6.48 %. Batista et al. [18] claimed FRR of 8.33 %, FARR of 0.50 %, FARU of 0.50 %, and FARS of 15.50 % using pixel density features and an SVM classifier. Dominique et al. [19] used stroke in addition to Spatial Distribution (SD) features and obtained FRR of 9.77%, FARR of 0.02%, FARU of 0.32%, and FARS of 10.65% through SVM classifier. The features related to the texture and signature shape were utilized by Kumar et al. [20].

Literature survey reveals that different feature sets, as mentioned above, are considered by the researchers while designing offline HSV system using WI approach. However, it is found that LBP features are not given due attention by researchers in the development of such systems using WI approach.

The present study focuses on exploring the effectiveness of SVM-POLY classifier system with LBP feature set for the development of WI approach based offline HSV system.
3. **RESEARCH METHODOLOGY**

The major steps to develop the WI approach based HSV system are creation of signature database used for training and testing, preprocessing and feature extraction, the creation of DFV set, and the classifier system's training and testing.

3.1. **Signature database**

Two signature databases of SD-100 and SD-260 are in used verification of signatures. Database utilized in research contain the genuine, unskilled forgery, and simulated forgery signatures. Signatures of the 60 Individuals and 40 Individuals of the SD-100 are utilized for training and testing, respectively, whereas SD-260 separated into 160 and 100 Individual’s signatures for the training and testing. To accomplish the training/testing process, six genuine, four unskilled forgery, and four simulated forgery samples of the signature per writer are utilized.

3.2. **Preprocessing and feature extraction**

A median filter is used to eliminate noise from the signature image during the preprocessing step. The grey level signature image is then converted to a binary image. After that, the signature image is cropped and scaled to 256 x 512 pixels. The intensity value of the central pixel is compared with the intensity value of 8 pixels in a 3 x 3 window lying on a circular path of radius 1 in the neighborhood in an anticlockwise orientation to extract the LBP features. The LBP feature vector of length 256 is produced in this manner.

3.3. **Creation of DFV set**

The positive (genuine) plus negative (forgery) Dissimilarity Feature (DF) vector subsets are utilized to train the classifiers. The Positive Feature Vector (PFV) subset is obtained by calculating the dissimilarity among six genuine signatures per writer; consequently, 15 different combinations are found. In this manner, 900 PFVs are produced from the 60 writers for the SD-100 signature database. Likewise, 2400 PFVs from the 160 writers are produced for the SD-260 signature database.

In this work, “only random forgery signatures are utilized to form the Negative Feature Vector (NFV) subset. To generate the NFV subset for the SD-100 signature database, the dissimilarity between first 4 genuine signatures of the first 5 writers and first 4 genuine signatures of 50 writers from the remaining training set is computed. This results into 1000 negative feature vectors”. Further, to produce the NFV subset for the SD-260 signature database, “the dissimilarity between first 4 genuine signatures of the first 5 writers and first 4 genuine signatures of 140 writers from the remaining training set is computed. In this way, 2800 negative feature vectors are obtained”. Thus, Dissimilarity Feature (DF) vector set of total 1900 (900 PFVs plus 1000 NFVs) DF vectors for the SD-100 signature database and DF vector set of 5200 (2400 PFVs plus 2800 NFVs) DF vectors for the SD-260 signature database are utilized to train the classifiers.
3.4. Classifier system's training and testing

The classification system is developed by using the SVM-POLY classifier to classify the questioned signature. The DF vector sets of 1900 DF vectors and 5200 DF vectors are utilized to train the classifier system. Then, all trained classifier system is utilized to classify the genuine, simulated forgery, unskilled forgery, and random forgery samples of signature of the test set. To classify the test set signatures, “3, 5, 7, 9, 11, 13, and 15” reference signatures are used. The system performance is assessed by means of the FRR, FARR, FARU, and FARS.

4. EXPERIMENTAL RESULTS

The experiments are carried out using MATLAB 2013a. Experiments are conducted with “3, 5, 7, 9, 11, 13, and 15” reference signatures. The mean fusion strategy is used to combine the classifiers' incomplete decisions. The writers who work on the training stage are not involved in the system's testing phase in the current method.

The performance of SVM-POLY classifiers on the SD-100 and SD-260 signature databases is shown in Tables 1 and 2. FRR, FARR, FARU, and FARS are used to describe the performance. The results of a comparison of proposed and existing WI offline systems are shown in Table 3. Only existing WI offline systems whose performance was reported in FRR, FARR, FARU, and FARS performance matrices were used to compare the performance of proposed WI offline systems. Fig. 3 and 4 illustrate the elapsed time of all reference signatures for the SD-100 signature database and the SD-260 database, respectively. Experiments were conducted on a configuration of 4 GB RAM with core 2 dual processor system and time elapsed in experiment is recorded for training/testing of the classifier systems. The elapsed time supposed to differ due to the configuration of the system.

![Elapsed Time for FRR](image1.png)

![Elapsed Time for FARR](image2.png)

![Elapsed Time for FARU](image3.png)

![Elapsed Time for FARS](image4.png)

Fig.3. Elapsed time of all reference signatures for SD-100 signature database
Fig. 4. Elapsed time of all reference signatures for SD-260 signature database

Table 1. Performance of the SVM-POLY classifier in FRR, FARR, FARU, and FARS for SD-100 signature database.

<table>
<thead>
<tr>
<th>RS</th>
<th>FRR (%)</th>
<th>FARR (%)</th>
<th>FARU (%)</th>
<th>FARS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7.50</td>
<td>10.00</td>
<td>12.50</td>
<td>22.50</td>
</tr>
<tr>
<td>5</td>
<td>7.50</td>
<td>10.00</td>
<td>15.00</td>
<td>27.50</td>
</tr>
<tr>
<td>7</td>
<td>7.50</td>
<td>5.00</td>
<td>10.00</td>
<td>27.50</td>
</tr>
<tr>
<td>9</td>
<td>5.00</td>
<td>2.50</td>
<td>10.00</td>
<td>32.50</td>
</tr>
<tr>
<td>11</td>
<td>2.50</td>
<td>2.50</td>
<td>10.00</td>
<td>30.00</td>
</tr>
<tr>
<td>13</td>
<td>0.00</td>
<td>2.50</td>
<td>10.00</td>
<td>25.00</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>2.50</td>
<td>12.50</td>
<td>25.00</td>
</tr>
</tbody>
</table>
Table 2. Performance of the SVM-POLY classifier in FRR, FARR, FARU, and FARS for SD-260 signature database.

<table>
<thead>
<tr>
<th>RS</th>
<th>FRR (%)</th>
<th>FARR (%)</th>
<th>FARU (%)</th>
<th>FARS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2.00</td>
<td>5.00</td>
<td>7.00</td>
<td>16.00</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>5.00</td>
<td>7.00</td>
<td>17.00</td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>3.00</td>
<td>6.00</td>
<td>18.00</td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>0.00</td>
<td>8.00</td>
<td>16.00</td>
</tr>
<tr>
<td>11</td>
<td>0.00</td>
<td>0.00</td>
<td>8.00</td>
<td>19.00</td>
</tr>
<tr>
<td>13</td>
<td>0.00</td>
<td>0.00</td>
<td>7.00</td>
<td>18.00</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>0.00</td>
<td>7.00</td>
<td>20.00</td>
</tr>
</tbody>
</table>

Table 3. Comparative results of existing and proposed WI offline systems

<table>
<thead>
<tr>
<th>SN</th>
<th>Authors</th>
<th>Classifier</th>
<th>Feature set</th>
<th>FRR (%)</th>
<th>FARR (%)</th>
<th>FARU (%)</th>
<th>FARS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bertolini et al. [2]</td>
<td>Ensemble of SVM</td>
<td>GF</td>
<td>11.32</td>
<td>4.32</td>
<td>3.00</td>
<td>6.48</td>
</tr>
<tr>
<td>2</td>
<td>Batista et al. [18]</td>
<td>Ensemble of SVM, Gentle AdaBoost</td>
<td>Pixel Density</td>
<td>8.33</td>
<td>0.50</td>
<td>0.50</td>
<td>15.50</td>
</tr>
<tr>
<td>3</td>
<td>Dominique et al. [19]</td>
<td>Ensemble of DT</td>
<td>Stoke, SD</td>
<td>9.77</td>
<td>0.02</td>
<td>0.32</td>
<td>10.65</td>
</tr>
<tr>
<td>4</td>
<td>Kumar et al. [20]</td>
<td>MLP-ANN, Ensemble of SVM-RBF</td>
<td>Texture, Shape</td>
<td>8.33</td>
<td>-</td>
<td>-</td>
<td>8.33</td>
</tr>
<tr>
<td>5</td>
<td>Proposed Approach</td>
<td>SVM-POLY</td>
<td>LBP</td>
<td>0.00</td>
<td>0.00</td>
<td>7.00</td>
<td>18.00</td>
</tr>
</tbody>
</table>

For both signature databases, the necessary elapsed time appears to be increasing as the number of reference signatures grows. It's also worth noting that the SD-260 signature database needed elapsed time is longer than the SD-100 signature database.

The performance of the SD-260 and SD-100 signature databases is compared, and it is discovered that the SD-260 database outperforms the SD-100 database in terms of FRR, FARR, and FARS. The SVM-POLY classifier gives the best results utilizing the 13 reference signatures for the signature database SD-260, with FRR of 0.00%, FARR of 0.00%, FARU of 7.00%, and FARS of 18.00%.
The study shows that increasing the number of reference signatures and writer signatures improves the performance of the WI offline HSV system in terms of FRR, FARR, FARU, and FARS while simultaneously increasing the elapsed time.

5. CONCLUSION

In this study, the performance of an SVM-POLY classifier trained with the LBP feature set is tested using two signature databases. The writers who took part in the testing are not included in the training. The SVM-POLY classifier used in the study are able to classify the signatures in the testing set without retraining, implying that the suggested classifier system for WI offline signature verification can absorb the signature of a new writer without retraining. Experiments revealed that the SD-260 signature database's classifiers have a higher classification accuracy than the SD-100 signature database. As a result of the preceding, it is reasonable to conclude that increasing the number of writers involved in the training and testing processes will enhance the system's classification accuracy in terms of FRR and FAR.

As indicated in the comparison table of the proposed and existing WI offline HSV systems, the proposed classifier system beats the existing WI offline HSV systems in terms of FRR. In conclusion, an effective WI offline HSV system can be developed by integrating the SVM-POLY classifier with the LBP features.

REFERENCES


Offline Handwritten Signature Verification Using SVM and LBP

Journal on Document Analysis and Recognition, 83-103.

