Future Computing and Informatics Journal

Volume 5
Issue 1 (2020) *Volume 5 Issue 1*

Article 3

2020

Signature identification and verification systems: a comparative study on the online and offline techniques

nehal hamdy al-banhawy
Al-Azhar University - Egypt, nehalhamdy2017@gmail.com

Heba Mohsen
Faculty of Computers and Information Technology, Future University in Egypt, New Cairo, Egypt, hmohsen@fue.edu.eg

Neveen I. Ghali Prof.
neveen.ghali@fue.edu.eg, Neveen.Ghali@Fue.edu.eg

Digital Commons

Future Computing and Informatics Journal

Homepage: https://digitalcommons.aaru.edu.jo/fcij/

SIGNATURE IDENTIFICATION AND VERIFICATION SYSTEMS: A COMPARATIVE STUDY ON THE ONLINE AND OFFLINE TECHNIQUES

Nehal Hamdy Al-banhawy[1,a], Heba Mohsen[2,b], Neveen Ghali[2,c]
[1]Al-Azhar University, Egypt
[2]Future University in Egypt

[a] nehalhamdy2017@gmail.com, [b] hmohsen@fue.edu.eg,
[c] neveen.ghali@fue.edu.eg

ABSTRACT

Handwritten signature identification and verification has become an active area of research in recent years. Handwritten signature identification systems are used for identifying the user among all users enrolled in the system while handwritten signature verification systems are used for authenticating a user by comparing a specific signature with his signature that is stored in the system. This paper presents a review for commonly used methods for pre-processing, feature extraction and classification techniques in signature identification and verification systems, in addition to a comparison between the systems implemented in the literature for identification techniques and verification techniques in online and offline systems with taking into consideration the datasets used and results for each system.

1. INTRODUCTION

Handwritten signature analysis is one of the most common techniques for determining the identity of individuals which we are often exposed to in our daily lives. There are other biometric systems as face, fingerprint, iris, vein and DNA but, the signature remains the most acceptable for identification in some tracks. However, this study is one of the biggest challenges, since the signature of one person may not be identical [1].

Biometric systems are applied in two main scenarios: identification and verification. Identification systems are used for identifying the user among all users registered in the system. In case of signature, it tries to answer the question 'who owns this signature'. Verification systems are used for

authenticating a user by comparing one specific biometric stored in the system. In case of the handwritten signature systems, it is used to classify the query signature as genuine or forged [2]. Forgery is categorized into three types: random forgery, simple forgery (not skilled), skilled forgery. In random forgery, the forger doesn't have any information about the user's name or his signature and instead of that he uses his own signature. In this case, his signature shape is completely different from user's signature. In simple forgery, the forger knows the user's name only not the signature shape. In this case, his signature shape may be similar to the user's signature. In skilled forgery, the forger knows both the user’s name and signature with some

Published by Arab Journals Platform, 2020

practising to perform a good faking for the user's signature. It is harder to detect this type of forgery because it has higher similarity to the genuine signature [3].

There are two main types of signature identification and verification systems: Offline (static) systems and Online (dynamic) systems. In the offline systems, signature is captured or scanned from a document such as bank checks then the system must read and extract features from the image of the signature. This is in contrast with online systems, where the system uses devices like tablets and smart phones for capturing additional information such as time, pressure, pen up and down, azimuth while the user is signing which allow to extract more features [4].

In signature identification, a user's signature is provided to the system then system compares this signature with all other signatures enrolled in dataset and calculates the similarity results. The best result refers to identified user [2]. While signature verification has two basic approaches: writer dependent and writer independent. In writer independent, a single model is trained for all users and is responsible for matching the query signature to reference signatures in a similarity/dissimilarity space as shown in figure 1. Most researchers prefer this approach because there is no need to retrain the system when adding a new writer.

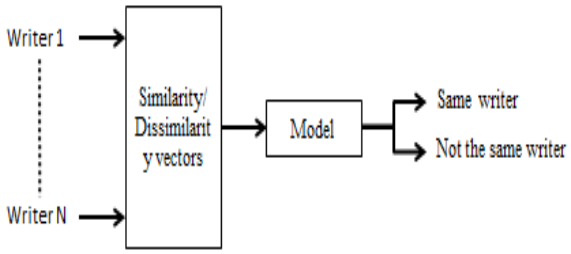

Fig. 1: writer independent system [5].

In writer dependent, one specific model is trained for each user and is responsible for authenticating her signatures, so the system needs to be updated (retrained) when we add a new writer (signer) as shown in figure 2 [5].

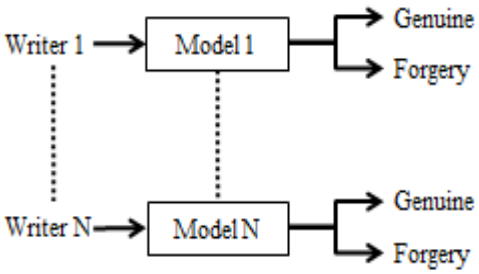

Fig. 2: writer dependent system [5].

The signature identification system was first developed in 1965. One of the first attempts to present a new approach in signature identification was found in [6] which introduced the use of the revolving active deformable model as an effective way for capturing the unique characteristics of the whole signature. Many reviews have been published like [7, 8 and 9] to explain improvements in signature identification. One of the first studies on signature verification was [10] in 1977, they worked on characteristics extracted from signatures sectioned into vertical and horizontal zones, then studies continued.

Many surveys have been published like [1, 3, 4, 11, 12 and 13] to explain improvements and several research directions in signature verification.

This paper is arranged as follows: Firstly, the three stages of the identification and verification systems are discussed and the techniques that are used for every stage are explained. Secondly, the evaluation method of identification and verification systems performance is discussed. Then a comparison between 29 recent proposed techniques for identification and verification systems is made. Finally, the important points found in the comparison are discussed.

2. IDENTIFICATION AND VERIFICATION SYSTEM

Signature identification and verification problem goes through three stages: Dataset preprocessing, Feature extraction and Classification. The common datasets that was used in researches and the common used techniques in preprocessing, feature extraction and classification stages will be discussed in detail.

2.1. Datasets

There are 10 common datasets used in researches for the two approaches of signature identification and verification systems: online and offline [14] that are explained as follows:

- CEDAR dataset: It contains data from 55 writers: 24 genuine signatures plus 24 skilled forgeries for each writer in a gray scale PNG format [15], [16] and [17].
- MCYT: there are two subsets of the MCYT signature, namely MCYT-100 (contains data from 100 writers: 25 genuine and 25 forged online samples for each writer) and MCYT-75 (contains data from 75 writers: 15 genuine and 15 forged offline signatures for each writer) [17], [18], [19], [20], [21] and [22].
- GPDS Signature: is a Spanish offline signature dataset. There are many subsets of GPDS Signature. GPDS-100 contains data from 100 writers. GPDS-150 contains data from 150 writers. GPDS-960 contains data from 960 writers (but it is no longer available).

- GPDS-Synthetic contains data from 4000 writers. All of this datasets have 24 genuine signatures plus 30 forged signature for each writer in a black and white and gray scale version [15], [16], [17], [23], [24], [25], [26], [27] and [28].
- UTSig: (university of Tehran Persian Signature dataset) is a Persian offline Signature dataset that contains data from 115 writers: 27 genuine and 45 forged signatures for each writer in gray scale TIF files [29], [30].
- SigComp2009: is a signature verification competition which provided an online and offline dataset containing training and evaluation sets. The training set contains data from 12 writers: 5 genuine and 5 forged signatures for each writer. The evaluation set contains data from 100 writers: 12 genuine and 6 forged signatures for each writer [31].
- SigComp11: includes two subsets of dataset: Chinese and Dutch signature samples. It contains both online and offline signature samples in RGB colored image. Number of signatures in online dataset is different from those in offline dataset and signatures number of Chinese dataset is different from those in Dutch dataset [22], [32].
- BHSig260 signature dataset: contains signatures of 260 writers, 100 were signed in Bengali and 160 were signed in Hindi dataset. Each of them contains 24 genuine and 30 forged signatures for each writer [15].
- SVC2004: was the first international signature verification competition. It has two datasets of online handwritten signatures; the first dataset contains only coordinate information and the second dataset contains additional information such as pen orientation and pressure. Each dataset contains data from 100 writers with 20 genuine signatures plus 20 skilled forgeries for each writer [18], [33], [34].
- SUSIG: (Sabanci University Signature) is divided into two subcorpora: visual and blind. In the Visual Subcorpus, signatures were obtained using an Interlink Electronics tablet which has a pressure-sensitive LCD screen that could see signatures writer while signing. It contains data from 100 writers with 20 genuine and 10 forged samples (5 skilled and 5 very skilled) for each writer. In the Blind Subcorpus, signatures were collected using Wacom's Graphire2 tablet and pen without visual feedback. It contains data from 100 writers with 10 genuine samples from 70 writers, 8 genuine samples from 30 writers and forged samples as in the Visual Subcorpus. Data stored for both subcorpora contains the x-y coordinates, pressure level and time stamp for each writer [33].
- ATVS dataset: It is an online signature dataset which contains data from 350 users with 25 signatures for each user [33], [35].

Published by Arab Journals Platform, 2020

2.2. Dataset preprocessing

Preprocessing of database is the process of improving the signature image after reading it. It is a very important stage in both online and offline identification and verification systems whereas it affects the accuracy and minimizes the computational time.

Commonly used steps in preprocessing are:

- **Convert to gray:**
  If signature images are colored, it has to be converted to gray scale as in references [28], [36]. A grayscale image is a simple image in which the only colors are shades of gray. It needs less information to be provided for each pixel. Most references which converted images to gray scale image used the following equation:

  Gray color = 0.299*Red + 0.5876*Green + 0.114*Blue **(1)**

- **Filtering or noise removal:**
  Some images may be corrupted with noise in many cases. There are many filtering techniques to remove impulse noise which are commonly classified as linear (such as: averaging or Gaussian filters) and non-linear filtering (such as: median filter and fuzzy filter)[37]. Many references did this step in preprocessing [16], [28], [34], [36] and [38].

- **Cropping:**
  Cropping is the removal of unwanted outer areas from an image, thus reducing the dimension of it as in references [23], [28], [29], [30] and [34].

- **Rotation:**
  Some signatures may be rotated at a certain angle in a clockwise direction, so it needs to rotate it at the same angle in the opposite direction as in references [16], [30] and [34].

- **Binarization:**
  One of the important steps in image processing is image binarization. It converts the image into black and white image (0 or 1) using different algorithms which divides into two categories: global binarization, local binarization. Some important global binarization methods are: Fixed Thresholding Method, Otsu Method and Kittler Method. They used single thresholding value for the whole image. Some important local binarization methods are: Niblack Method, Adaptive Method, Sauvola Method and Bernsen Method. They calculated the thresholding value locally pixel by pixel [39]. Many researchers did this step in preprocessing [16], [23], [28], [29], [30], [36] and [38].

- **Thinning:**
  Thinning is the morphological operation approach used to reduce a digital image to the minimum size. It is implemented using various algorithms such as: Zhang Suen Thinning algorithm, Canny Edge detection, Edge Based thinning algorithm, Optimized iterative algorithm using successive erosion, Guo and Hall's parallel Thinning algorithm [40]. Many researchers did this step in preprocessing [16], [23], [28], [36] and [38].

- **Normalization:**
  Normalization is the process of making all images in the dataset of fixed size as researchers in references [16], [23], [28], [29], [30], [34] and [36], did.

2.3. Signature features extraction

Signature features represents magnitudes that can be extracted from the whole pen path, with aim of describing each signature as a vector of values. Features must allow us to distinguish between signatures for different users. Signature features is divided into: global features and local features.

Global features are features that are extracted from whole signature. It measures various sides of the image such as shape, color or texture. Common global features for offline handwritten signature verification are: Density, Width, Height and

aspect ratio (is the ratio of width of signature in bounding box/ height of signature in bounding box) etc. [2], [28], [35] and [38]. There are more global features in [41].

Local features extracted from specific parts of the signature. Using of local features gives better performance than global features provide in image search [42]. Common local features for offline signature acceleration, x and y coordinate and pressure [22], [43].

Big efforts were made by researchers to find good feature representation for signature image. The main descriptors proposed for the problem are:

- Geometrical features describe the geometrical properties of the signature image [24], [44].
- Mathematical transformations: many researchers used some kind of mathematical transformations (such as: Wavelet transform [26], [34] and [45], Contourlet transform [29], Gabor wavelet transform [30], discrete Radon transform [20], [21] and Fourier transform [45]) as feature extractors.
- Directional features: describe image in terms of the strokes direction in the signature using histogram of oriented gradients technique [27], [28], [36].
- Extended shadow-code: extracts information of the spatial distribution of the signature [18].
- Texture features: includes end points, crossing points and branch points. End points are the starting and ending points of the signature stroke. Crossing points are intersection points between one signature stroke and another stroke. Branch points are points where one signature stroke branch into two strokes [16] and [28].
- Interest point matching: such as SURF (Speeded Up Robust Features) [46] and SIFT (Scale-Invariant Feature Transform) [32] became used for computer vision tasks.

In recent years, there has been an increased interest in techniques that learns features automatically from signature images instead of designing feature extractors as in Deep Learning Models.

verification are: horizontal and vertical projections, center of gravity, normalized area of black pixels, area of black pixels in each grid region, gradient and concavity features etc. [2], [19], [23] and [38].

Some features are extracted from online systems called Dynamic Function Features which are functions of time t such as: signing time, Number of pen lifts, x, y, total values of position, speed,

2.4. Classification for identification and verification systems

Identification (recognition) is the general process of detecting signature's owner which is a multi-class classification issue. The stages of identification system are shown in figure 3. Firstly, the input signatures are scanned and preprocessed then the specific features are extracted and saved in knowledge base. In the last stage (classification stage), the extracted features are compared to the template signature saved in knowledge base and studied an affiliation of tested signature to which class [9].

Published by Arab Journals Platform, 2020

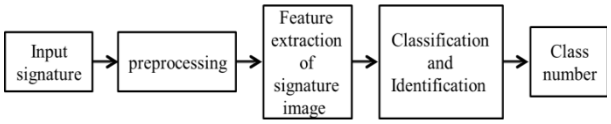

Fig. 3: signature identification system.

Verification is the general process of making a decision about the signature and determines whether it is genuine or forged, so it is a two-class classification issue. The stages of verification system are shown in figure 4. Signature verification system stages are the same as signature identification system but it differs in classification stage where we know the class of tested signature and check its authenticity to that class [4].

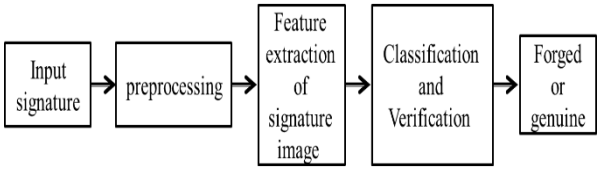

Fig. 4: signature verification system.

Researchers introduced many approaches for signature identification and verification techniques that were used in feature extraction and classification stages. For signature identification systems, Harsha [9] has classified signature identification techniques into seven types: Template Matching, Hidden Markov Models, Neural Networks, Statistical, Structural, Wavelet-Based and Support Vector Machine. For signature verification systems, Donato Impedovo [11] has classified signature verification techniques into three types: Template Matching, Statistical, Structural approaches. Researchers in [3], [4] have classified the signature verification techniques into seven types: Template Matching, Neural Network, Wavelet Based Approach, Structural Approach, Support Vector Machine, Hidden Marcov and Gaussian Mixture Model, Statistical Approach. Diaz et. al. [13] has classified them to five types: Template Matching, Statistical measures, Statistical models, Deep Learning, support Vector Machine (SVM)), Structural, Fusions. Then techniques that are used in identification systems can be used also in verification systems and vice versa.

Besides some statistical classifiers (e.g. correlation [45]), several classifiers have been used in online and offline identification and verification systems that we will discuss in detail as follows:

- **Template Matching:**

Template Matching is a recognition technique for finding areas of an image which are similar to a template image. It is used to identify numbers, printed characters and other small objects. The common algorithm used for template matching is dynamic time warping (DTW) [3], [43], mahalanobis distance [30], [38] and euclidean distance [15], [23], [29], [36] and [38].

- Neural Networks (NNs):

Neural networks are widely used in pattern recognition because it is powerful, easy to use and capable of learning and generalizing [4]. It consists of inputs, neurons, which are units that takes several observed inputs and produce a single output, weights and outputs [47]. Multi-layer perceptron (MLP) is the common form of NN which using back propagation algorithm to adjust the weights. We feed the feature vector into the input layer for training purpose, perform a pattern matching on the test set in the hidden layer and do a classification. It can implement a discrimination function that separates input data into classes. The output of the neural network is forged or genuine in verification system and class number in identification

system [16], [24], [28], [34], [35], [44] and [48].

- **Probabilistic Neural Network (PNN):**

A probabilistic neural network is a feedforward NN that is widely used in classification and pattern recognition problems. It consists of 4 layers: input, pattern, summation and output layers. The input layer receives the input vector, the pattern layer includes one neuron per each input vector and computes how close the inputs are to the training inputs, the summation layer contains one neuron per each class of inputs and sums these contribution for each class of inputs to generate a vector of probabilities, the output layer holds the maximum probability and produces the final outcome [20], [26].

- **Support Vector Machine (SVM):**

Support vector machines (SVMs) are a group of associated supervised learning methods used in classification and regression. It is a binary classifier responsible for finding the decision boundary to separate different classes [49] but sometimes it is used as multi-classifier. It is depending on the idea of finding a hyperplane (called margins) that best separates the features into different domains as shown in figure 5. The closest points to the margin are called support vector points. Margin Width is the distance of vectors from margin. In multi class SVM approach, SVM is established for each class in the system by discriminating that class with the remaining (k-1) classes so the number of used SVM is k. A test set is classified in which the class with the maximum value of the discriminant function f(x) is assigned to it [16], [25], [27], [32], [38] and [46].

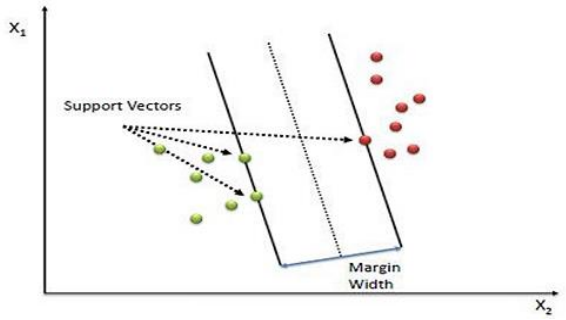

Fig. 5: Illustration of Linear SVM [50].

- **Hidden Markov Models (HMMs) and Gaussian Mixture Models (GMMs)**:

HMMs and GMMs are statistical methods widely used for developing signature identification and verification systems. HMMs are modeled consisting of Markov processes with finite number of states. It is used in sequence analysis where each point in signature path gives vector of values that used in HMM. In HMM, matching of signatures and models happens. This matching is achieved by calculation probability distribution of features of signature. GMMs can use several multidimensional Gaussian probability distributions for clustering low dimensional data. It is can be considered as HMM with a single state [18], [19].

- **K- nearest neighbor (K-NN):**

The k-nearest neighbor algorithm is a method for classifying objects depending on closest training samples in the feature space. It is commonly used in pattern recognition. In signature identification and verification systems, it is used to classify signatures based on their nearest neighbors classes with

Published by Arab Journals Platform, 2020

taking more than one neighbor into account. K-NN classification in identification systems has two stages; the assigning of the nearest neighbors and the assigning of the class that is using those neighbors. When a query signature enters the system, it is compared to its nearest neighbors. In the verification systems, when a query signature enters the system, it is compared to all of the reference signatures in the class of its nearest neighbors. If the resulting dissimilarity measurement is lower than or equals to a threshold value of the classifier, the person is verified; otherwise it is rejected [17], [21] and [30].

- **Structural approach:**

In the structural approaches, symbolic data structure such as trees, graphs and strings are used to represent patterns of the signature. Structural approach is based on relational organization of low-level features into higher-level structures then these structures are matched with models stored in database. Each signature in the system is represented as a symbolic representation (number of graphs or trees). When a tested signature enters, its symbolic representation is compared with stored representations of signatures in database and the best match is found in the class whose results have the greatest mean of all the results achieved [2], [4].

- Deep Learning:

In recent years, there has been an increased interest in techniques that learned features from raw data (pixels in images) automatically and are represented hierarchically in multiple levels [1]. This is the strong point of deep learning against traditional machine learning approaches. It consists of several layers between the input and output layer which allows for many stages of non-linear information processing units that are exploited for feature learning and pattern classification. Deep learning approaches can be categorized into: Supervised (that uses labeled data), semi-supervised (Reinforcement Learning) (that occurs based on partially labeled datasets) and unsupervised learning (that can occurs without labeled data). Deep Neural Networks (DNN), Convolutional Neural Networks (CNN) [15], [17], [51] and Recurrent Neural Networks (RNN) are the common models of supervised learning. Generative Adversarial Networks (GAN) and RNN are used for semi-supervised learning. Auto-Encoders (AE) [33], Restricted Boltzmann Machines (RBM) [25] and deep belief network (DBN) are common models for unsupervised learning [52].

3. Evaluation of signature identification and verification systems performance

To decide if the system is good or not, we have to evaluate its performance. Biometric systems are dependent to several kinds of errors. Each system whether identification or verification system, is measured in a specific way.

3.1. Identification system performance metrics

Performance of identification systems is measured using one of the following metrics:

- **Identification rate =**

$$\textbf{Number of correct identification for all writers} \;/(\textbf{total number of query signature for all writers}) \times \mathbf{100}\% \qquad \textbf{(2)}$$

- **The Correct Classification Rate (CCR)**

$$= (\mathbf{TP} + \mathbf{TN})/(\mathbf{TP} + \mathbf{TN} + \mathbf{FP} + \mathbf{FN}) \times \mathbf{100}\% \qquad \textbf{(3)}$$

Where:

True positive (TP): number of users that have been correctly authenticated.

True Negative (TN): number of impostors that have been correctly authenticated.

False positive (FP): number of users that have been incorrectly authenticated.

False Negative (FN): number of impostors that have been incorrectly authenticated.

3.2. Verification system performance metrics:

Verification systems performance is usually evaluated based on the two terms, False Rejection Rate (FRR) and False Acceptance Rate (FAR).

- False Rejection Rate (FRR): Ratio of authentic users that are incorrectly rejected. It is calculated as:

FRR = 1-TP/ (number of genuine signatures) **(4)**

- False Acceptance Rate (FAR): Ratio of impostors that are accepted by the biometric system. It is calculated as:

FAR = FP/ (number of forged signatures) **(5)**

We can use FRR and FAR to find Receiver operating characteristic curve (ROC curve) and equal error rate as follows:

- Receiver Operating Characteristic curve (ROC)

The ROC curve is the most commonly used method to measure a performance for biometric systems. It is plotted to represent FAR according to FRR (figure 6). The advantage of this method is that it gives a precise representation for the performance of a biometric system through a single curve allowing the comparison of different biometric systems [53]. Then we can compute the area under the curve (AUC) and the equal error rate (ERR). The higher AUC and the lower equal error rate mean that the performance is good. The optimal result is obtained if the AUC equals 1 and the ERR equals 0.

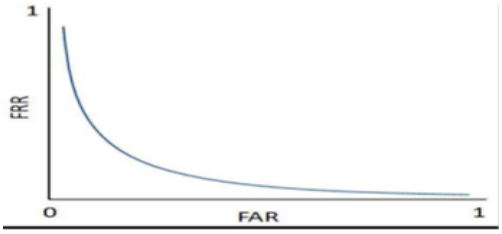

Fig. 6: Simple example of ROC curve: FAR versus FRR [54].

- **Equal Error Rate (EER):**

This error rate coincides with the point at which the FAR and FRR cross, as shown in figure 7. It is widely used to evaluate and compare biometric authentication systems. Whenever the EER is near to 0.0 %, it means that the performance of the target system is better [54]. We can calculate the verification accuracy using EER as follows:

Accuracy = 1- equal error rate **(6)**

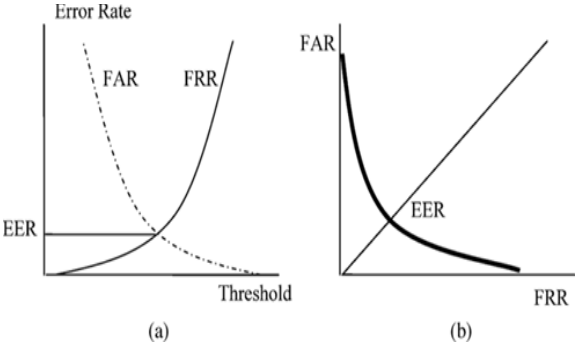

Fig. 7: Example of EER (a) by using the FRR and FAR curves. (b) by using ROC curve [10].

Published by Arab Journals Platform, 2020

4. Comparison of recent proposed techniques for identification and verification systems

In this section a comparison is presented between 29 recent proposed techniques for identification and verification systems. This comparison is made between techniques used in researches for feature extraction and classifiers in identification and verification systems with mentioning the used between 7 recent papers for identification systems is presented in table 1, and then a comparison between 18 recent papers for verification systems is made in table 2. Finally we comparison between 4 recent papers for both identification and verification systems is presented in table3.

5. Discussion

From the previous tables, some important points are found:

- Online systems achieve better results than offline systems whereas it takes dynamic features that are obtained during the writing into account beside other features. However offline systems are the most applied in research because its applicability and ease of use. Beside, online systems need a special hardware which offline systems don't need to like digitizers and pressure sensitive tablets.
- Whenever the number of users in dataset increased, the performance decreases in identification and verification systems; while if the number of samples per user increases, the system has better performance.

When comparing the feature extraction techniques, it is found that:

- Using of histogram of oriented gradients technique, SIFT and SURF techniques and mathematical transformation gave better performance with all classifiers.
- The use of deep learning techniques in feature extraction achieved good performance but it is not the best.
- Discrete wavelet transform and probabilistic neural network system achieved best results on GPDS dataset (which is the most used in research).

When comparing between classifiers performance on different databases, it is found that:

- All previous classifiers can be used for offline and online systems except DTW which used only in online systems.
- All previous classifiers can be used in both identification systems and verification systems.
- SVM achieved better performance when comparing with euclidean distance with the same dataset.
- Gaussian mixture models (GMM) gave better performance when comparing with hidden Marcov models, Gaussian mixture models and DTW with the same dataset (MCYT-100).
- Structural approach achieved a good performance but it costs a lot of calculations.
- Euclidean distance and nearest neighbor gave near performance with the same dataset (UTSig).
- When comparing between MLP and SVM in the same system, we found that MLP achieved better performance when using big dataset but SVM achieved better performance when using smaller dataset.
- Neural network is better when comparing with k- nearest neighbor with the same dataset (MCYT) by using mathematical transformations for feature extraction

Table 1- Comparison of 7 recent proposed techniques for identification systems

reference	dataset	Extracted features	classifier	evaluation
[17] 2019 (offline)	GPDS-4000 MCYT and CEDAR datasets	CNN	Used 1-Nearest Neighbor (1-NN) for classification task obtained from fully-connected layers	Achieved 96.91%, 96.41% and 98.30% accuracy for the GPDS-4000, MCYT and CEDAR datasets respectively
[46] 2018 (online)	Two datasets, the first dataset contains 240 signatures that was taken from ten writers and the second dataset contains 768 signatures that was taken from 32 writers	Using speed up robust features (SURF)	Support vector machine(SVM)	98.75% accuracy for the first dataset and 97.7% accuracy for the second dataset
[32] 2018 (offline)	SigComp2011	extracted local features using SIFT (the Scale Invariant feature Transform)[49] and K-means algorithm	SVM with RBF kernel	98.86% accuracy
[35] 2015 (online)	Two sets of ATVS dataset are collected, dataset I contains 25 signature samples per each writer. Dataset II contains 46 signature samples per each writer	9 global features	feed forward neural network	98% accuracy for dataset I and 89% accuracy for dataset II
[6] 2011 (online)	small dataset of 27 users	graph theory	graph norm	94.25% accuracy
[48] 2011 (offline)	collected from 35 persons	Using adaptive learning vector quantization (LVQ) neural network compact architecture	adaptive learning vector quantization (LVQ) neural network compact architecture	98% accuracy
[38] 2009 (offline)	dataset consisting of 600 persons of genuine and forgery writers	extract local and global features	1. Euclidean distance 2. Mahalanobis distance 3. Gaussian empirical rule 4. Fusion of the three classifiers achieved by SVM	92.61%, 93.36% and 91.52%, 97.17% for accuracy respectively

Table 2- Comparison of 18 recent proposed techniques for verification system

reference	dataset	Extracted features	classifier	evaluation
[51] 2018 (offline)	Dataset consists of 6000 signatures with 1000 genuine and 1000 forged signatures per subject , two classes ware established for each writer (genuine and forgery)	Used CNN to extract features	CNN	Achieved 98.11 % training accuracy and 98.23% validation accuracy with 80-20 data split ratio
[43] 2018 (online)	dataset consists of 10 writers with 10 genuine signatures and 10 forged signatures per each user	Extracted some features as (coordinates, pressure, altitude and azimuth) which are function of time t	DTW algorithm was used to calculate warping distance to differentiate between a genuine signature and its forgery and make a decision about verification result.	This system can detect fake signatures with an accuracy of 90.4%.
[19] 2017 (Online)	MCYT-100	extract the local features	Gaussian Mixture Model (GMM) and Longest Common Sub-Sequences (LCSS)	0.4% Equal Error rate for GMM-LCSS model
[26] 2017 (offline)	GPDS-960 and another data from 20 writers	Using Discrete Wavelet Transform	probabilistic neural network [39]	92.06 % accuracy result for GPDS dataset, 92.87% for database B.
[36] 2017 (offline)	Dataset contains 10 writers	adopted Histogram of Oriented Gradient (HOG) for features extraction technique	K-NN classifier with using of Euclidean distance as a distance computation measure	They record their results in Confusion matrix
[15] 2017 (offline)	CEDAR, GPDS300, GPDS Synthetic Signature Dataset, and BHSig260 signature corpus	Using convolutional Siamese network	Euclidean distance	Achieved 100% accuracy for CEDAR dataset, 76.83% accuracy for GPDS 300 Signature Corpus, 77.76% accuracy for GPDS Synthetic Signature Corpus, 86.11 % accuracy for Bengali dataset, 84.64% accuracy for Hindi dataset.

Table 2 (Continue)

reference	dataset	Extracted features	classifier	evaluation
[20] 2016 (offline)	MCYT signature dataset and another database that was collected from 100 writers and 10 forgers, containing 1000 genuine signatures,500 random forgeries and 500 skilled forgeries	Using discrete radon transform for feature extraction then Principal component analysis is used to reduce the number of transformed values after DRT	Probabilistic neural network	9.87% EER was reported for MYCT dataset and 1.51%, 3.23% and 13.07% EER of random, casual and skilled forgeries, respectively for the other dataset
[22] 2016 (online)	Chinese dataset from SigComp2011 and MCYT 100 dataset	Extracted 15 function features	DTW was used to compute distance values between query signature and all reference signatures in the template dataset	Achieved 1.69% and 1.77% EER for Chinese and MCYT-100 datasets respectively.
[45] 2015 (online)	Japanese online dataset from ICDAR2013 which contains data from 11 writers for training set and 20 writers for evaluation set.	extracted using a combination of: Fourier Transform based features, Wavelet Transform based features and Global features	correlation	27.48 % FAR , 25.54% FRR and 73.49% accuracy
[33] 2015 (online)	ATVS, SVC, SUSIG datasets	The features have been learned from ATVS dataset by using a sparse autoencoder with one hidden layer	One-class classifier	0.83% EER for SVC2004 dataset and 0.77% EER for SUSIG dataset
[23] 2015 (offline)	GPDS dataset	Extract 4 features : Pixel Density, Cell Angle, Pixel Angle, Pixel Length	Euclidean distance	12.53% EER
[21] 2013 (offline)	MCYT-75 signature CORPUS	Extract global feature using discrete radon transform, then the extreme points warping (EPW) algorithm was used to find the distance between each training signature and reference signatures belonging to the claimed ID	k-nearest neighbor classifier	obtained 80% overall performance

Table 2 (Continue)

reference	dataset	Extracted features	classifier	evaluation
[34] 2013 (online)	SVC	Wavelet transform	neural network (NN)	3.5% EER
[24] 2012 (offline)	GPDS dataset	Extract these geometric features: Normalized area of signature, Aspect Ratio, Maximum horizontal and vertical histogram, Center of mass (centroid), Tri surface feature, the Six fold surface feature and Transition feature.	neural networks (NN)	Achieved 82.66% accuracy When the system is tested on genuine and forged signatures that it wasn't trained on
[25] 2011 (offline)	GPDS-300 dataset	extract fourteen different features which are global and local feature using Restricted Boltzmann Machine model(RBM)	SVM	9.24 FAR and 26.42 FRR
[44] 2011 (offline)	dataset of 100 signatures from 3 writers	Five geometric features are: Area, Centroid Coordinates, Kurtosis, Eccentricity and Skewness.	artificial neural network (ANN)	93% accuracy
[16] 2011 (offline)	GPDS300 and CEDAR dataset	proposed a feature set that is depending on neighboring pixel surroundedness which represent shape and texture attributes of the signature	multilayer perceptron and SVN	The best results are 91.67% accuracy for CEDAR dataset and 86.24% accuracy for GPDS dataset by using multilayer perceptron
[18] 2007 (online)	MCYT dataset and SVC 2004 dataset	Basic functions, Geometric normalization, Extended functions, Time derivatives, Signal normalization	Hidden Markov Models (HMM)	0.74% and 0.05% EER to skilled and random forgeries, respectively for MCYT dataset, 6.9% and 3.02% EER to skilled and random forgeries for SVC 2004 dataset, respectively

Table 3 - Comparison of 4 recent proposed techniques for identification and verification systems

reference	dataset	Extracted features	classifier	evaluation
[28] 2013 (offline)	GPDS dataset	extracted global, mask and texture features	neural network	100 % accuracy for identification rate, 12.5% FAR and 10% FRR in verification
[30] 2011 (offline)	Persian, South African, Turkish and Spanish signatures datasets	Gabor wavelet	nearest neighbor classifier for identification and Mahalonobis distance for verification	100%, 77.3% accuracy for Persian and Spanish datasets respectively in identification and 15%, 16.8%, 9% EER for Persian, South African and Turkish databases respectively in verification.
[27] 2010 (offline)	GPDS and GAVAB (which collected from 28 persons with 4 signatures per person) signature datasets	used Pyramid histogram of oriented gradients (PHOG) to extract features	1. KNN 2. logistic regression 3. MLP 4. SVM	The best results were achieved by using SVM classifier that is 99% and 96% accuracy for GPDS and DAVAB datasets, respectively in identification. Beside 4.0% FRR, 3.25% FAR for GPDS dataset and 16.4% FRR, 14.2% FAR for DAVAB dataset in verification
[29] 2009 (offline)	Persian signature dataset contains 20 classes with 20 genuine signature and 10 forgery signature for each class. English signature dataset contains 22 classes with 30 genuine signature and 12 forgery signature for each class.	used Contourlet transform	euclidean distance	100% and 93% accuracy for Persian dataset and English dataset respectively for identification , 14.00% EER for skilled forgery set in Persian dataset and 23% EER for skilled forgery set in English dataset for verification

- Neural network is better when comparing with hidden Marcov model with the same dataset (SVC 2004).
- Probabilistic neural network gave better performance than multi-layer perceptron with the same dataset.

Finally, it is noticed that neural network is the best classifier. Until now, deep learning techniques didn't achieve the expected performance because it was not applied sufficiently in research.

6. Conclusion

This paper presented a review for commonly used methods for preprocessing, feature extraction methods and techniques for online and offline handwritten signature identification and verification systems. In addition to a comparison between proposed systems that were used in 29 recent papers for feature extraction, identification techniques and verification techniques in online and offline systems with displaying of database used and results for each system. From the comparison, it was concluded that using of histogram of oriented gradients technique, SIFT and SURF techniques and mathematical transformation for feature extraction gives better performance. All mentioned classifiers can be used in both online and offline identification systems and verification systems except DTW (used only for online systems). The use of support vector machine (SVM), neural network (NN) and deep learning for classification achieves remarkable performance. So the combination of one of these features extraction techniques with one of these classifiers (except deep learning because it doesn't need a feature extractor) in one system, leads to better system with better performance.

References

[1] Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2017, November). Offline Handwritten Signature Verification—Literature Review. In 2017 Seventh International Conference On Image Processing Theory, Tools And Applications (Ipta) (Pp. 1-8). Ieee.

[2] Fotak, T., Bača, M., & Koruga, P. (2011). Handwritten Signature Identification Using Basic Concepts Of Graph Theory. Wseas Transactions On Signal Processing, 7, 117-129.

[3] Mohammed, R. A., Nabi, R. M., Sardasht, M., Mahmood, R., & Nabi, R. M. (2015, December). State-Of-The-Art In Handwritten Signature Verification System. In 2015 International Conference On Computational Science And Computational Intelligence (Csci) (Pp. 519-525). Ieee.

[4] Al-Omari, Y. M., Abdullah, S. N. H. S., & Omar, K. (2011, June). State-Of-The-Art In Offline Signature Verification System. In 2011 International Conference On Pattern Analysis And Intelligence Robotics (Vol. 1, Pp. 59-64). Ieee.

[5] Bhattacharyya, S., Mukherjee, A., Pan, I., Dutta, P., & Bhaumik, A. K. (Eds.). (2017). Book Chapter 3 " Artificial Immune Recognition System For Offline Handwritten Signature Verification". Hybrid Intelligent Techniques For Pattern Analysis And Understanding. Crc Press.

[6] Pavlidis, I., Papanikolopoulos, N. P., & Mavuduru, R. (1998). Signature Identification Through The Use Of Deformable Structures. Signal Processing, 71(2), 187-201.

[7] Vala, K. A., & Joshi, N. P. (2014). A Survey On Off-Line Signature Recognition And Verification Schemes. International Journal Of Advanced Research In Electrical, Electronics And Instrumentation Engineering, 3, 7735.

[8] Impedovo, D., Pirlo, G., & Russo, M. (2014, September). Recent Advances In Offline Signature Identification. In 2014 14th International Conference On Frontiers In Handwriting Recognition (Pp. 639-642). Ieee.

[9] Harsha G. Chavan, Pradnya A. Vikhar (2018), “A Survey: Offline Handwritten Signature Recognition System”, Multidisciplinary Journal Of Engineering And Technology Issn: 2348 - 6953, Volume 5, Issue 3&4, Pp.08-15.

[10] Nagel And Rosenfeld(1977),“Off-Line System” Ieee T. Comp.

[11] Impedovo, D., & Pirlo, G. (2008). Automatic Signature Verification: The State Of The Art. Ieee Transactions On Systems, Man, And Cybernetics, Part C (Applications And Reviews), 38(5), 609-635.

[12] El-Henawy, I., Rashad, M., Nomir, O., & Ahmed, K. (2013). Online Signature Verification: State Of The Art. International Journal Of Computers & Technology, 4(2c2), 664-678.

[13] Diaz, M., Ferrer, M. A., Impedovo, D., Malik, M. I., Pirlo, G., & Plamondon, R. (2019). A Perspective Analysis Of Handwritten Signature Technology. Acm Computing Surveys (Csur), 51(6), 117.

[14] Bezerra, B. L. D., Zanchettin, C., Toselli, A. H., & Pirlo, G. (2017). Handwriting: Recognition, Development And Analysis. Nova Science Publishers, Inc.
[15] Dey, S., Dutta, A., Toledo, J. I., Ghosh, S. K., Lladós, J., & Pal, U. (2017). Signet: Convolutional Siamese Network For Writer Independent Offline Signature Verification. Arxiv Preprint Arxiv:1707.02131.
[16] Kumar, R., Sharma, J. D., & Chanda, B. (2012). Writer-Independent Off-Line Signature Verification Using Surroundedness Feature. Pattern Recognition Letters, 33(3), 301-308.
[17] Çalik, N., Kurban, O. C., Yilmaz, A. R., Yildirim, T., & Ata, L. D. (2019). Large-Scale Offline Signature Recognition Via Deep Neural Networks And Feature Embedding. Neurocomputing.
[18] Fierrez, J., Ortega-Garcia, J., Ramos, D., & Gonzalez-Rodriguez, J. (2007). Hmm-Based On-Line Signature Verification: Feature Extraction And Signature Modeling. Pattern Recognition Letters, 28(16), 2325-2334.
[19] Shashidhar, S., & Sravya, A. (2017). Online Handwritten Signature Verification System: Using Gaussian Mixture Model And Longest Common Sub-Sequences.
[20] Ooi, S. Y., Teoh, A. B. J., Pang, Y. H., & Hiew, B. Y. (2016). Image-Based Handwritten Signature Verification Using Hybrid Methods Of Discrete Radon Transform, Principal Component Analysis And Probabilistic Neural Network. Applied Soft Computing, 40, 274-282.
[21] Abdelrahaman, A. A., & Abdallah, M. A. (2013, August). K-Nearest Neighbor Classifier For Signature Verification System. In 2013 International Conference On Computing, Electrical And Electronic Engineering (Icceee) (Pp. 58-62). Ieee.
[22] Chen, Z., Xia, X., & Luan, F. (2016, August). Automatic Online Signature Verification Based On Dynamic Function Features. In 2016 7th Ieee International Conference On Software Engineering And Service Science (Icsess) (Pp. 964-968). Ieee.
[23] Al-Maqaleh, B. M., & Musleh, A. M. Q. (2015). An Efficient Offline Signature Verification System Using Local Features. International Journal Of Computer Applications, 975, 8887.
[24] Pansare, A., & Bhatia, S. (2012). Handwritten Signature Verification Using Neural Network. International Journal Of Applied Information Systems, 1(2), 44-49.
[25] Ribeiro, B., Gonçalves, I., Santos, S., & Kovacec, A. (2011, November). Deep Learning Networks For Off-Line Handwritten Signature Recognition. In Iberoamerican Congress On Pattern Recognition (Pp. 523-532). Springer, Berlin, Heidelberg.
[26] Daqrouq, K., Sweidan, H., Balamesh, A., & Ajour, M. (2017). Off-Line Handwritten Signature Recognition By Wavelet Entropy And Neural Network. Entropy, 19(6), 252.
[27] Zhang, B. (2010). Off-Line Signature Verification And Identification By Pyramid Histogram Of Oriented Gradients. International Journal Of Intelligent Computing And Cybernetics, 3(4), 611-630.
[28] Sthapak, S., Khopade, M., & Kashid, C. (2013). Artificial Neural Network Based Signature Recognition & Verification. International Journal Of Emerging Technology And Advanced Engineering (Ijetae), 2(8), 191-197.
[29] Pourshahabi, M. R., Sigari, M. H., & Pourreza, H. R. (2009, December). Offline Handwritten Signature Identification And Verification Using Contourlet Transform. In 2009 International Conference Of Soft Computing And Pattern Recognition (Pp. 670-673). Ieee.
[30] Sigari, M. H., Pourshahabi, M. R., & Pourreza, H. R. (2011). Offline Handwritten Signature Identification And Verification Using Multi-Resolution Gabor Wavelet. International Journal Of Biometrics And Bioinformatics (Ijbb), 5(4), 234-248.
[31] Blankers, V. L., Van Den Heuvel, C. E., Franke, K. Y., & Vuurpijl, L. G. (2009, July). Icdar 2009 Signature Verification Competition. In 2009 10th International Conference On Document Analysis And Recognition (Pp. 1403-1407). Ieee.
[32] Kurnaz, Sefer, And Adnan Al-Khdhairi.(2018). "Offline Signature Identification System To Retrieve Personal Information From Cloud." Iosr Journal Of Computer Engineering (Iosr-Jce).
Volume 20, Issue 1, Pp 56-64.
[33] Saffar, M. H., Fayyaz, M., Sabokrou, M., & Fathy, M. (2018). Online Signature Verification Using Deep Representation: A New Descriptor. Arxiv Preprint Arxiv:1806.09986.
[34] Nilchiyan, M. R., & Yusof, R. B. (2013, December). Improved Wavelet-Based Online Signature Verification Scheme Considering Pen Scenario Information. In 2013 1st International Conference On Artificial Intelligence, Modelling And Simulation (Pp. 8-13). Ieee.
[35] Babita, P. (2015). Online Signature Recognition Using Neural Network. Journal Of Electrical & Electronics, 4(3), 1.
[36] Rajput, G. G., & Patil, P. (2017). Writer Independent Offline Signature Recognition

Based Upon Hogs Features. A Ugc Recommended Journal. Volume-9 • Number-1. Pp. 59-67.
[37] Varghese, J. (2013). Literature Survey On Image Filtering Techniques. International Journal Of Engineering Research And Technology.
[38] Kisku, D. R., Gupta, P., & Sing, J. K. (2009, December). Fusion Of Multiple Matchers Using Svm For Offline Signature Identification. In International Conference On Security Technology (Pp. 201-208). Springer, Berlin, Heidelberg.
[39] Garg, N. (2013). Binarization Techniques Used For Grey Scale Images. International Journal Of Computer Applications, 71(1).
[40] Abhishek, Lakshmesha K.N,(2017). "Thinning Approach In Digital Image Processing",Https://Www.Ijltet.Org/Journal_Details.Php?Id=922&J_Id=4110, Special Issue - Sacaim , 326-330, #Ijltetorg.
[41] Martinez-Diaz, M., Fierrez, J., & Hangai, S. (2015). Signature Features. Encyclopedia Of Biometrics, 1375-1382.
[42] Hassaballah, M., Abdelmgeid, A. A., & Alshazly, H. A. (2016). Image Features Detection, Description And Matching. In Image Feature Detectors And Descriptors (Pp. 11-45). Springer, Cham.
[43] Patil, B. V., & Patil, P. R. (2018, February). An Efficient Dtw Algorithm For Online Signature Verification. In 2018 International Conference On Advances In Communication And Computing Technology (Icacct) (Pp. 1-5). Ieee.
[44] Karouni, A., Daya, B., & Bahlak, S. (2011). Offline Signature Recognition Using Neural Networks Approach. Procedia Computer Science, 3, 155-161.
[45] Tahir, M., & Akram, M. U. (2015, November). Online Signature Verification Using Hybrid Features. In 2015 Second International Conference On Information Security And Cyber Forensics (Infosec) (Pp. 11-16). Ieee.
[46] Hamadly, I., Khaleel, A., Munim, A., Hassan, H. E., & Mohamed, H. K. (2018). Online Signature Recognition And Verification Using (Surf) Algorithm With Svm Kernels. Journal Of Al-Azhar University Engineering Sector, 13(49), 1332-1344.
[47] Corradin, R. S. (2008). Signature Verification In Consignment Notes. Department Of Computer Science.
[48] Khuwaja, G. A., & Laghari, M. S. (2011). Offline Handwritten Signature Recognition. World Academy Of Science, Engineering And Technology, 59, 1300-1303.
[49] Evgeniou, T., & Pontil, M. (1999, July). Support Vector Machines: Theory And Applications. In Advanced Course On Artificial Intelligence (Pp. 249-257). Springer, Berlin, Heidelberg.
[50] Https://Towardsdatascience.Com/Support-Vector-Machines-Svm-C9ef22815589.
[51] Gideon, S. J., Kandulna, A., Kujur, A. A., Diana, A., & Raimond, K. (2018). Handwritten Signature Forgery Detection Using Convolutional Neural Networks. Procedia Computer Science, 143, 978-987.
[52] Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2019). A State-Of-The-Art Survey On Deep Learning Theory And Architectures. Electronics, 8(3), 292.
[53] Cherifi, F., Hemery, B., Giot, R., Pasquet, M., & Rosenberger, C. (2010). Performance evaluation of behavioral biometric systems. In Behavioral Biometrics for Human Identification: Intelligent Applications (pp. 57-74). IGI Global.
[54] El-Abed, M., & Charrier, C. (2012). Evaluation of biometric systems, book chapter. Intech open science.