République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Larbi Tébessi –Tébessa
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie
Département de Mathématiques et d'Informatique

MEMOIRE DE MASTER

Domaine : Mathématiques et Informatique
Filière : Informatique
Option : Systèmes et Multimédia

***Thème* :**

Vérification des Signatures Manuscrites et Détection de Falsifications en Utilisant des Caractéristiques Texturales

Présenté par :

Walid BOUAMRA
Mohamed El Amine ZOGHBI

Devant le jury :

Abdelhakim GHARBI	MAA	Président	Université Larbi Tébessi- Tébessa
Chawki DJEDDI	MCB	Rapporteur	Université Larbi Tébessi- Tébessa
Lakhdar LAIMECH	MAA	Examinateur	Université Larbi Tébessi- Tébessa

Date de Soutenance: 30-05-2016

Année Universitaire : 2015-2016

الملخص

إن العمل المقدم في هذا المخطوط يندرج في إطار التحليل والتعرف على الوثائق، وبصفة أخص التحقق من التوقيعات والكشف عن التزوير. يكمن الهدف منه، في تعزيز وتقوية قدرات أنظمة التحقق من التوقيع، بواسطة منحها القدرة على العمل بشكل واقعي عن طريق تدريبها وتعليمها بنفس الطريقة التي يتدرب بها البشر. أي، بمعاينة العينات الإيجابية فقط (توقيعات أصلية لكل شخص)، دون الحصول على أية عينة لإمضاء مزور. الطريقة المقترحة تستند أساسا على توزيعات أطوال القطع، والتي تقارن بالطرق والخصائص الأغلب شيوعا والأكثر كفاءة حاليا. يتم إجراء التصنيف باستعمال الفواصل ذات الهامش الواسع بدرجة واحدة (OC-SVM). النتائج التجريبية المحصلة والمنجزة على صور التوقيعات الموافقة لعدد 881 موقع من القاعدة GPDS 960 تثبت أن الطريقة المقترحة تسمح بالحصول على كفاءة وفعالية عالية.

Abstract

The work presented in this manuscript can be placed within the field of document analysis and recognition, and more precisely, the off-line signature verification and forgery detection. The objective is to enhance the capabilities of automatic signature verification systems allowing them to work in a realistic fashion by training them the way humans are trained, by using, merely the positive samples without getting any forged ones. The proposed method is based on a set of run-length features which are compared with the well-known state-of- the-art features. Classification is carried out using One-Class Support Vector Machines (OC-SVM). The experimental results obtained on Signature images corresponding to the 881 writers of the GPDS960 database show that the proposed scheme achieves interesting performances.

Résumé

Le travail présenté dans ce manuscrit se situe dans le domaine de l'analyse et la reconnaissance de documents, et plus précisément, la vérification hors-ligne de signatures et la détection de falsification. L'objectif est de renforcer les capacités des systèmes de vérification de signature en leur permettant de travailler de façon réaliste en les formant de la même la manière dont les humains sont formés, à savoir, en regardant seulement les échantillons positifs (signatures authentiques de chaque personne) sans accès à aucun échantillon de signature falsifié. La méthode proposée est basée sur des distributions de longueurs de segments qui sont comparées avec les méthodes les plus connues et les plus performantes de l'état de l'art. La classification est réalisée en utilisant les séparateurs à vaste marge One-Class (OC-SVM). Les résultats expérimentaux obtenus sur des images de signature correspondant aux 881 signataires de la base GPDS 960 montrent que la méthode proposée permet d'obtenir des performances intéressantes.

Dedication

قال تعالى: { يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَ الَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ} [المجادلة : 11]

قال رسول الله صلى الله عليه و سلم: ((إن الله و ملائكته و أهل السموات و الأرض حتى النملة في جحرها و حتى الحوت في البحر ليصلون على معلمي الناس الخير)),

I thank God all the merciful, for the great blessings he gave us.
Thanks to my mother and my father, source of tenderness and factory of manhood.

I dedicate this modest work, to the pure spirits of my dear brothers
***Tarek** and **Sofiane**, whose I miss very much. I never forgot you my brothers,*
Even for a second, you're always present in my heart, in my mind and
in my imagination,
That God the most merciful gathers us in his paradise,

Thanks to everyone I have known.

Walid BOUAMRA

Acknowledgement

I would like to gratefully and sincerely thank ***DR. Chawki DJEDDI*** *for his guidance, understanding, patience and most importantly his friendship during my post-graduation studies at LARBI TEBESSI UNIVERSITY. His supervision was paramount in providing a well-solid ground for research and investigation to reach the objective and fulfill the goal and came up with such a masterpiece. I would like also to express my best hopes and wishes for him in his professional and personal life,*

Secondly, I'll never forget the kindness and the generosity of my cousin ***Abdullah BOUAMRA****, my teacher in English and in etiquette. I hope for him a good homeliness where he is. That God protects his wife and his little prince "Mahdi ",*

I would like also to thank all my teachers, members and staff of the faculty of mathematics and computer science as well as all the staff of LARBI TEBESSI University,

Finally I would like to express my respect and appreciation for my fellow students and friends that I proudly consider them as a valuable part of my life,

Without forgetting to thank everybody that helped me to make a step in my life,

Thank you everyone...

Walid BOUAMRA

Dedication

In the name of Allah, all Merciful, the Most Merciful. Praise be to Allah gave me the strength and courage to carry out this work.

I take this opportunity to extend my sincere thanks to our Dr. Chawki Djeddi for their help, their availability.

Also I dedicate this modest work to the most beautiful pearls: My Mother and Father who have always encouraged me to succeed in my studies give me the will and overcome all obstacles.

In addition, I dedicate this work all my colleagues of (SYM 2016) and all my friends.

To those I love ... and those who love me.

Mohamed El Amine ZOGHBI

Acknowledgement

Praise be to God, who made his grace deeds

Praise be to God, who help us to complete this work

I thank all those who helped make this work and specifically ***Dr. Djeddi Chwaki*** *who has always remained a firm support for me during this entire study endeavor. His advice, correction, and encouragement.*

I would like also to thank all my teachers, members and staff of the faculty of mathematics and computer science as well as all the staff of LARBI TEBESSI University

In the last but not the least, I really really admire the support of my family. Without this support it was never possible for me to think of this day. A huge bouquet of love, thanks, and gratitude for my family.

Mohamed El Amine ZOGHBI

Contents

Figures list

Tables list

Chapter 1

General Introduction

General Introduction

The last few years have witnessed a significant increase in research in different areas of biometrics. Notable advancements in this area have resulted in many biometric modalities such as palm vein recognition [1], face recognition [2], palm print recognition [3], fingerprint recognition [4], DNA recognition [5], speaker recognition [6], keystroke dynamics authentication [7] and gait recognition [8]. These recognition modalities can be divided into *physiological* and *behavioral* biometrics [9].

Physical biometrics employs some physical characteristics of individuals for their identification. Behavioral biometrics, on the other hand, is based on the behavioral traits learnt and acquired over time, which are exploited for authentication purposes.

Despite tremendous development in different biometric modalities, signatures have remained the most widely accepted authentication mechanism in legal documents and financial transactions. Automatic signature verification has remained an attractive pattern classification problem for several decades [10], [11], [12].

The principal task of a signature verification system involves judging whether the input signature image is genuine or forged. There exist three types of forgeries, which are linked to both inter, and intra writer variations [21]. The first type involves random forgeries where signatures of a writer different from that of the learned signature model are presented to the system. The second type concerns simple forgeries where the forger has knowledge of the name but not of the signature of an individual. The third type involves skilled forgeries where the forger imitates the genuine signature of an individual.

Recent advancements on this problem have been summarized in a number of survey papers [13], [14]. Signature verification can be performed using a writer-dependent approach where a separate classifier is trained for each writer or, a writer-independent approach where a single classifier is trained on genuine and forged signatures of all individuals in the database [15]. Based on the technique employed, signature verification methods are also categorized into static and pseudo-dynamic methods. The former techniques rely on extracting geometrical measures from the signature while the later

estimate dynamic information from the signature image [16]. Signature modeling has been effectively carried out using hidden Markov models (HMM) [17] and graph models [18], [19]. For matching, Dynamic Time Warping (DTW) has been most widely used with function features while Support Vector Machines (SVMs) have been very effective on parameter features [20].

Because of all these advancements, and newly, organization of International signature verification competitions [22, 23, 24 and 25], has been a regular activity in conjunction with the ICDAR and ICFHR, the two most notable publishing platforms for the document analysis and recognition community. The increasing number of participants in these competitions from all over the world speaks of the tremendous research attention this problem has continued to attract.

The aim of this thesis is to introduce a novel offline signature verification method for detection of skilled forgeries. The proposed method comprises three main stages: feature extraction, design and classification (signature verification). Feature extraction relies on extracting textural measures, these measures include run-lengths distributions extracted from black and white pixels of the signature image. During classification, features of the genuine signature image are matched with the features corresponding to the signature image in question. We have used One-Class Support Vector Machines (SVMs) as a classifier. The proposed method evaluated on the 881 writers of the GPDS960 signature corpus [26] reports very promising results.

We have also taken into consideration the increasing demand from forensic signature analysts, to enhance the capabilities of automatic signature verification systems allowing them to work in a realistic fashion by training them the way humans are trained, i.e., only by looking at the positive specimens (genuine signatures of each person) without access to any forged samples. The proposed method consists to classify a questioned signature as being genuine or not. From the perspective of pattern recognition, this corresponds to the scenarios where classifiers are to be trained with training data of one class (genuine signatures only) for every individual. This thesis is structured as follows:

The second chapter discusses the concepts of biometrics (definitions, some biometric modalities, comparison of biometric technologies) followed by the presentation of automatic signature verification systems and their different steps (preprocessing, feature

extraction and classification), in the same chapter we presents the state-of-the-art of signature verification systems as well as a summary of organized signature verification competitions accompanied with an overview on the feature extraction methods. We will also describe some performance measures.

In the third chapter, we introduce a novel signature verification system for detection of skilled forgeries. The proposed system comprises three main stages: feature extraction, design and evaluation (signature verification). Feature extraction relies on extracting run-lengths distribution from the signature images. The last section of the third chapter presents the experimental settings, the evaluation protocols and the obtained results.

Finally, in the last part of this thesis, we present the concluding remarks with a discussion on possible research directions on this problem

Chapter 2

Signature Verification, Concepts and tools

1. Introduction to Biometrics

1.1. Definition

Biometrics is the technique used to recognize people from their physical and behavioral characteristics.

There are different physical or behavioral means, which allow a recognition of the individual. As already mentioned, the impression, iris, face and shape of the hand are physical means called biometric modalities'. One can also cite the example of the hand vein and retina. In terms of behavioral modalities, include the signature (dynamic or static).

Biometric features are an alternative to old ways Verification of identity. The advantage of these biometric is to be universal, That is to say, present in all people is identify. On the other hand, they are Measurable and unique: two people cannot have the same Feature. They are also the permanent, which means they do little or vary over time. For the collected features can be qualified arrangements Biometric [28], they must be:

- Universal (exist in all individuals).
- Unique (Able to differentiate an individual compared to another).
- Permanent (allow evolution over time).
- Recordable (collect characteristics = an individual with his agreement).
- Measurable (allow future comparison).

The main interest of biometrics is to recognize and identify =automatically the identities of individuals by using their physiological characteristics alternatively, behavioral.

The physiological characteristics may include the face, the iris, Fingerprints, hand geometry. Behavioral characteristics include voice, Signature, etc.

2. Biometrics Modalities

No single biometrics could answer effectively to the needs of all verification systems or applications. A number of techniques have been biometrics proposed, analyzed, and evaluated.

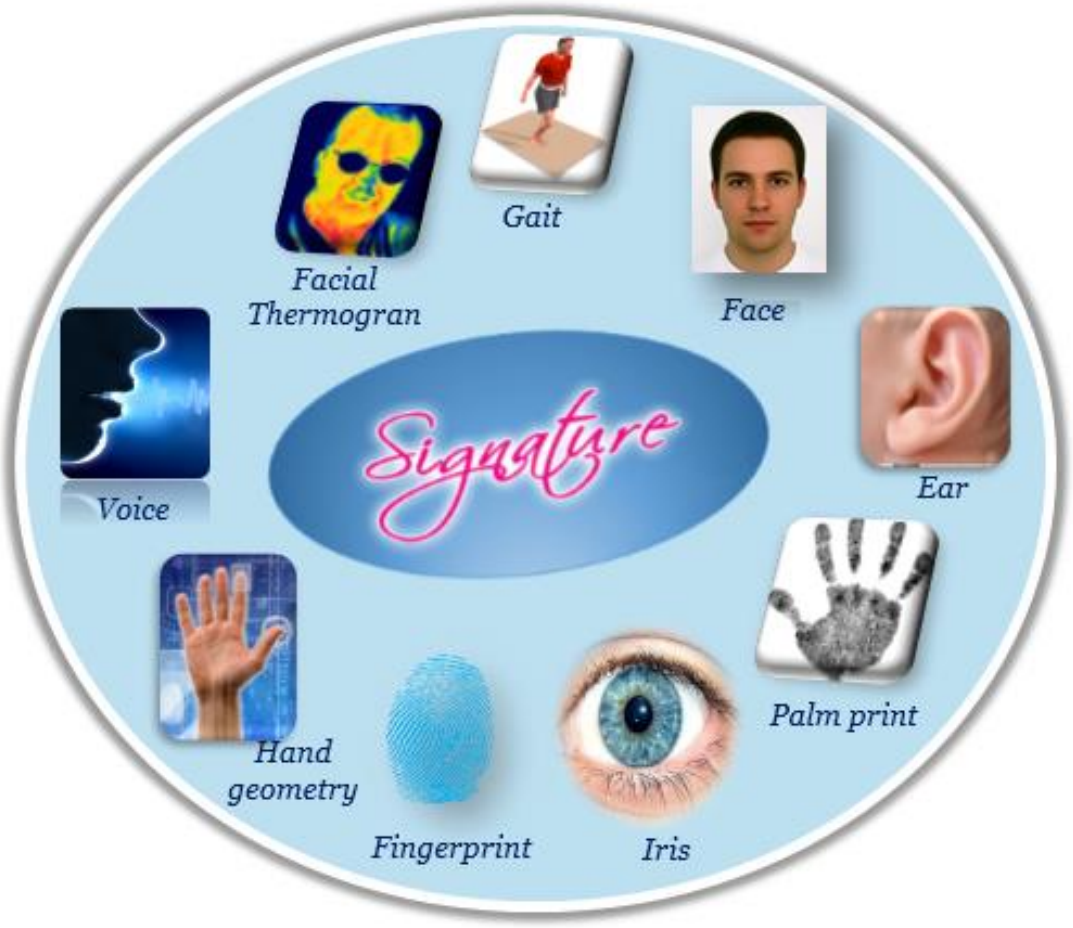


Figure 2.1: Examples of biometric traits that can be used for authentication.

Each biometrics to its strengths and limitations, and accordingly, each biometrics is used in a particular application. For the physical, we describe the face recognition, fingerprints, the hand geometry and iris. For behavioral characteristics, we describe the biometrics-based voice and signature. (*Figure 2.1*). There are other biometric methods like A.D.N, vein pattern, the shape of the ear, the strike rate on a keyboard...etc.

There are other biometric methods like A.D.N, the shape of the ear, the strike rate on a keyboard, approach, which will not be spoken in this chapter.

2.1. The Fingerprints

Currently, the fingerprint recognition is the method the most commonly used biometric. Fingerprints are made up of lines locally parallel with singular points (minutiae) and constitute a unique pattern, universe and permanent.

For an image of a fingerprint, technological advances helped to automate the task using integrated sensors, replacing the use Traditional ink and paper. [29].

These sensors operate according to different mechanisms measuring (pressure, electric field , temperature) to measure the footprint of a fixed finger positioned on the latter (matrix sensor) or movement (sensors scanning).(*Figure2.2*).

- *Domains of application*

Fingerprint image based is the most successful and popular method for individual identification because it's very easy to use and the cost of small-size acquisition devices is low, allowing its use in different domains e.g., electronic commerce, physical access, PC logon, law enforcement applications, background checks for employment or licensing, airports, access control to secure areas (*Figure 2.2*),..etc.

Figure 2.2: The fingerprint sensors

2.2. The voice

Of all the human traits used in biometrics, voice is that humans learn to recognize at an early age. The speaker recognition systems can be divided into two categories: the dependent systems of the spoken text and systems independent of the text (*Figure 2.3*). In the first case, the user must use a text (A word or phrase) fixed predetermined during the learning sessions and recognition.

Figure 2.3: the Voice is a kind of biometry

While for a system independent of the text, the speaker speaks freely without predefined text. Recognition is growing, as it does not require expensive equipment, since most PCs today are equipped with a microphone.

However, the poor quality and ambient noise may influence the verification and hence reduce its use in biometric systems. In a speaker, recognition system. The signal is first measured and then divided into several frequency channels pass stringer then the important characteristics of the speech signal are extracted from each band.

2.3. The Iris

Iris recognition is the process of recognizing a person by analyzing the random pattern of the iris (*Figure 2.4*).

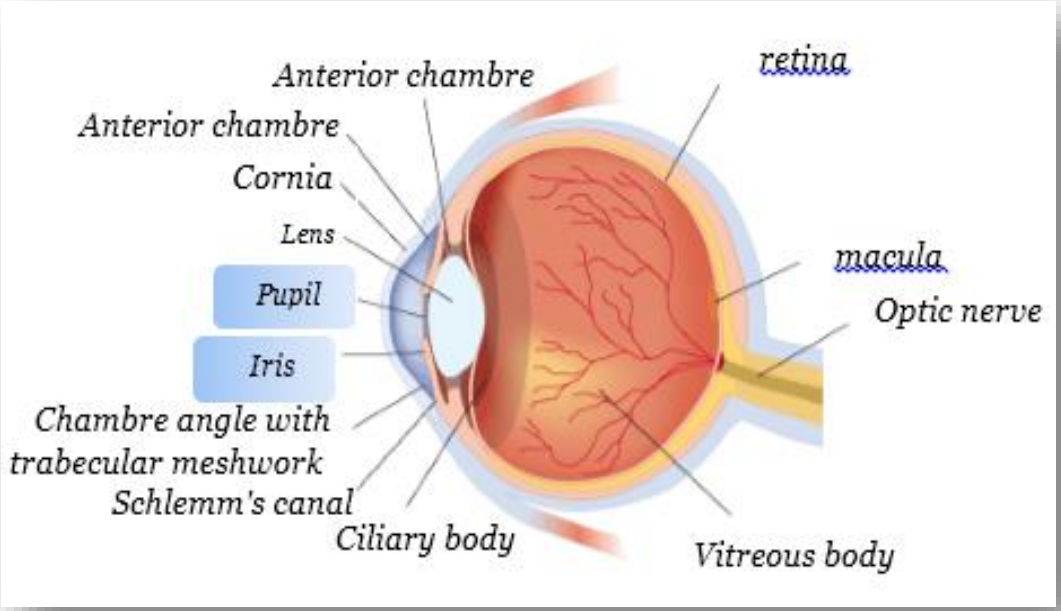


Figure 2.4: Anatomy of the Eye

The automated method of iris recognition is relatively young, existing in patent only since 1994.

The iris is a muscle within the eye that regulates the size of the pupil, controlling the amount of light that enters the eye. It is the colored portion of the eye with coloring based on the amount of melatonin pigment within the muscle as shown in *Figure 4*. Although the coloration and structure of the iris is genetically linked, the details of the patterns are not.

The iris develops during prenatal growth through a process of tight forming and folding of the tissue membrane. Prior to birth, degeneration occurs, resulting in the pupil opening and the random, unique patterns of the iris. Although genetically identical, an individual's iris is unique and structurally distinct, which allows it to be used for recognition purposes. Iris Recognition Process The process of capturing an iris into a biometric template is made up of three steps:

1. Capturing the image.
2. Defining the location of the iris and optimizing the image.
3. Storing and comparing the image.

2.4. Face Recognition

Major advancements and initiatives in the past ten to fifteen years have propelled face recognition technology into the spotlight. Face recognition used for verification Considerate of the public's social and privacy concerns. Today, face recognition technology is being used to combat passport fraud, support law enforcement, and identify missing children.

Face detection in the image is an indispensable and crucial treatment before the recognition phase. Indeed, the face recognition process can never become completely automatic if it was not preceded by an effective detection step.

The treatment is to look at a picture the faces and the position extract in the form of a set of thumbnails in order to facilitate further processing. A face is considered correctly detected if the extracted thumbnail size does not exceed 20% of the actual size of the facial region, and it essentially contains the eyes, nose and mouth [30, 31] (*Figure 2.5*).

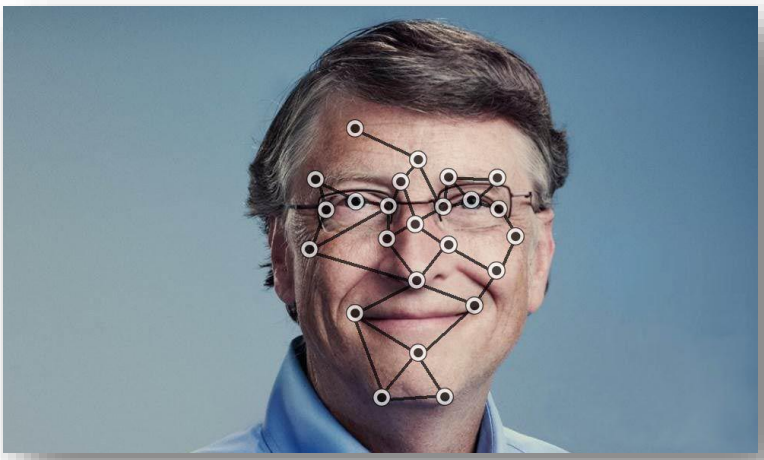


Figure 2.5: Face recognition

2.5. The Hand geometry

The hand geometry is a new biometric technology. As the name suggests, it is to analyze and measure the shape of the hand, that is to say the measure length, width and

height of the hand of a user and create a 3-D image. Of the infrared LEDs and a digital camera is used to acquire data hand. (*Figure 2.6*).

Figure 2.6: hand geometry sensor.

Hand geometry systems have the longest implementation history of all biometric modalities. David Sidlauskas developed and patented the hand geometry concept in 1985.[32] and the first commercial hand geometry recognition systems became available the next year.[33] The 1996 Olympic Games implemented hand geometry systems to control and protect physical access to the Olympic Village.' Many companies implement hand geometry systems in parallel with time clocks for time and attendance purposes. Walt Disney World has used a similar "finger' geometry technology system for several years to expedite and facilitate entrance to the park and to identify guests as season ticket holders to prevent season ticket fraud.[34]

This technology offers a reasonable level of accuracy and is relatively difficult to use. However, twins or people with forms of close hand can easily deceive it. The most popular uses hand geometry include record keeping and access control. By cons, capture systems hand geometry are relatively large and heavy, which limits their use in other applications such as the authentication in embedded systems: mobile phones, cars, laptops, etc.

2.6. The Signature

Signature: is the way a person signs his/her name and its known to be unique and characterizes that individual. Signature is a behavioral biometric that changes over a period of time and it is easily influenced by the physical and emotional conditions of the signatories. Although signatures require contact with the writing instrument and an effort

on the part of the user, they have been accepted in government, legal, and commercial transactions as a method of authentication. (*Figure 2.7*)

Figure 2.7: different signatures froms (GPDS960)

3. Comparison of biometric technologies

There is no perfect biometric system. And we find that the International Biometric Group conducted a comparison of different technologies based on four criteria (*Table 2.1*):

- Effort: effort by the user during authentication.
- Intrusion: Information on the acceptance of the system by users.
- Cost: cost of technology (readers, sensors, etc.).
- Accuracy: effectiveness of the method (related to the error rate).

Technologies	Advantage	Disadvantages
Digit prints	Cost, average ergonomics	Optimal quality measuring devices, (Reliability), average acceptability possibility of attacks (Persistence of the footprint ...)
Hand shape	Ergonomic, good acceptability	Cumbersome system, cost, possible disruption by injuries and the authentication of the same family
Face	Cost, compact, good acceptability	Twins, psychology religion, disguise Vulnerability to attacks
Iris	Reliability	Use low acceptability of lighting stress
Voice	Ease	Vulnerable to attacks
Signature	Ergonomic	Depending on the person's condition, Reliability

Table 2.1: Advantages and disadvantages of various biometric technologies.

4. Handwritten signature identification and Verification

Is the process of recognizing or verifying an individual's identity using his handwritten signature. The identification (or recognition) is finding the signature owner while verification is confirming or denying claimed identity by deciding whether the signature is genuine or forgery. During this process the samples of signature are taken then compared to the samples of signatures stored in the database.

In case of identification the comparison is one-to-many process and the result usually between 0 and 1, which represents a matching ratio.

In case of verification the comparison is one-to-one process and the result is Boolean (yes/no).

Signature verification has several advantages over other technologies as an identity verification mechanism.

Firstly, signature analysis can only be applied when the person is conscious and willing to write in the usual manner, although it is possible that individuals may be forced to submit the handwriting sample. To give a counter example, a fingerprint may also be used when the person is in an unconscious (e.g., drugged) state [35].

Secondly, Measurement of signature characteristics is noninvasive (compare this with other potential techniques such as iris scanning) and has no negative or undesirable health connotations (as might be the case with, say, fingerprint checking, which is often considered to raise civil liberties issues and which, in use, involves direct physical contact with a possibly contaminated surface) [36].

Thirdly, most of the new generation of portable computers and personal digital assistant (PDAs) use handwriting as the main input medium.

Handwritten signatures come in many different forms and there is a great deal of variability even in signatures of people that use the same language. Some people simply write their name while others may have signatures that are only vaguely related to their name and some signatures may be quite complex while others are simple and appear as if they may be forged easily [37]. It is also interesting that the signature style of individuals relates to the environment in which the individual developed their signature. For example, people in the United States tend to use their names as their signature whereas Europeans attend away from directly using their names. Systems which rely directly on the American style of signing may not perform as well when using signatures of Europeans or signatures written in different languages [38].

- ***Genuine signature***

It's known that signatures from same person are never the same, they differ in both globally and locally and may also differs in scale and orientation [38].circumstances in which the signature was written also has impact like: size of signing space, careless signatures, different pens, physical and psychological condition of the person, surface (in case of signing a paper), in fact even asking people their signatures for training/testing set use ,that may produces self-conscious, unnatural signature. In addition, a person's signature often changes over time but those changes are slight and can be over came by updating that person's signatures in database.

➢ ***Forgery signature:***

The process of forging a signature, if it is to be successful, involves a double process requiring the forger to not only copy the features of the writing imitated but must also hiding the writer's own personal writing characteristics[38]. In fact the over effort on the signature is what makes it mostly Forgery.

Some signature experts note that if two signatures of the same person written on paper were identical they could be considered forgery by tracing, and of course there is the skilled forgeries those who can deceive the system

4.1. Types of forgeries

The objective of signature verification system is to discriminate between two classes' **authentic** and **forgery**. In parallel with real life scenarios, research databases define three types of forgeries:

1. **Skilled forgery** refers to a forgery, which is signed by a person who has had access to some number of genuine signatures and practiced them for some time. Often, the imposter is simply one of the enrolled users who has been asked to forge the signature of another user. [39].
2. **Random forgery** is typically collected from other people's real signatures, simulating the case where the impostor does not even know the name, nor shape of the target signature and hence uses their own in forgery. [40].
3. **Simple forgery** here in this type the forger has no access to the sample of the signature but he/she knows the author's name and the forger produces the signature in his/her own style.

4.2. Data acquisition

Depending on the signature acquisition method used, automatic signature verification systems can be classified into two groups: **online** and **offline** signature.

- Offline(Static)
- Online (Dynamic)

In the verification of **static** signature, only the geometric shapes of the signature is used to authenticate a person (*Figure 2.8*). In this approach, the extraction of dynamic

data is not so easy because the input to this system will be a 2D image of the signature and the dynamic information will not be available.

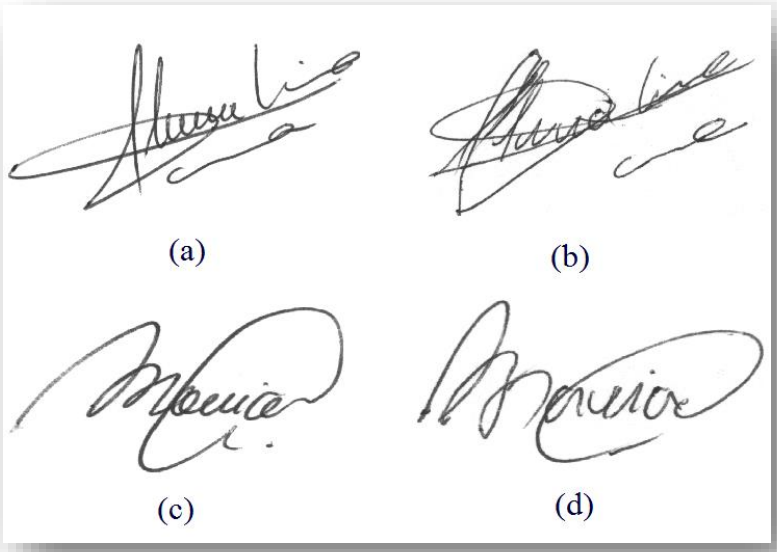


Figure 2.8: Offline signature samples : genuine (a), skilled forgery (b) genuine (c), skilled forgery (d) of GPDS-960.

For the second approach (Dynamic) to the verification of signatures, it uses, in addition to the geometric shape, the dynamic characteristics such as acceleration, speed and trajectory profiles of signing. The signature evolves over time and is influenced by the physical and emotional conditions of the person. The two types of variation found in the signatures are: Inter personal and Intra personal variability. The variation among the signatures of the same person is called as Intra personal variation.

This variation can be due to:

- Age.
- Illness.
- The wounds.
- The time constraints.
- Drugs.
- Temperature.

The variation between the originals and the forgeries is called Inter personal.

4.3. Image processing

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information

from it, i.e. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image.

4.3.1 Image pre-processing

The off-line signature recognition requires applying several preprocessing steps on both training and testing sets to make it ready for feature extraction process. In preprocessing stage, the image of the signature goes through: scanning in gray (or converted into gray scale), then background elimination, noise reduction, size normalization and thinning (*Figure2.9*).

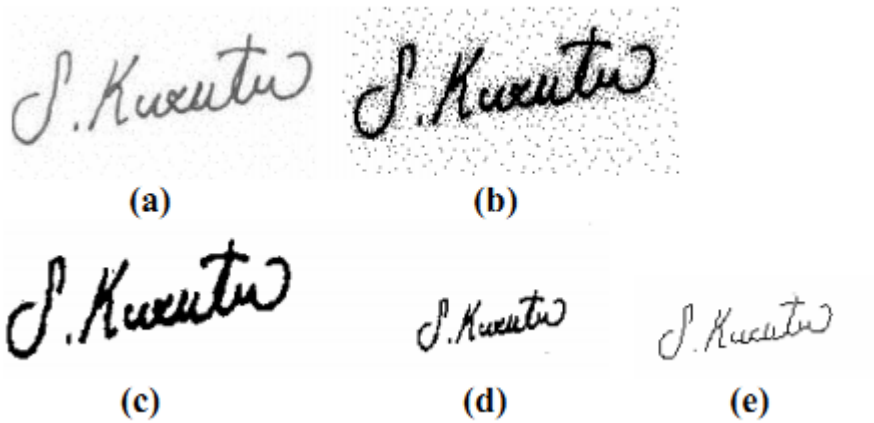


Figure 2.9: Signature Pre-processing
(a) scanned signature, (b) after binarization, (c) after noise reduction, (d) after size normalization, (e) after thinning.

In preprocessing stage, the image of the signature goes through: scanning in gray (or converted into gray scale), then background elimination, noise reduction, size normalization and thinning.

4.3.2. Filtering

Is a technique for modifying or enhancing an image. For example, you can filter an image to emphasize certain features or remove other features. Image processing operations implemented with filtering include smoothing, sharpening, and edge enhancement. (*Figure 2.10*)

Figure 2.10. Example of filtering image

Besides, filtering is a *neighborhood operation,* in which the value of any given pixel in the output image is determined by applying some algorithm, to the values of the pixels in the neighborhood of the corresponding input pixel.

4.3.3. Image binarization

The binarization it means converts an image of up to 256 gray levels to a black and white image. Frequently, binarization is used as a pre-processor before OCR. (*Figure 2.11*)

Figure 2.11: Binarized image

4.3.4. Image Enhancement

Image enhancement is among the simplest and most appealing areas of digital image processing. (*Figure 2.12*).

Figure 2.12: Image Enhancement

The idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. Such as, changing brightness etc.

4.3.5. Segmentation

Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. (*Figure 2.13*)

A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually.

Figure 2.13: Image segmentation

4.3.6. Morphological image processing

Is a collection of non-linear operations related to the shape or morphology of features in an image. According to Wikipedia, morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images.

4.2. Feature extraction:

In this section, we describe the extraction methods features used in signature verification. Feature extraction step reduces the dimension of original signature images while preserving and extracting the important information encoded in the image.

A carefully selected set of features will transform the images so that it becomes easier to distinguish between genuine and forgery classes.

4.2.1 Global features:

The Global features describe the entire signature image such as width, height, aspect ratio. These features are used in combination with other features. These features are less sensitive to noise, and can obtained by considering all points within a region, or by points in the boundary of a region (Signature area). It includes characteristics of regions in images, moments, Fourier descriptors, perimeter[41].Global feature means looking the whole image while local means focusing on something.

4.2.2 Local features:

Local features refer to a pattern or distinct structure found in an image, such as a point, edge, or small image patch. They are usually associated with an image patch that differs from its immediate surroundings by texture, color, or intensity. What the feature actually represents does not matter, just that it is distinct from its surroundings. Examples of local features are blobs, corners, and edge pixels.

4.3. Classification:

Classification is a type of supervised machine learning in which an algorithm "learns" to classify new observations from examples of labeled data.

To explore classification models interactively, use the Classification Learner app. For greater flexibility, you can pass predictor or feature data with corresponding responses or labels to an algorithm-fitting function in the command-line interface.

5. Classifiers

Many different classifiers have been applied to offline signature verification so far. Solar et al. [38] use Bayes classifier. We mention some known classifiers:

5.1. Neural networks approach

Neural Networks is a type of artificial intelligence that attempts to imitate the way a human brain works. It consists of interconnected processing elements neurons that work in parallel to produce an output function from an input [42]. this approach is widely used in signature verification systems, due to its power(have the ability to learn complex nonlinear input-output relationships), ease of use(as NNs learn by example it is only necessary for a user to gather a highly representative data set and then invoke training algorithms to learn the underlying structure of the data(off line approaches),capabilities in learning and generalizing(use sequential training procedures and adapt themselves to the data).the NNs are very robust against failure or error because they can function even if some neurons are not functioning anymore.

5.2. K-Nearest Neighbors

K nearest neighbors is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique. [43]

5.3. Hidden Morkov Models (HMM)

A hidden Markov model (HMM) is one in which you observe a sequence of emissions, but do not know the sequence of states the model went through to generate the emissions. Analyses of hidden Markov models seek to recover the sequence of states from the observed data.

5.4. Support vector machines (SVMs)

Are a set of supervised learning methods used for classification, regression and outlier's detection.

An SVM classifies data by finding the best hyper plane that separates all data points of one class from those of the other class. [27].

The best hyper plane for an SVM means the one with the largest margin between the two classes.

Margin means the maximal width of the slab parallel to the hyper plane that has no interior data points.

✓ **Advantages of support vector machines**

Effective in high dimensional spaces.

Still effective in cases where number of dimensions is greater than the number of samples.

Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.

Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.

✓ **Disadvantage of support vector machines**

If the number of features is much greater than the number of samples, the method is likely to give poor performances.

6. Approaches

The design of an (HSVS) can be performed according two approaches: writer-dependent (WD) and writer-independent (WI).

The (WD) models extract features from genuine signatures of a specific writer and are trained for that writer. The questioned signature is compared against the model for that writer.

This is the standard approach to signature verification [44].Based on a writer-independent approach to determining whether two handwritten documents not just signatures were written by the same person or not.

In the writer independent model, the probability distributions of within writer and between-writer similarities, over all writers, are computed in the training phase. These distributions are used to determine the likelihood of whether a questioned signature is authentic.

7. Overview

7.1 Dimensional autoregressive coefficients (ARCoeff)

First, two-dimensional auto regression models were introduced by K. Deguchi for the representation of images as well as the characterization of textures. Since then, they have been successfully applied to the segmentation and modeling of textures. In a relatively recent study, auto regression models have been adapted to characterize and identify the writers of manuscripts and to determine their sex.

For our task of signature verification and falsification detection, we characterize a signature given by a set of coefficients dimensional auto regression extracts the binary image of the signature.

7.2 Local Binary Pattern (LBP)

The original LBP operator forms labels for the image pixels by thresholding the 3 x 3 neighborhood of each pixel with the center value and considering the result as a binary number. [45]

7.3 Local Derivative Pattern (LDP)

Is a general framework to encode directional pattern features based on local derivative variations. The $n^{(th)}$-order LDP is proposed to encode the (n-1) (th) order local derivative direction variations, which can capture more detailed information than the first-order local pattern used in local binary pattern. [46]

7.4 Edge-direction distribution (ED12)

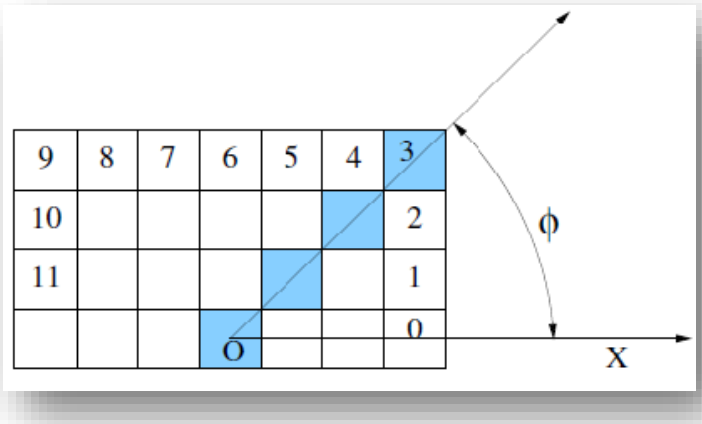


Figure 2.14: Example of extraction of edge-direction distribution using 12 directions.

Feature extraction starts with conventional edge detection (convolution with two orthogonal differential kernels, we used Sobel, followed by thresholding) that generates a binary image in which only the edge pixels are "on". We then consider each edge pixel in the middle of a square neighborhood and we check (using logical AND operator) in all directions emerging from the central pixel and ending on the periphery of the neighborhood for the presence of an entire edge fragment. (*Figure 2.14*)

All the verified instances are counted into a histogram that is finally normalized to a probability distribution p (ϕ) which gives the probability of finding in the image an edge fragment oriented at the angle ϕ measured from the horizontal.

7.5. Edge-hinge distribution (EH5)

The method of feature extraction is similar to the one previously described, but it has added complexity. *(Figure 2.15)* The central idea is to consider in the neighborhood, not one, but two edge fragments emerging from the central pixel and, subsequently, compute the joint probability distribution of the orientations of the two fragments.

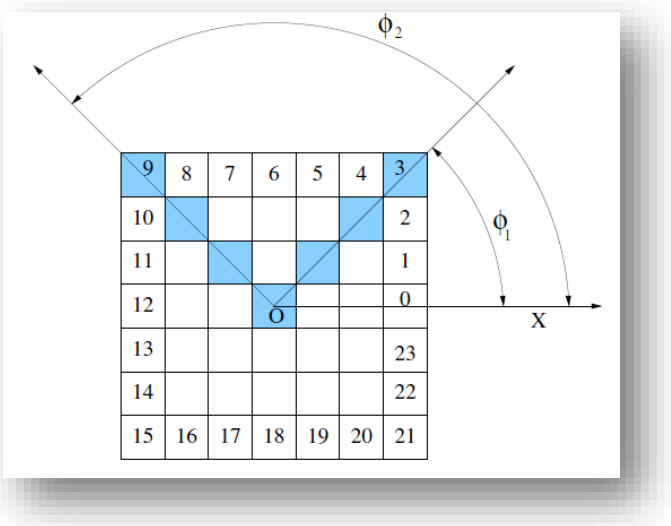


Figure 2.15: Extraction of edge-hinge distribution.

7.6. Local Phase Quantization (LPQ)

The LPQ [47] operator is applied to texture identification by computing it locally at every pixel location and presenting the resulting codes as a histogram. Generation of the codes and their histograms is similar to the local binary pattern method.

7.7. Binarized Statistical Image Features (BSIF)

The method computes a binary code string for the pixels of a given image. [48]. The code value of a pixel is considered as a local descriptor of the image intensity pattern in the pixel's surroundings. Further, histograms of pixels' code values allow characterizing texture properties within image sub regions.

7.8. Histogram of oriented gradients (HOG)

Histogram of oriented gradients [49]. It involves first computing the gradient information at each pixel inside a particular grid zone (either Cartesian or Polar). Next, histogram of gradient orientations in that zone is computed we can conclude that HOG features utilize a coarse shape of signature by modeling local directions of gradients with histograms.

8. Performance measures

- **False Rejection Rate (FRR) and False Acceptance Rate (FAR):**

To evaluate the performance of a signature verification system, two parameters are generally used, False Rejection Rate (**FRR**) and a False Acceptance Rate (**FRR**).

The (**FRR**) is the ratio of the number of genuine test signatures rejected to the total number of genuine test signatures submitted which corresponds to the genuine signature rate rejected by the system (*Figure 2.16*).

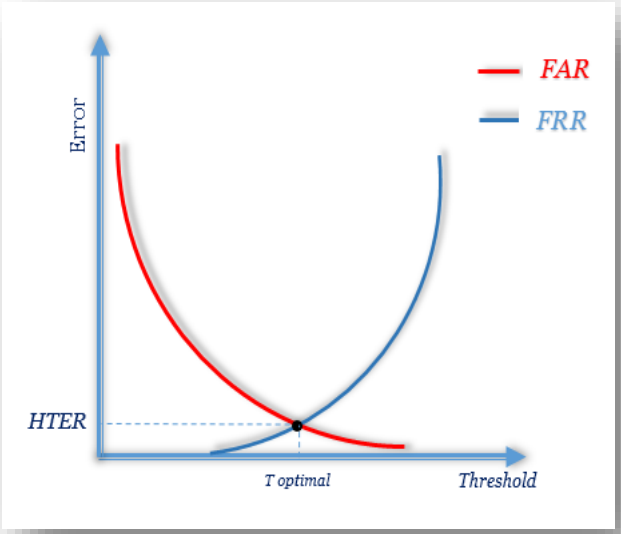


Figure 2.16: the principle for selecting the optimal decision threshold from FAR and FRR

$$frr = \frac{totalnumberofGenuinesrejected}{totalnumberofGenuinessubmitted}$$

- The (**FAR**) is the ratio of the number of forgeries accepted to the total number of forgeries submitted which corresponds to the fictitious signature rate, accepted by the system.

$$far = \frac{totalnumberofforgeriesaccepted}{totalnumberofforgeriessubmitted}$$

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

It describes the point at which genuine and forged error rates are closest to zero. EER can be represented as a percentage with time/unit factors. EER is not useful in assessing actual system performance, but can be helpful as a first-order performance indicator for verification systems.

- **The Accuracy (ACC) :**

It simply measures the percentage of correct judgments with respect to all the judgments passed by a verifier, as given in Equation.

$$Accuracy = \frac{numberofcorrectjudgments}{numberoftotaljudgemetns}$$

- **Cost of Log Likelihood Ratio (Cllr) and $\hat{C}_{llr}^{min}$:**

Cllr is interpreted as an average decision cost for all prior probabilities and costs involved in the decision process. It is an average over costs and priors, and therefore is not giving the performance for a given value of the prior, but for an average of all possible priors. Cllr value can be expressed as the sum of a minimum Cllr values referred to as discrimination loss. $\hat{C}_{llr} = \hat{C}_{llr}^{min} + \hat{C}_{llr}^{cal}$.

- **The min Cllr ($\hat{C}_{llr}^{min}$):**

The $\widehat{C}_{llr}^{min}$ is a measure for the discrimination performance of the system, with lower values representing better discrimination ,the $\widehat{C}_{llr}^{min}$ can also be seen as a validity measure of a biometric system, in that it indicates the quality.

9. Conclusion

Biometrics is based on the principle of recognition of physical characteristics. Fingerprints and the range of indices generally covered by biometrics, including iris, hand, voiceprints and signature offer irrefutable proof of the identity of a person because they are unique biological and behavioral characteristics that distinguish one person from another and can only be associated with one person. Peoples are familiar with signature verification.

The signature verification process becomes a primary task, and for this, several methods have been developed to verify the authenticity of signatures, using universal databases, which help for forgeries detection, and by the other hand, to enhance the performance of these methods.

Chapter 3

A novel handwritten signature verification system to detect forged signatures

1. Introduction

Today, there is an increasing demand from different application areas, especially, forensic signature analysts, to enhance the capabilities of automatic signature verification systems allowing them to work in a realistic fashion. Therefore, many competitions have been contributed in order to develop the signature verification systems accuracy. A very important challenge in these verification systems is to train them how humans train, by the employment of genuine signatures as the single existed sample for training. The system must answer by ranking a questioned signature if this signature is authentic or not.

Our contribution consists to investigate the performance of classifiers trained on genuine samples only so that resulting system would closely match real world scenarios.

In this chapter we present a novel signature verification system for detection of skilled forgeries. Feature extraction relies on extracting run-length distributions from the signature images; during classification, features of the genuine signature image are matched with the features corresponding to the signature image in question (*Figure 3.1*). One Class Support Vector Machine (OC-SVM) was used as classifier. The proposed method was evaluated on all the 881 writers of the GPDS960 signature corpus and the reported results were promising.

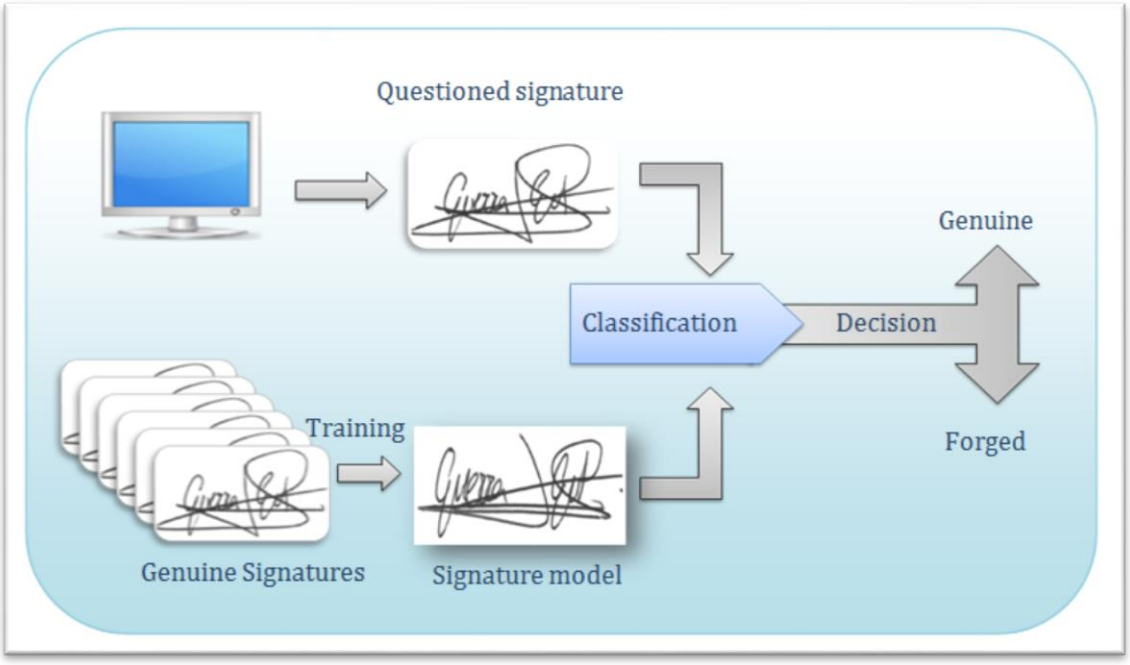


Figure 3.1. Block diagram of a signature verification system.

In the next section, we will present a preview of one class support vector machine, the computation of run-length distributions that we employed to characterize signatures of individuals. For comparison purposes, we also briefly outline few of the well-known state-of-the-art features that have been effectively applied to the signature verification problem. Section IV describes the performance evaluation measures employed in our study, while Section V presents the experimental settings and the realized results, and in section VI we study the system stability. Finally, we present the concluding remarks about the consequences of the results obtained in the precedents sections.

2. Review of the OC-SVM classifier

With binary or multi-class support vector machines, we always have positive and negative examples, i.e. examples and counterexamples. Such information is not available in all instances of application. Sometimes, it is very difficult and costly, may be impossible to find counterexamples that truly represent the negative class. Taking the example of recognizing a particular category of parts by a robot in a factory, it is easy to have sufficient examples of this piece, but it is difficult to have examples of all the different parts. Likewise, for handwriting signatures, counterexamples (forgeries) are not available all the time. It is desirable in such cases to have a decision model to recognize many possible examples of this category and rejects all others.

This problem is called "single-class classification", "Novelty detection" or detection of "newness", it's a practice (training) algorithm developed by Schölkopf et al. [50]. It allows classifying only objects of a single class, and differentiating them from all other possible objects. The classifier groups well the objects, but will consider others as outliers [51], while the decision model knows a set of examples and detects all that is new (in our case well verifying a questioned signature if it is genuine or not).

One class classification distinguishes the target class from all other classes using only training data from the target class, i.e. we have only genuine signatures in training set. The goal is to find a boundary between the examples of the target class from the rest of the space, i.e. a border around the target class that accepts as many examples as possible targets [52]. This boundary is represented by a decision function which is positive within a class S and negative outside in the complement of S: ($\bar{S}$). (*Figure 3.2*) shows an example of separating a class from any other class. In other words, the Schölkopf algorithm, returns a function f that takes the value +1 in a “small” region capturing most of the data vectors, and -1 elsewhere. It could be summarized by mapping the data into a feature space H (Hyper plan), employing an appropriate kernel function, and then trying to separate the mapped vectors from the origin with maximum margin.

$$f(x) = \begin{cases} +1 \; if \; x \in S \\ -1 \; if \; x \in \bar{S} \end{cases}$$

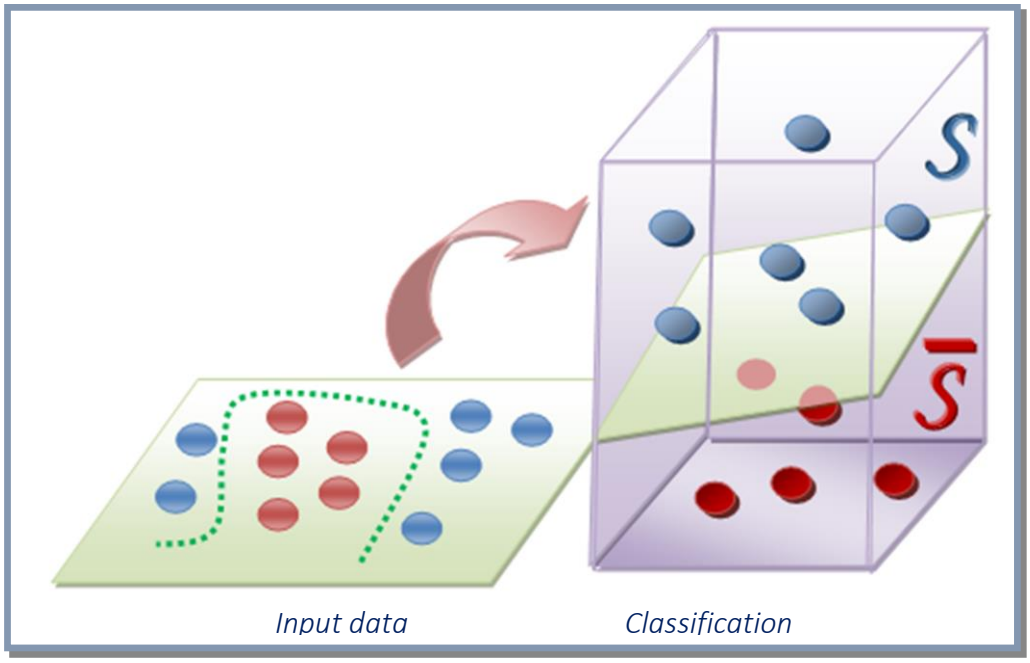


Figure 3.2. One-class SVM classification.

3. Feature Extraction

To characterize the signature of an individual, we have employed black and white run-length distributions [53],[54]. These features have been effectively applied to problems like writer recognition and it would be interesting to study its performance on the more challenging task of signature verification and forgeries detection. For comparison purposes, we have implemented some of the latest state-of-the- art methods that have shown good results on this problem.

Run-length distributions are computed on binary images of signatures taking into account the black pixels which correspond to the ink trace of the signature and the white pixels which correspond to the background of the signature image. A 'run' is defined as a sequence of connected pixels in a given direction all having the same intensity. The run-length matrix is defined as a matrix P_{ij} where the value at (i, j) in the matrix is the number of pixel runs of color i and length j in a given direction. The size of the matrix is $M \times K$ where M represents the number of unique colors (intensities) in the image while K is the maximum possible length of a run in a given direction.

In our study, we consider the horizontal, vertical, left-diagonal and right-diagonal run-lengths on black and white pixels of the binarized signature image. The extraction of the run-length distributions is illustrated for a 9×10 binary image in (*Figure 3.3)*.

The elements of the matrices represent the number of times, runs of different length occur in the four directions. For example, the elements in the first row of the horizontal run-length matrix indicate that for pixel value 0 "Black Pixels", there are 2 runs of length 1, 4 runs of length 2, 3 runs of length 3, 1 run of length 4, 1 run of length 5, and so on in the horizontal direction. The second row indicates the similar values for runs of 1 "white pixels", where we have 4 runs of length 1, 8 runs of length 2, 3 runs of length 3, 2 run of length 4, 1 run of length 5, and so on in the horizontal direction, no runs of length 6, 7, 8, 9 and 10, and lastly we get 2 runs containing 10 pixels in Fig 3.a, and likewise for other directions illustrated in Figures 3.b, 3.c and 3.d.

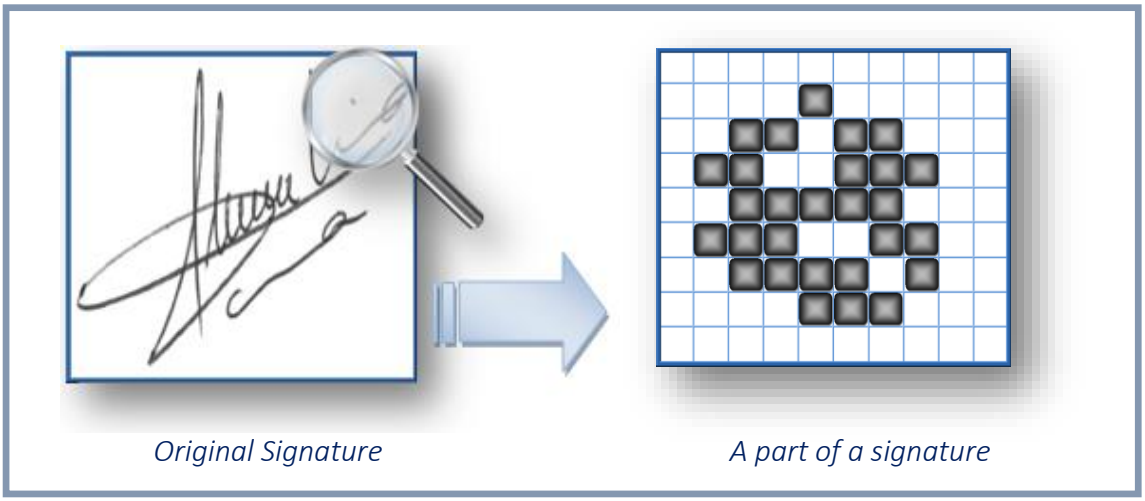
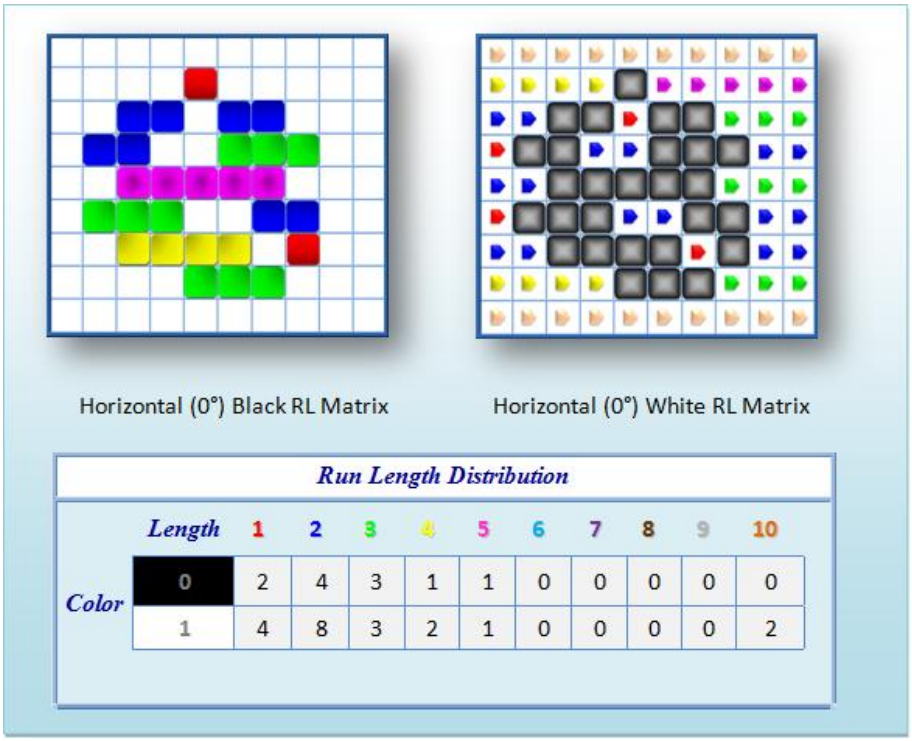


Figure 3.3. Run-lengths computation on a part of a signature image extracted from the GPDS database, (a) Horizontal run-lengths distributions (0°), (b) Vertical run-lengths distributions (90°), (c) Right-diagonal run-lengths distributions (45°), (d) Left-diagonal run-lengths distributions (135°).



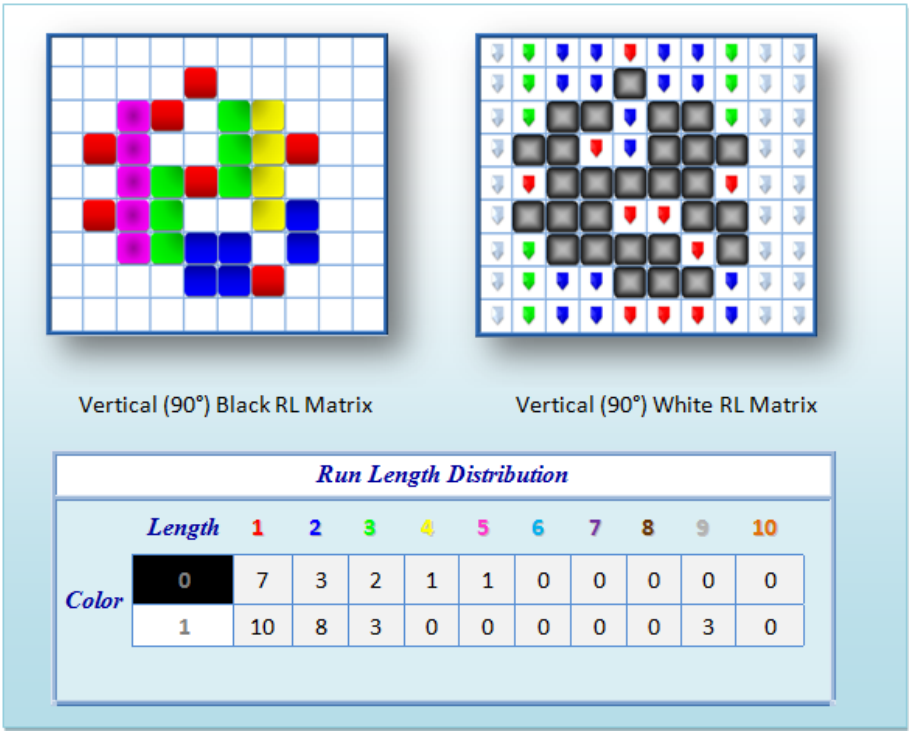
Horizontal (0°) Black RL Matrix

Horizontal (0°) White RL Matrix

Run Length Distribution

Color \ Length	1	2	3	4	5	6	7	8	9	10
0	2	4	3	1	1	0	0	0	0	0
1	4	8	3	2	1	0	0	0	0	2

(a)



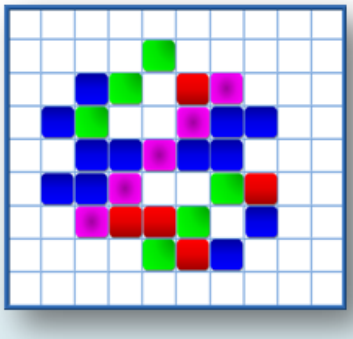
Vertical (90°) Black RL Matrix

Vertical (90°) White RL Matrix

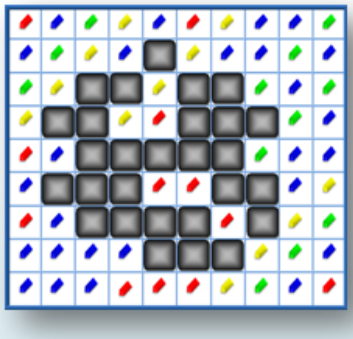
Run Length Distribution

Color \ Length	1	2	3	4	5	6	7	8	9	10
0	7	3	2	1	1	0	0	0	0	0
1	10	8	3	0	0	0	0	0	3	0

(b)



Inclined (45°) Black RL Matrix

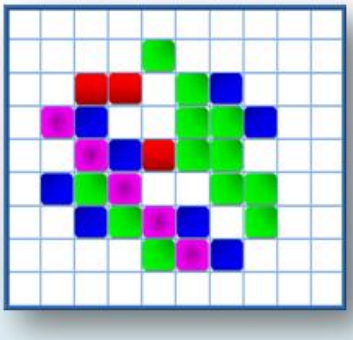


Inclined (45°) White RL Matrix

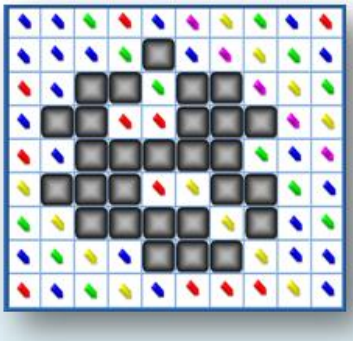
Run Length Distribution

Color \ Length	1	2	3	4	5	6	7	8	9	10
0	5	6	2	0	1	0	0	0	0	0
1	12	13	4	3	0	0	0	0	0	0

(c)



Inclined (135°) Black RL Matrix



Inclined (135°) White RL Matrix

Run Length Distribution

Color \ Length	1	2	3	4	5	6	7	8	9	10
0	3	4	4	0	1	0	0	0	0	0
1	11	11	4	3	1	0	0	0	0	0

(d)

The size of the run-length matrices is a function of the image size. However, it can be noticed that the informative non-zero values occur in the initial columns of the matrix only. Consequently, for each matrix, we only keep the first 400 columns. In other words, a run of a maximum of 400 pixels (determined empirically) is considered in our study. The four run length matrices are normalized and are converted to vectors (each of dimension 400) which are concatenated together to form a single feature vector, by the same way this vector of black pixels of (400 values) is concatenated with the vector of white pixels of (400 values) likewise calculated. So, the final vector is about (400 + 400 = 800 values). Further details on run-length distributions can be found in previous work [54].

Previously, the run-length features were evaluated in a number of writer identification contests held in conjunction with ICDAR 2011 [55, 56, and 57] and ICFHR 2012 [58, 59]. These features realized interesting results in these competitions. The present study is intended to explore their effectiveness on the more challenging task of signature verification where a very limited amount of text is available as opposed to traditional writer recognition methods.

In an attempt to compare the effectiveness of run-length distributions in detecting skilled forgeries, we have implemented some of the latest state-of-the- art methods that have shown good results on this problem. These include Autoregressive Coefficients, Local Directional Pattern features, Local Binary patterns, Local Derivate Pattern, contour-direction distributions, contour-hinge distributions proposed in [60] and Curvelet transforms... Table I summarizes, for each of the used features, the corresponding dimension.

Feature	Description	Dimension
f1	Black Run-lengths Distributions	400
f2	White Run-lengths Distributions	400
f3	Black and White Run-lengths Distributions	800
f4	Auto regressive Coefficients	24
f5	Local Directional Pattern (LDP)	672
f6	Local Binary Pattern (LBP)	3060
f7	Local derivative pattern (LDerive)	12240
f8	Contour-direction	12
f9	Contour-hinge	1042
f10	Curvelet transforms	10

Table 3.1. Summary of features employed in our study

4. Design of signature verification system

The design stage of the proposed system is composed of three steps: selecting a set of signers, constructing the signature models and finally locating the optimal decision threshold. More precisely, a set of writers is selected from the used dataset, each one having N genuine signatures. To build the signature models, the set of signatures for each writer is divided into two subsets namely Subset 1 and Subset 2. The first subset (Subset 1) containing Np genuine signatures are used for finding the parameters of the OC-SVM during training step. While the second subset (Subset 2) containing Nt genuine signatures is used for finding the optimal decision threshold. (*Figure 3.4*) shows the concept for selecting the optimal parameters of the HSVS [61].

The optimal decision threshold is deduced from FRR and FAR curves using the Half Total Error Rate (HTER) as defined in the next formula:

$$HTER = \frac{FRR + FAR}{2}$$

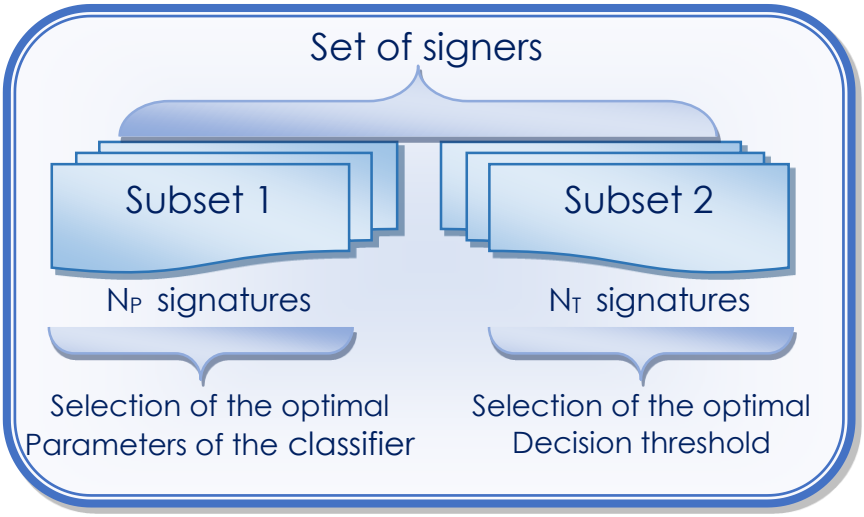


Figure 3.4. Design Stage

5. Performance Evaluation

The performance of the proposed methodology is quantified by computing the standard performance measures. These measures include the Type I error or the False Rejection Rate (FRR), which represents the ratio of the number of genuine test signature images rejected to the total number of genuine test signature images and, the Type II error or the False Acceptance Rate (FAR) which represents the ratio of the number of accepted forgeries to the total number of forgeries. The performances are also measured by calculating the accuracy (ACC) that represents the percentage of correct decisions with respect to all disputed signatures, and Finally we measured the cost of the log-likelihood rations (Cllr) in his minimal possible value $\hat{C}_{llr}^{min}$, that value is a minimal final assessment value, it denotes a better performance of the method.

6. Experimental Settings and Results

To evaluate the effectiveness of the proposed features on detection of forgeries, we have used the GPDS960 signature database. The database contains 24 genuine signatures and 30 forged signatures for a total of 881 different individuals. For experiments reported

in this chapter, signature images belonging to the first 281 individuals are considered for the design step and the remaining 600 writers for the evaluation Step.

To create a signature model that can reject forgeries efficiently and at the same time is tolerant to intra-writer variations, it is important to use a sufficient number of genuine samples in the training step. The influence of number of genuine signature samples in the training set on the overall performance is studied by considering three experimental scenarios. The first of these scenarios includes for each individual, only 4 genuine signatures in the training set whereas 20 genuine signatures as well as 30 skilled forgeries in the test set. In the second scenario, we have used, for each writer, 8 genuine signatures in the training set and 16 genuine signatures as well as 30 skilled forgeries in the test set whereas in the last scenario, we have used, for each writer, 12 genuine signatures in the training set and 12 genuine signatures as well as 30 skilled forgeries in the test set. The distribution of training and test sets in the three scenarios is summarized in Table II. It should be noted that these experimental settings match closely to the real world scenarios where only genuine signatures could be used for training and skilled forgeries are encountered in the test phase only.

Scenario	Design step				Evaluation step			
	#Writers	GR	GQ	FQ	#Writers	GR	GQ	FQ
I	281	4	20	30	600	4	20	30
II	281	8	16	30	600	8	16	30
III	281	12	12	30	600	12	12	30

Table 3.2. Different Experimental Scenarios

* *GR :* Genuine Reference, * *GQ :* Genuine Questioned, * *FQ :* Forged Questioned.

Three series of experiments are carried out corresponding to the scenarios listed in Table II. For each of the features we report the ACC, FAR, FRR and $\hat{C}_{llr}^{min}$.

Scenario I				
Feature	Accuracy (%)	FAR (%)	FRR (%)	$\hat{C}_{llr}^{min}$
f1	91.94	11.17	5.98	0.27
f2	85.47	17.05	12.84	0.50
f3	92.48	10.41	5.59	0.25
f4	72.44	26.22	28.45	0.78
f5	68.78	28.22	33.21	0.83
f6	77.82	21.36	22.73	0.68
f7	75.99	21.84	25.45	0.71
f8	67.55	28.59	35.03	0.85
f9	72.62	24.02	29.62	0.78
f10	71.34	29.51	28.10	0.82

Table 3.3 . Results of the the first experimental scenario.

Scenario II				
Feature	Accuracy (%)	FAR (%)	FRR (%)	$\hat{C}_{llr}^{min}$
f1	93.17	9.42	5.45	0.24
f2	87.61	16.51	10.19	0.45
f3	93.80	8.77	4.82	0.22
f4	74.51	23.75	26.42	0.75
f5	69.48	25.67	33.10	0.80
f6	79.99	20.11	19.96	0.63
f7	78.37	18.59	23.24	0.64
f8	68.37	27.31	33.93	0.85
f9	73.62	22.72	28.40	0.77
f10	71.40	28.59	28.61	0.83

Table 3.4. Results of the the second experimental scenario.

Scenario III				
Feature	Accuracy (%)	FAR (%)	FRR (%)	$\hat{C}_{llr}^{min}$
f1	94.10	7.72	5.17	0.22
f2	89.06	15.40	9.15	0.42
f3	94.63	7.05	4.70	0.20
f4	73.83	22.79	27.52	0.75
f5	68.71	24.93	33.84	0.81
f6	80.86	18.89	19.23	0.61
f7	77.48	17.19	24.66	0.63
f8	66.28	26.47	36.63	0.86
f9	72.10	22.01	30.14	0.77
f10	70.34	28.32	30.20	0.83

Table 3.5. Results of the the third experimental scenario.

The performance of the proposed features as well as that of the state-of-the-art features when evaluated using the three scenarios is summarized in Tables III, IV and V.

A number of interesting observations can be drawn from the achieved verification errors. For scenario I, it can be seen that the false acceptance (FA) and false rejection (FR) rates of different features vary significantly with run-length features (f3) outperforming all other features reporting a FAR of 10.41% and a FRR of 5.59%. Similar trends can be observed for scenario II where the run-length features again outperform all other features reporting a FAR of 8.77% and a FRR of 4.82%. Likewise, in the last experimental scenario, the run-length features report the minimum FAR of 7.05% as well as the minimum FRR of 4.70%. The relatively low performance of contour-direction features, for example, can be attributed to their low dimensionality (12) as compared to other features.

The proposed run-length features come out to be the most effective in terms of error rates. For the $\hat{C}_{llr}^{min}$ the big difference between Run Length distributions and other features is clear, we see for example that $\hat{C}_{llr}^{min}$ value for RL is 0.25 in the first scenario, 0.22 in the second, and 0.20 for the last one, while the best value for the other features is

0.61 was realized by LBP in the third scenario, then 0.630 for the second scenario of the same feature, then 0.632 for the third scenario of LDP, then 0.684 for LBP again, whereas all the rest $\hat{C}_{llr}^{min}$ values exceed 0.7. By balancing the Accuracy values, the highest value is 94.626 %, observed in the third scenario of RL distribution of course, followed by a closer value : 94.098 % of Black RL feature, then 93.803 % and 93.169 % for the second scenario, and for the first scenario we got : 92.478 % and 91.941 % for the same distributions respectively. The white RL comes in the third place having ACCs: 89.862 %, 87.610 % and 85.473 % for Scenarios III, II and I. Whereas, all the other features didn't exceed an ACC of 80.865 % accompanied by 79.988 % and 77.821 % by counting the three scenarios of LBP, Starting by Scenario III to scenario I. the next order returns to the LDERIVE feature, getting: 77.477 %, 78.375 for scenario III then II, and for 75.993 % for the first scenario, the other features had less values.

Comparing the performance measurements of different features across the three evaluation scenarios, it can be seen that the error rates and the Accuracy reduce as the number of genuine signatures in the training set is increased, when the $\hat{C}_{llr}^{min}$ increases. This observation is consistent across all the features. The performance enhancement is more significant in case of run-length features as opposed to any of the other features. It can therefore be concluded that run-length features are effective for detecting skilled forgeries and can realize low error rates with a 'sufficient' number of genuine samples in the training.

The results based on other features, however, do not seem to be as effective and may be explored further by considering different context (neighborhood) sizes for possible improvements.

7. Stability of the proposed features

One of the most important parameters that influence the performance of signature verification systems is the available genuine samples that will be used for training. The number of signers involved in the evaluation step is also an important parameter that can significantly affect the performance of run length features on signature verification. These two parameters will be considered to study the stability of the proposed features and system

In order to simulate the real conditions and to cover the limited number of writers and genuine signatures, which is really the main problem for constructing a robust signature verification system, we will evaluate the influence of number of genuine signature samples in the training set by applying a special scenario. It includes for each individual, only one genuine signature in the training set whereas the remaining 23 genuine signatures as well as the available 30 skilled forgeries have to be used in the test set. The results of this scenario are reported in Table VI.

It should be noted that these experimental settings also match closely to the real world scenarios where only one genuine signature could be used for training and skilled forgeries are encountered in the test phase only, we have used each signature alone and applied the feature from the first until the 24th one, to see the difference of competence results every time, the evaluation measurements are calculated and an average is given as following table.

Signature	Accuracy (%)	FAR (%)	FRR (%)	$\hat{C}_{llr}^{min}$
1	90.031	10.464	9.589	0.333
2	90.588	10.072	8.905	0.321
3	91.057	10.297	7.904	0.296
4	91.387	11.239	6.597	0.273
5	91.903	07.188	8.794	0.274
6	91.951	09.290	7.097	0.266
7	92.577	08.348	6.713	0.249
8	92.243	07.920	7.631	0.264
9	92.611	07.493	7.308	0.255
10	92.448	08.768	6.619	0.252
11	92.234	07.725	7.798	0.268
12	92.473	07.775	7.336	0.256
13	92.224	08.507	7.214	0.258
14	92.265	08.580	7.086	0.256
15	92.297	08.797	6.864	0.253
16	92.316	08.203	7.286	0.257
17	92.215	08.710	7.075	0.264
18	92.158	08.529	7.314	0.261
19	91.881	08.985	7.453	0.271
20	91.642	09.406	7.553	0.277
21	91.771	09.239	7.453	0.273
22	91.504	09.645	7.614	0.283
23	91.592	09.080	7.892	0.283
24	91.246	09.450	8.221	0.295
Overall	91.859	08.905	7.555	0.272

Table 3.6. Results when using only one genuine sample in training step.

We see clearly that the overall values, by using only one genuine signature, are encouraging and give better results than using more than one genuine signature for the other features, which indicates the robustness of the proposed run-length features.

Lastly, to evaluate our system stability, we have studied the influence of changing the signer's number using 4, 8 and 12 genuine samples for each signer, in order to check the cases when we use a few or many samples of the signatures dataset. The recorded results prove the stability and the robustness of our features, which is clear and evident in the next curves whose represent the variation of the performance measurement ACC, FAR, FRR and $\hat{C}_{llr}^{min}$ by varying the number of signers, we used always 281 writers in the HSVS design for GPDS, and the remaining (600 writers) are used for evaluating its performance by changing them from 02 until 600 writers, so we got these results:

ACC:

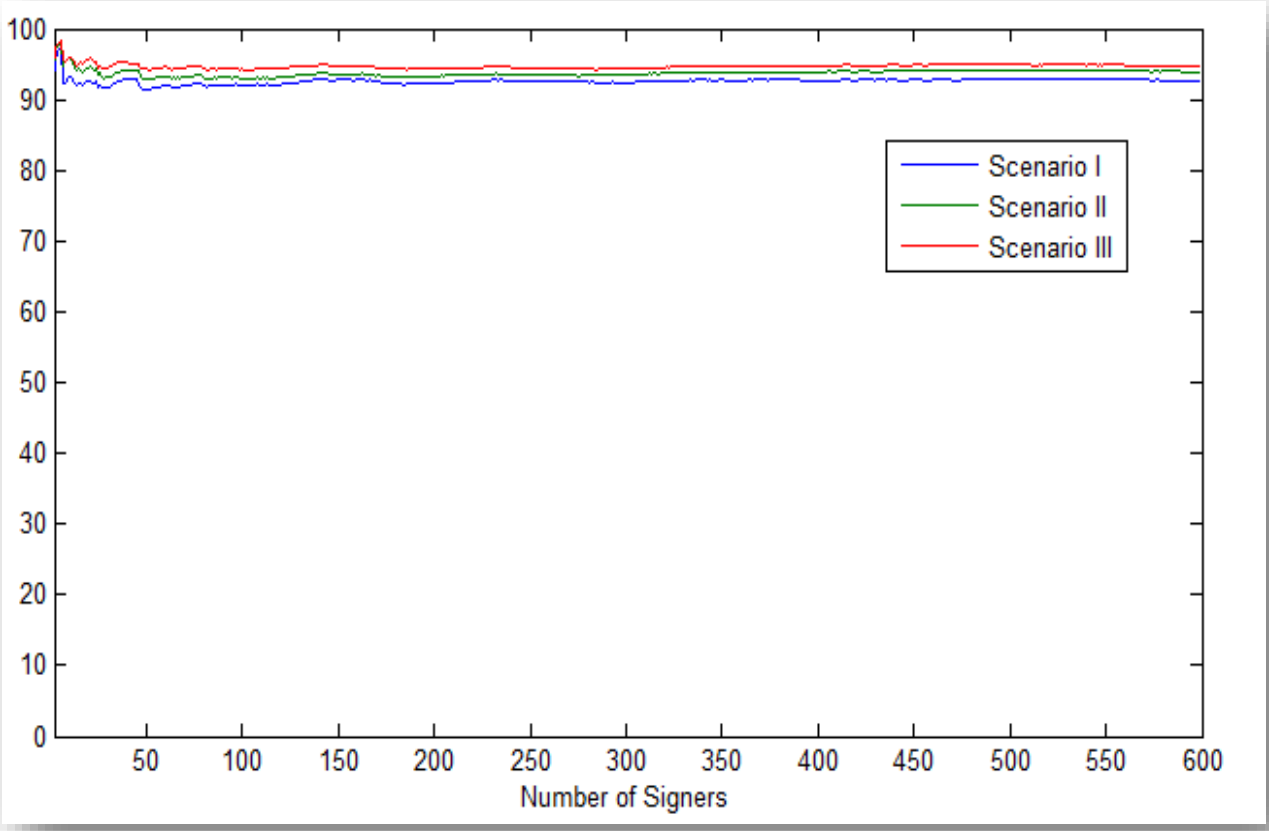


Figure 3.5. ACC values by changing Number of Signers

It is Evident that the curves of the three scenarios, tolerate some weak changes for the first 50 variations, but gradually, it seems to be staying steady, getting the lead of 600 signers, which appears not influenced very well by changing the number of signers.

FAR – FRR:

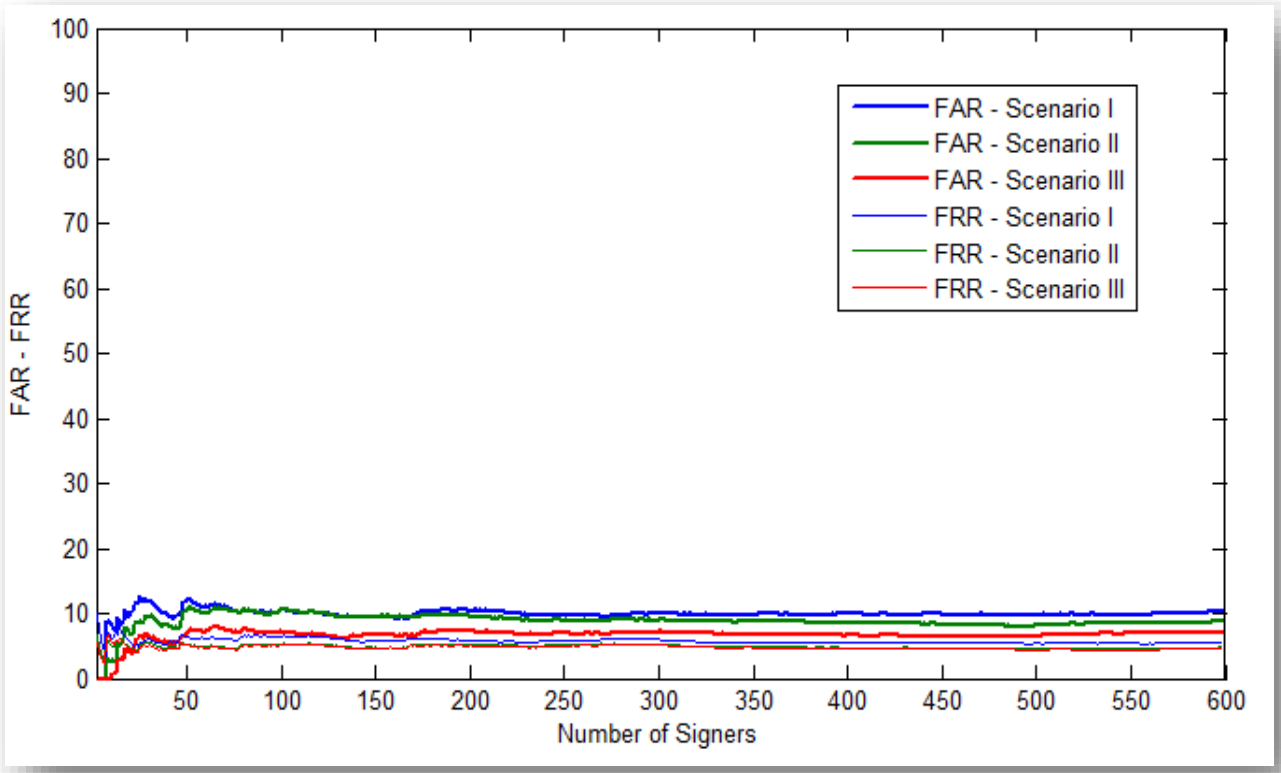


figure 3.6. FAR and FRR values by changing Number of Signers

As we said for the ACC measurement, the curves suffer some weak changes for the first 50 until 100 variations, but step by step, it adjusts to steadiness, until having 600 signers. The influence of number of signers is so weak on the robustness of system.

$\hat{C}_{llr}^{min}$:

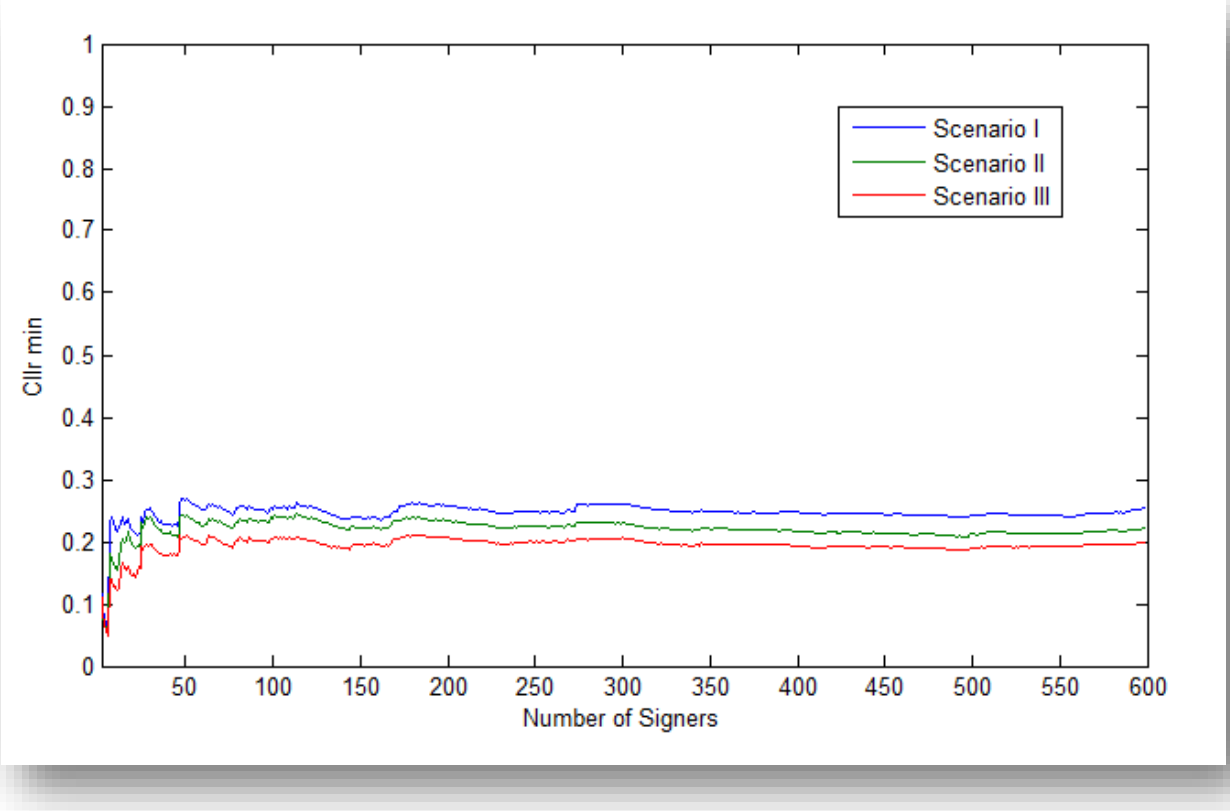


figure 3.7. $\hat{C}_{llr}^{min}$ values by changing Number of Signers

For $\hat{C}_{llr}^{min}$, for the three scenarios, its values change for some variations before having about 170 signers, then the curve stays stable for all the rest variations.

The Figures 5, 6 and 7 indicate clearly the stability and the good steadiness of the proposed features, when varying the signer's number from 02 to 600 signers, because it is transparent to see that curves are almost converge near to a straight line. As a result: the influence of the number of signers is almost ignored on the exactitude of our system.

We clearly can note that when the signature number increases, the ACC does not affect significantly the performance of the HSVS, and so on for the other evaluation factors: FAR, FRR and $\widehat{C}_{llr}^{min}$. Therefore, few signatures are sufficient for designing the HSVS.

8. Conclusion

In this chapter, we treated the feature extraction and the application of our proposed system for handwritten signature verification that is Run Length Distributions. Then using the OC-SVM, we've employed only genuine signatures in training step.

The validity of the HSVS start up with good choice of the optimal threshold, from genuine and adulterate signatures, and by carefully adapting the OC-SVM appropriate kernel, in the design step, to get an accurate classification.

Relatively to the state of the art, the results generated by our system and computed on GPDS960 dataset, indicate its important performances, and interesting validity.

The influence of applying only one genuine signature in training step, and varying the number of signers, are other witnesses to the robustness and the reliability of the HSVS proposed, that hadn't been well impacted.

General Conclusion

General Conclusion

We presented three texture-based features, run-length distributions: Black, white and Black and White, applied to the problem of signature verification. The system was evaluated on signatures of 881 individuals from the GPDS960 database including skilled forgeries. The performance of these features was also compared with some well-known signature verification features proposed in the literature. It was observed that for a sufficient number of genuine signatures in the training set, the run-length features outperform other features considered in this study.

The robustness of the system is more clear when using only one genuine signature in training set, which is the extreme case in reality, and finally, our system is sufficiently stable when changing the number of signers .In our further study on this subject, we intend to complement these features by other textural measurements and evaluate the system on much larger signature repositories. We also intend to extend the study to the verification of online signatures by incorporating dynamic features into the system.

Bibliography

Bibliography

[1] Kuang-Shyr Wu, Jen-Chun Lee, Tsung-Ming Lo, Ko-Chin Chang, and Chien-Ping Chang. A secure palm vein recognition system. Journal of Systems and Software, 86(11):2870–2876, 2013.

[2] Weihua Ou, Xinge You, Dacheng Tao, Pengyue Zhang, Yuanyan Tang, and Ziqi Zhu. Robust face recognition via occlusion dictionary learning. Pattern Recognition, 47(4):1559–1572, 2014.

[3] GS Badrinath and Phalguni Gupta. Stockwell transform based palm- print recognition. Applied Soft Computing, 11(7):4267–4281, 2011.

[4] Dongjin Fan, Peng Yu, Peng Du, Wenda Li, and Xiaofei Cao. A novel probabilistic model based fingerprint recognition algorithm. Procedia Engineering, 29:201–206, 2012.

[5] Tomas Obsil and Veronika Obsilova. Structural basis for dna recognition by foxo proteins. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1813(11):1946–1953, 2011.

[6] Li Zhu and Qing Yang. Speaker recognition system based on weighted feature parameter. Physics Procedia, 25:1515–1522, 2012.

[7] M Karnan, M Akila, and N Krishnaraj. Biometric personal authenti- cation using keystroke dynamics: A review. Applied Soft Computing, 11(2):1565–1573, 2011.

[8] Chin Poo Lee, Alan WC Tan, and Shing Chiang Tan. Gait recognition via optimally interpolated deformable contours. Pattern Recognition Letters, 34(6):663–669, 2013.

[9] Shi-Lin Wang and Alan Wee-Chung Liew. Physiological and behavioral lip biometrics: A comprehensive study of their discriminative power. Pattern Recognition, 45(9):3328–3335, 2012.

[10] Rejean Plamondon and Guy Lorette. Automatic signature verification and writer identification: the state of the art. Pattern recognition, 22(2):107–131, 1989.

[11] Franck Leclerc and Rejean Plamondon. Automatic signature verifica- tion: The state of the art 1989-1993. International Journal of Pattern Recognition and Artificial Intelligence, 8(03):643–660, 1994.

[12] Rejean Plamondon and Sargur N Srihari. Online and off-line handwrit- ing recognition: a comprehensive survey. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(1):63–84, 2000.

[13] Donato Impedovo and Giuseppe Pirlo. Automatic signature verifica- tion: the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 38(5):609–635, 2008.

[14] Donato Impedovo, Giuseppe Pirlo, and Rejean Plamondon. Hand- written signature verification: new advancements and open issues. In Proceedings of International Conference on Frontiers in Handwriting Recognition, pages 367–372, 2012.

[15] Sargur N Srihari, Ahuja Xu, and Meenakshi K Kalera. Learning strategies and classification methods for off-line signature verification. In Proc. of 9th International Workshop on Frontiers in Handwriting Recognition, pages 161–166, 2004.

[16] Julian Fierrez-Aguilar, N Alonso-Hermira, G Moreno-Marquez, and Javier Ortega-Garcia. An off-line signature verification system based on fusion of local and global information. In Biometric Authentication, pages 295–306. 2004.

[17] Bao Ly Van, Sonia Garcia-Salicetti, and Bernadette Dorizzi. On using the viterbi path along with hmm likelihood information for online signature verification. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 37(5):1237–1247, 2007.

[18] Kaiyue Wang, Yunhong Wang, and Zhaoxiang Zhang. On-line signature verification using graph representation. In Proc. of the 6th International Conference on Image and Graphics, pages 943–948, 2011.

[19] Marcin Piekarczyk. Hierarchical random graph model for off-line handwritten signatures recognition. In Proc. of International Conference on Complex, Intelligent and Software Intensive Systems, pages 860–865, 2010.

[20] Donato Impedovo, Giuseppe Pirlo, and Rejean Plamondon. Hand- written signature verification: new advancements and open issues. In Proceedings of International Conference on Frontiers in Handwriting Recognition, pages 367–372, 2012.

[21] Edson JR Justino, Flavio Bortolozzi, and Robert Sabourin. Off- line signature verification using hmm for random, simple and skilled forgeries. In Proc. of the sixth International Conference on Document Analysis and Recognition, pages 1031–1034, 2001.

[22] M.I. Malik, M. Liwicki, L. Alewijnse, W. Ohyama, M. Blumenstein, B. Found, “ICDAR 2013 Competitions on Signature Verification and Writer Identification for On- and Offline Skilled Forgeries (SigWiComp 2013)”, In Proceedings of the 12th International Conference on Document Analysis and Recognition, Washington, DC, USA, pp: 1477 -1483.

[23] M. Liwicki, E. van den Heuvel, B. Found and M.I. Malik., "Forensic Signature Verification Competition 4NSigComp2010 - Detection of Simulated and Disguised Signatures", In Proceeding of the 12th International Conference on Frontiers in Handwriting Recognition, Kolkata, India, pp: 715 – 720, 2010.

[24] M. Liwicki, M.I. Malik, E. van den Heuvel and X. Chen., "Signature Verification Competition for Online and Offline Skilled Forgeries (SigComp2011)", In Proceeding of the 11th International Conference on Document Analysis and Recognition, Beijing, China, pp: 1480 – 1484, 2011.

[25] M. Liwicki, M.I. Malik, L. Alewijnse, C. Elisa van den Heuvel, B. Found., “ICFHR 2012 Competition on Automatic Forensic Signature Verification (4NsigComp 2012)”, In Proceedings of the International Conference on Frontiers in Handwriting Recognition, Bari, Italy, pp: 823 - 828, 2012.

[26] Miguel A Ferrer, JFrancisco Vargas, Aythami Morales, and Ord. Ro- bustness of offline signature verification based on gray level features. IEEE Transactions on Information Forensics and Security, 7(3):966– 977, 2012.

[27] Tee Wilkin, Ooi Shih Yin “State of The Art: Signature Verification System“7th International Conference on Information Assurance and Security (IAS),2011

[28] IlseGiesingchapter5”biometrics”university of Pretoria 2003.

[29] National science and technology council (NSTC): committee on technology, Subcommittee on biometrics, 2005, http://www.biometrics.org/forums/showthread.php/74-NSTC-Subcommittee-onBiometrics-Reference-Documents.

[30] Boudjellal. S. Détection et identification de personne par méthode biométrique.Mémoire de Magister : Electronique-Télédétection. Tizi-ouzou University Mouloud Mammeri, 95p.

[31] Ming-Hsuan Yang, David J. Kriegman et Narendra Ahuja. Detecting faces Images: A survey. Dans IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002, vol.24 (1), 34-58

[32] United States Patent and Trademark Office, "Patent 4,736,203: 3D hand profile identification apparatus," 5 April 1988 >.

[33] IR Recognition Systems <http://recogsys.com/index.shtml>.

[34] "Finger Scanning at Disney Parks Causes Concern," 15 July 2005 <htto://www.local6.cominews/4724689/detail.html>.

[35] Suvarna Joshi and AbhayKumar “Feature Extraction Using DWT with Application to Offline Signature Identification” Volume 222, 2013.

[36] Meenakshi S Arya, Vandana S Inamdar “A Preliminary Study on Various Off-line Hand Written Signature Verification Approaches” International Journal of Computer Applications (0975 – 8887) ,Volume 1 – No. 9 ,2010.

[37] J. Brault, R. Plamondon “A Complexity Measure of Handwritten Curves” Modeling of Dynamic Signature Forgery. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 23, No. 2, 1993.

[38] OndrejRohlik “Handwritten text analysis” thesis, university of west bohemia in pilsen, page 16 (mazall) March 2003.

[39] Andreas Schlapbach. Writer identification and verificationClearway Logistics Phase 2-3 (November 14, 2007), 2007.

[40] MadasuHanmandlu, Mohd. HafizuddinMohd. Yusof, and Vamsi Krishna Madasu. Off-line signature verification and forgery detection using fuzzy modeling. Pattern Recognition. 38:341–356, March 2005.

[41] H. Weiping, Y. Xiufen, and W. Kejun, “A survey of off-line signature verification,” in Intl. Conf. on Intelligent Mechatronics Automation, pp. 536–541, Aug. 26-31 2004.

[42] J.Ruiz-Del-Solar, C. Devia, P. Loncomilla, and F. Concha, “Offline signature verification using local interest points and descriptors,” in CIARP '08: Proceedings of the 13th Ibero american congress on Pattern Recognition. Berlin,

Heidelberg: Springer-Verlag, 2008, pp. 22–29.

[43] Savanna Thirumuruganathan, a Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm, May 17, 2010.

[44] wd,wi Sargur N Srihari, Aihua Xu, and Meenakshi K Kalera. Learning strategies and classification methods for off-line signature verification. In Proc. of 9th International Workshop on Frontiers in Handwriting Recognition, pages 161–166, 2004.

[45] T. Ojala, M. Pietikainen, and D. Harwood. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. Pages A: 582-585, 1994.

[46] M. Ferrer, F. Vargas, C. Traviesto and J. Alonso. Signature veri_cation using local directional pattern".In International Carnahan Conference on Security. Technology, pp. 336{340. IEEE, Oct 2010.

[47] V. Ojansivu and J. Heikkil¨a. Blur insensitive texture classification using local phase quantization. In ISP,2008.

[48] M. Crosier and L. Griffin. Using basic image features
for texture classification. IJCV, 3(88):447–460, 2010.

[49] N. Dalai and B. 'Diggs. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Com-puter Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - *hone 01, CVPR '05, pages 886-893, Washington, DC, USA, 2005. IEEE aim-puler Society.

[50] B. Schölkopf, J. Platt, J. Shawe-Taylor, A. Smola, R. Williamson, Estimating the support of a high dimensional distribution, Neural Comput. 13 (7) (2001)1443–1472.

[51] D. Impedovo, G. Pirlo, Automatic signature verification: the state of the art,IEEE Trans. Syst. Man Cybern Part C: Appl Rev. 38 (5) (2008).

[52] L. Hamel. Knowledge discovery with support vector machines. Wiley Edition, 2009.

[53] Chawki Djeddi, Imran Siddiqi, Labiba Souici-Meslati, and Abdellatif Ennaji. Text-independent writer recognition using multi-script hand- written texts. Pattern Recognition Letters, 34(10):1196–1202, 2013.

[54] Chawki Djeddi, Labiba Souici-Meslati, and Abdellatif Ennaji. Writer recognition on arabic handwritten documents. In Image and Signal Processing, pages 493–501. 2012.

[55] G. Louloudis, N. Stamatopoulos, and B. Gatos. Icdar 2011 writer identification contest. In Proc. of International Conference on Document Analysis and Recognition, pages 1475–1479, 2011.

[56] A. Hassaine, S. Al-Maadeed, J.M. Aljaam, A. Jaoua, and A. Bouri- dane. The icdar2011 arabic writer identification contest. In Proc. of International Conference on Document Analysis and Recognition, pages 1470–1474, Sept 2011.

[57] A. Fornes, A. Dutta, A. Gordo, and J. Llados. The icdar 2011 music scores competition: Staff removal and writer identification. In Proc. of International Conference on Document Analysis and Recognition, pages 1511–1515, Sept 2011.

[58] G. Louloudis, B. Gatos, and N. Stamatopoulos. Icfhr 2012 competition on writer identification challenge 1: Latin/greek documents. In Proc. of International Conference on Frontiers in Handwriting Recognition, pages 829–834, 2012.

[59] A. Hassaine and S.A. Maadeed. Icfhr 2012 competition on writer identification challenge 2: Arabic scripts. In Proc. of International Conference on Frontiers in Handwriting Recognition, pages 835–840, 2012.

[60] Almudena Gilperez, Fernando Alonso-Fernandez, Susana Pecharroman, Julian Fierrez, and Javier Ortega-Garcia. Off-line signature verification using contour features. In Proc. of the 11th International Conference on Frontiers in Handwriting Recognition, 2008.

[61] Y Guerbai, Y Chibani,B Hadjadji Y. The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters / Pattern Recognition 48 (2015) 103–113.