Offline writer authentication

by

Sandura Shumba

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science (Applied Mathematics) in the Faculty of Science at Stellenbosch University

Supervisor: Dr. J. Coetzer

December 2021

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: .... December 2021 ....

Abstract

Offline writer authentication

S. Shumba

Division of Applied Mathematics,
Stellenbosch University,
Private Bag X1, Matieland 7602, South Africa.

Thesis: MSc

December 2021

In this thesis a number of systems are proposed for the purpose of offline writer authentication. A text-dependent approach is adopted, since a very specific targeted handwritten word is considered for authentication purposes. Feature extraction is facilitated by calculating a number of projections of the targeted word from different angles. Two distinct categories of systems are proposed. The first category employs template matching and is based on the computation of the Euclidean distance and a dynamic time warping (DTW) distance between corresponding feature vectors, while the second category relies on machine learning techniques, that is support vector machines (SVMs) and quadratic discriminant analysis (QDA). Within the context of the proposed machine learning-based systems, a writer-independent protocol is followed. This is achieved by employing a DTW-based dichotomy transformation which converts a feature set in feature space into a dissimilarity vector-based representation in dissimilarity space. This dichotomy transformation is followed by writer-specific dissimilarity vector normalisation which significantly improves interclass separability. The DTW-based dichotomy transformation and writer-specific dissimilarity vector normalisation are novel within the context of offline writer authentication. The systems developed in this study are evaluated on a subset of the CEDAR-LETTER data set. It is demonstrated that the proficiency of the systems developed in this study are at least on par when compared to existing systems. The most proficient SVM-based system developed in this study achieves an AUC of 93% and an equal error rate (EER) of 14.93%.

Uittreksel

Vanlyn skrywerverifikasie

S. Shumba
Afdeling Toegepaste Wiskunde,
Universiteit Stellenbosch,
Privaatsak X1, Matieland 7602, Suid-Afrika.

Tesis: MSc

Desember 2021

In hierdie tesis word 'n aantal stelsels vir die doel van vanlyn skrywerverifikasie voorgestel. 'n Teksafhanklike benadering word gevolg, aangesien 'n baie spesifieke handgeskrewe teikenwoord vir verifikasiedoeleindes beskou word. Kenmerkonttrekking word moontlik gemaak deur 'n aantal projeksies van die teikenwoord vanuit verskillende hoeke te bereken. Twee aparte kategorieë van stelsels word voorgestel. Die eerste kategorie gebruik templaatpassing en is gebaseer of die berekening van die Euklidiese afstand en 'n dinamiese tydsverbuiging (DTW) afstand tussen die ooreenstemmende kenmerkvektore, terwyl die tweede kategorie op masjienleertegnieke staatmaak, m.a.w. ondersteuningsvektormasjiene (SVMs) en kwadratiese diskriminant-analise (QDA). Binne die konteks van die voorgestelde masjienleergebaseerde stelsels word 'n skrywer-onafhanklike protokol gevolg. Dit word moontlik gemaak deur 'n DTW-gebaseerde tweeledigheidstransformasie te ontplooi wat 'n kenmerkstel in die kenmerkruimte na 'n verskilvektor-gebaseerde voorstelling in die verskilruimte omskakel. Hierdie tweeledigheidstransformasie word deur skrywerspesifieke verskilvektor-normalisasie gevolg, wat interklas-skeibaarheid aansienlik verhoog. Die DTW-gebaseerde tweeledigheidstransformasie en skrywerspesifieke verskilvektor-normalisasie is nuut binne die konteks van vanlyn skrywerverifikasie. Die stelsels wat in hierdie tesis ontwikkel is, word op 'n substel van die CEDAR-LETTER datastel geëvalueer. Dit word aangetoon dat die vaardigheid van die stelsels wat in hierdie studie ontwikkel is ten minste soortgelyk is aan dié van bestaande stelsels. Die mees vaardige SVM-gebaseerde stelsel wat in hierdie studie ontwikkel is behaal 'n AUC van 93% en 'n gelyke foutkoers (EER) van 14.93%.

Acknowledgements

I would like to express my sincere gratitude to the following people and organisations:

- My supervisor Dr Johannes Coetzer, for his patience, unwavering support, invaluable insight and contribution to my development as a researcher. This study would not have been possible without his input. I am grateful and honoured to have been his student.
- The Postgraduate Funding Department of Stellenbosch University, for their support and financial assistance.
- My Pastor for his spiritual guidance.
- My Family, for their love and support.
- God almighty for His grace and mercy in my life

Contents

List of Figures

List of Tables

List of Acronyms

AMF	Adaptive Median Filter
ANN	Artificial Neural Network
AUC	Area Under Curve
DA	Discriminant Analysis
DRT	Discrete Radon Transform
DTW	Dynamic Time Warping
EER	Equal Error Rate
FAR	False Acceptance Rate
FN	False Negative
FP	False Positive
FRR	False Rejection Rate
GITF	Global Improved Texture Features
GMM	Gaussian Mixture Model
GMM-UBM	Gaussian Mixture Model Universal Background Model
iKNN	Iterative K-Nearest Neighbour
LBP	Local Binary Pattern
(L/Q)DA	(Linear/Quadratic) Discriminant Analysis
LDCF	Local Directional Chain-Code Features
LPQ	Local Phase Quantization
mAP	Mean Average Precision
ML	Machine Learning
MVA	Majority Voting Algorithm
PCA	Principal Component Analysis
PDF	Probability Density Function
QDS	QDA-Based System
RBF	Radial Basis Function
RBK-SVM	Radial Basis Function Kernel Support Vector Machine
ROC	Receiver Operating Characteristic
SDC	Simple Distance Classifier
SIFT	Scale Invarient Feature Transform
SR-KDA	Spectral Regression Kernal Discriminant Anaysis
SVM	Support Vector Machine
TN	True Negative
TP	True Positive

Nomenclature

Image representation

d	Feature vector dimension
T	Feature set length/ dissimilarity vector dimension
I	Sample image
I^B	Binary image
$R_\theta(x')$	Continuous Radon transform
$\mathbf{X}$	Feature set
$\mathbf{x}_i$	Feature vector

Classifiers

τ	Threshold
$\mathbf{w}$	SVM weight vector
b	SVM bias
$\mathbf{Z}$	Dissimilarity vector set
$\mathbf{z}$	Dissimilarity vector
$\bar{\mathbf{z}}$	Normalised dissimilarity vector
C	SVM regularisation hyper-parameter
$\mathcal{E}$	SVM slack variable
$\phi(\mathbf{z})$	Mapping function of vector $\mathbf{z}$
$K(\mathbf{z}, \mathbf{z}')$	RBF kernel transformer for vector $\mathbf{z}$
σ	RBF kernel width

Modelling and verification

$\mathbf{C}$	Covariance matrix
$\mathbf{C}^{-1}$	Inverse of covariance matrix
$\lvert\mathbf{C}\rvert$	Determinant of covariance matrix
λ	Machine learning model
s	Confidence score

General

Ψ	Number of writers
r	Writer
$'$	Transpose
K	Number of reference samples

Chapter 1

Introduction

1.1 Background and motivation

In today's ever-evolving world, security systems have become critical to protect individuals and organisations. Identity management systems have emerged as unique and safe security measures to distinguish or authenticate an individual. The task of identity management systems is to verify an individual's claimed identity in order to prevent imposters from accessing resources they are not entitled to have access to. Traditional methods of establishing or verifying an individual's identity involve knowledge-based systems such as passwords or ID cards. However, the challenge with such forms of identification measures is that they can be easily manipulated. The aforementioned risks of compromising an identity management system do not exist within the context of biometric authentication systems as they establish an individual's identity based on who he or she is (Jain *et al.*, 2006) rather than what he or she possesses or remembers.

Biometrics have found application in many fields of research such as engineering, pharmaceuticals, agriculture and medicine as well as the study of biological behaviours with the assistance of statistical methods. With the development of advanced technological methods, biometrics have become more relevant. In today's research, biometrical methods automatically analyse behavioural and/or physiological characteristics of a human in order to recognise an individual or distinguish one person from another (Gracia *et al.*, 2006). A biometric can be defined as a unique, measurable characteristic of a human being that assists in automatically identifying the individual in question.

1.1.1 Biometric systems

Biometric systems can be categorised into those that employ either physiological or behavioural characteristics, as depicted in Figure 1.1. Physiological characteristics refer to the physical traits of an individual. These include fingerprints, palm prints, retinal patterns, and facial features. Behavioural characteristics refer to the behaviour of a human being, that is how the individual uses his/her body. These include handwriting, signatures and voice samples. Each type of biometric system relies on these characteristics to establish the identity of the individual. A number of systems supplement behavioural characteristics with physiological characteristics producing an extra level of security. Although physiological characteristics are considered to be more reliable, behavioural characteristics are non-intrusive and can be collected without the knowledge or consent of the subject (Yampolskiy and Govindaraju, 2008).

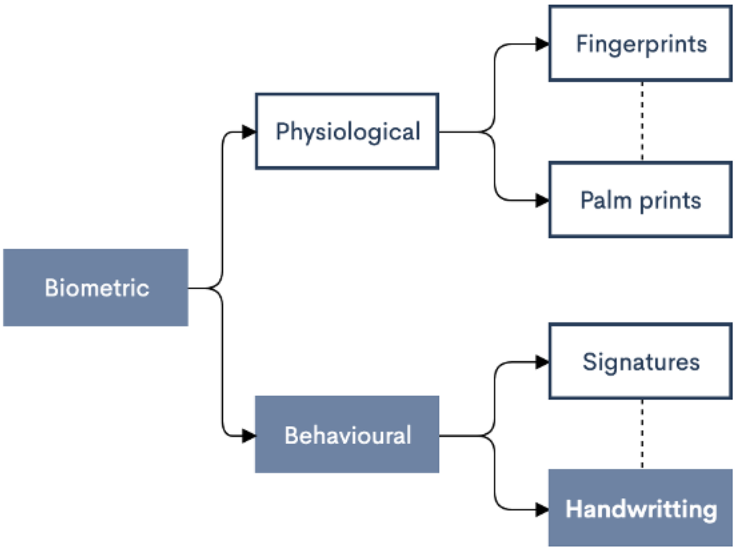


Figure 1.1: A graphical representation of the different types of biometrics. Those utilised in this study are denoted with a dark background.

1.1.2 Handwriting as a biometric

Handwriting is categorised as a behavioural biometric. Humans learn and develop this over time. Handwriting involves the coordination between the individuals' hands and eyes. Handwriting analysis is the science of identifying the individual characteristics of handwriting for the purpose of developing a recognition system. It has been the subject of considerable research interest for several decades, and spread across a variety of fields that include forensic analysis (Tapiador *et al.* (2004)), the analysis of historical documents (Fornés *et al.* (2008)) and security (Ballard *et al.* (2006)). In addition to this, writer recognition is also used to verify financial documents, wills, and pieces of evidence in criminal investigations, such as handwritten ransom notes. A typical handwriting recognition algorithm aims to assist experts by facilitating the semi-automated process of writer identification. These techniques produce a list of predicted authors of a queried sample, ranked in terms of confidence, which then allows the forensic expert to make the final decision as described in Fiel and Sablatnig (2012).

Identifying an individual based on their handwriting is a reliable identification tool. Handwriting is considered a valuable biometric since it meets the criterion of being a measurable behaviourable characteristic. It is universal in terms of practicality and is unique to every individual as demonstrated by Srihari *et al.* (2002). Although it has been observed that ageing and other physical conditions associated with ageing have some impact on the handwriting process (Walton (1997)), the effects of aging are negligible in preventing or adversely affecting the process of identification. Handwriting is therefore deemed to be stable throughout an individual's life.

1.2 Key concepts

This section presents a brief discussion on some of the fundamental concepts involved in handwriting authentication and recognition systems.

1.2.1 Sample acquisition

In order to conduct a machine-based analysis of handwritten samples, the data from the aforementioned samples first needs to be captured. This process involves the digitisation of the samples using a flatbed scanner. The handwritten sample is scanned, resulting in a static image where each pixel is either associated with the foreground (pen strokes) or the background of the document.

1.2.2 Preprocessing

Various instruments and writing surfaces may be used to produce handwritten samples. As a result, it may not always be possible to observe a clear distinction between the foreground and background pixels. In addition to this, the scanning of sample documents may also result in "noise" within the final scanned image. Several image processing techniques may be employed to "clean" the samples and prepare them for analysis. In addition to this, this study utilises a *text-dependent* approach to handwriting authentication, making it necessary for the targeted word to be segmented from the remainder of the document using preprocessing techniques. It is important to note that these static images do not contain any temporal information. As a result, the analysis of said samples is limited to the interpretation of the spatial information indicated by the pen-stroke pixels.

1.2.3 Recognition versus verification

It is important to emphasise the distinction between pattern recognition/identification and pattern verification/authentication. Within the context of pattern recognition, the system attempts to identify the pattern class to which the questioned sample belongs. The process of identifying the correct pattern class involves a one-to-many comparison since the questioned sample is compared to all available templates or models in the database. Within the context of pattern verification, on the other hand, the system compares the questioned sample to that of the *claimed* writer. The aforementioned comparison is therefore carried out on a one-to-one basis in the sense that the system compares the questioned sample to an existing template or model. The systems developed in this study perform pattern *authentication* rather than recognition. This effectively involves a two-class problem, where the aim is to distinguish between a positive class and a negative class (Jain *et al.*, 2005).

1.2.4 Feature extraction

Within the context of feature extraction, quantifiable features are extracted from each raw image sample. This process involves determining the exact features which serve as useful identifiers for classifying samples. Effective features are those that, with a high degree of confidence, can identify the class to which a sample belongs. In this study the features are represented as a feature set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\}$, that is a collection of T, d-dimensional observations. In this study, feature extraction is based on the calculation of the discrete Radon transform (DRT) (Coetzer, 2005).

1.2.5 Classifiers

The purpose of a classifier is to distinguish between different pattern classes. These can be either represented in feature space or dissimilarity space. Such a classifier can be constructed by calculating class-specific descriptive statistics or by training a sophisticated statistical model. Several of these classification techniques are depicted in Figure 1.2. This study employs two distinct types of classifiers, that is simple distance classifiers (SDCs), which is based on the Euclidean distance and a dynamic time warping (DTW) distance, as well as machine learning-based classifiers, specifically support vector machines (SVMs) and classifiers based on quadratic discriminant analysis (QDA).

1.2.6 Training and testing

Within the context of this study, training is described as the process by which a statistical classifier is "taught" to distinguish between samples from different writers. Training requires a sufficiently large collection of historical patterns

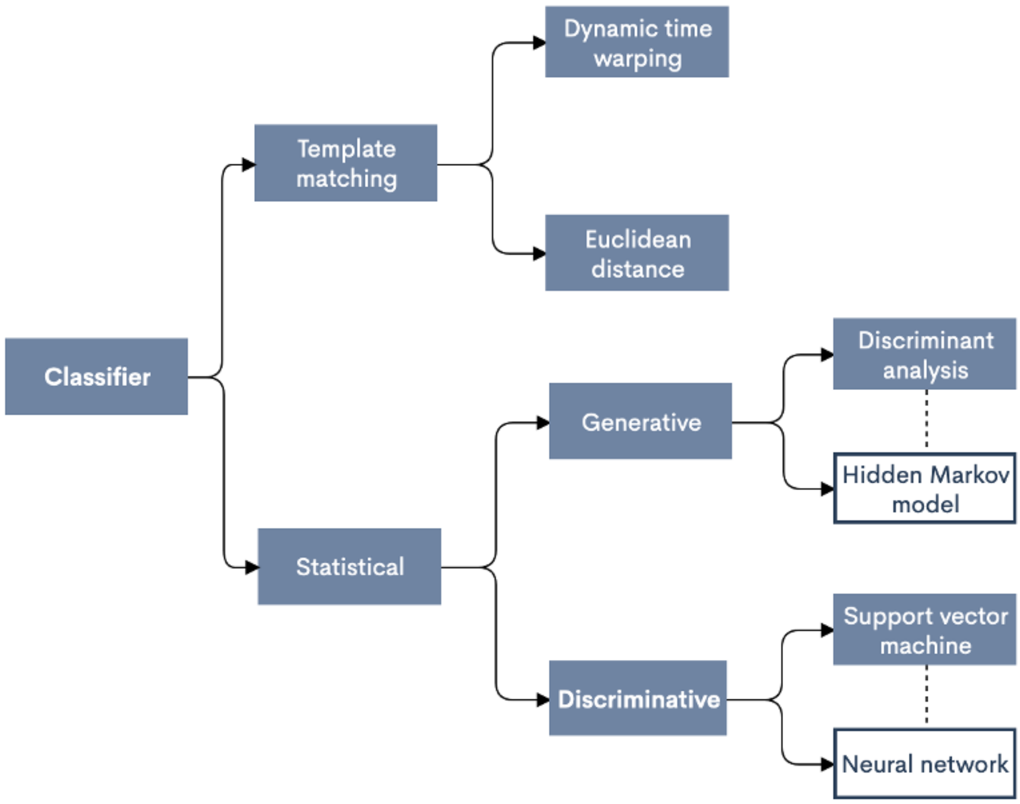


Figure 1.2: A conceptualisation of different types of classification techniques. Those utilised in this study are denoted with a dark background.

as a *training* set. The number of training samples required is dependent on the classification technique, however, larger training sets inevitably result in a superior classifier. Within the context of machine learning-based classification, this study employs what is known as *supervised learning*. Supervised learning is where the respective classes of the patterns are known beforehand, whereas, with unsupervised learning, the correct classes are unknown. A typical classifier receives a feature set and subsequently emits a numeric score. Using a suitable transformation technique, the score can be expressed on a confidence scale where 1 denotes a perfect match, and 0 indicates a complete mismatch.

Within the context of machine learning based-classification, the test phase is concerned with assigning a class label to a questioned sample whose label is known for the purpose of evaluating the system's proficiency. A questioned sample is provided to the trained classifier, which then produces a decision or confidence score.

1.2.7 Writer-independent modelling

Within the context of handwriting authentication, a writer-dependent approach implies that different classifiers are trained for different writers. This approach requires each writer to submit a relatively large number of training samples. In a writer-independent approach, a single classifier is trained across all writers. The classifier attempts to model *differences* between a questioned (positive or negative) and reference (positive) sample. The machine learning-based systems developed in this thesis employ a *writer-independent* approach to handwriting authentication, constructed in *disimilarity space* (Pekalska *et al.*, 2001).

1.2.8 Text-dependent versus text-independent handwriting authentication

Automatic handwriting authentication can be divided into text-dependent and text-independent approaches (Namboodiri and Gupta, 2006). In the text-dependent approach, the writer is identified using a *specific* sample of writing that the writer had to reproduce. This can be a single letter, word, phrase or sentence. *Text-independent* systems, on the other hand, are able to authenticate a writer irrespective of the content of the text provided. The text-dependent approach typically leads to more accurate systems, however, since a specific script has to be generated by each writer, it may also be less practical. This study adopts a *text-dependent* approach to handwriting authentication, with the sample text being the single word "the".

1.3 Scope and objectives

This study aims to achieve the following primary objectives:

- To develop a novel strategy for writer-independent handwriting representation with the specific objective of investigating the use of the discrete Radon transform (DRT) in conjunction with a DTW algorithm for the respective purposes of feature extraction and the subsequent dichotomy transformation.

- To develop a novel strategy for incorporating writer-specific information into a writer-independent modelling framework with the specific aim to investigate the feasibility of a writer-specific dissimilarity normalisation function.

- To design and implement a text-dependent handwriting authentication system using DTW distance-based and Euclidean distance-based template matching techniques.

- To design and implement a robust text-dependent handwriting authentication system using SVM-based and QDA-based machine learning techniques.

1.4 System overview

Two distinct types of systems are developed in this study, those that make use of template matching techniques, that is those based on the Euclidean distance and those based on a dynamic time warping (DTW) distance, as well as those that make use of machine learning-based approaches, that is support vector machines (SVMs) and classifiers that employ quadratic discriminant analysis (QDA). It is important to note that within the context of machine learning-based classification, DTW is however utilised, not as a classification technique but as a *dichotomy transformation* for the purpose of converting feature vectors in feature space into a dissimilarity vector representation in dissimilarity space. A typical template matching-based system developed in this study is conceptualised in Figure 1.3, while a typical machine learning-based system developed in this study is conceptualised in Figure 1.4.

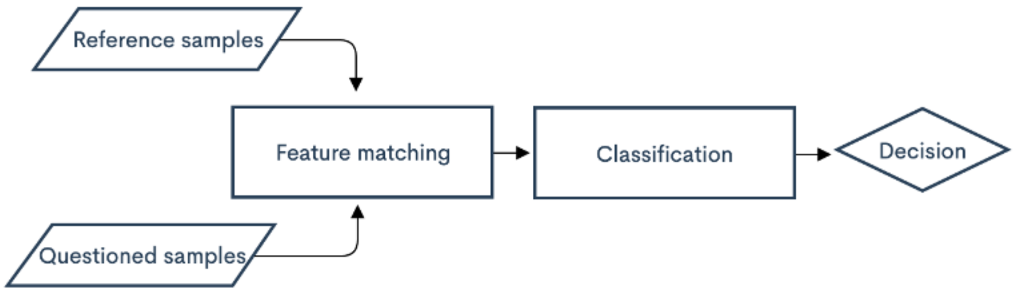


Figure 1.3: A conceptualisation of a typical template matching-based system developed in this study.

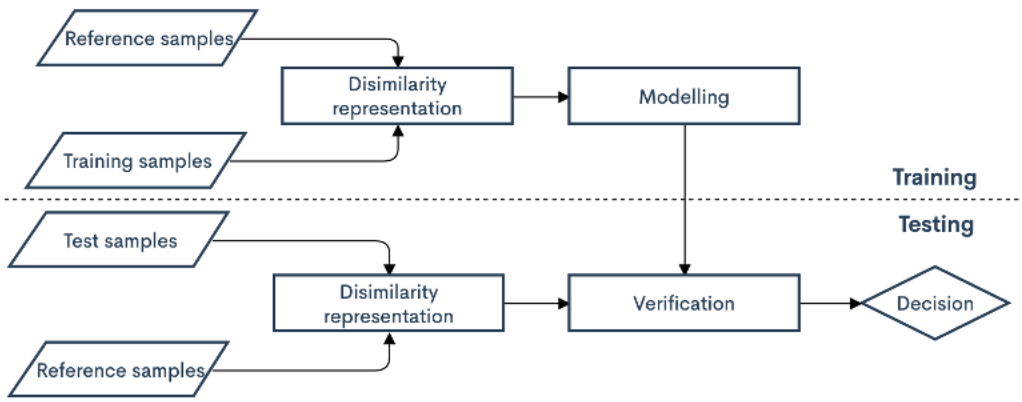


Figure 1.4: A conceptualisation of a typical supervised learning-based system developed in this study.

1.5 Data

The experiments conducted in this study are based on a *subset* of the CEDAR-LETTER data set consisting of 100 writers. The CEDAR-LETTER data set consists of sample letters that were reproduced three times by each candidate writer. As previously stated, this study uses a text-dependent approach to writer authentication, with the word "the" being the targeted word. For each of the sample documents provided, each instance of the word "the" is segmented from the rest of the document.

1.6 System design

In this section, the systems developed in this study are discussed in more detail. As previously mentioned, the systems developed in this study are divided into those that make use of simple distance classifiers (SDCs), that is those based on the Euclidean distance and those based on a DTW distance, as well as those that employ machine learning-based approaches, that is SVMs and QDA classifiers.

1.6.1 Simple distance classifiers

Feature extraction

Each of the extracted samples (instances of the word "the") undergo several preprocessing procedures so as to ensure that the features can be optimally extracted. Once the suitable binary images are obtained, the discrete Radon transform (DRT) extracts a set of features associated with T predetermined angles. These feature sets are then normalised in such a way that they are all of the same dimension d. This feature extraction technique therefore yields a feature set consisting of T, d-dimensional feature vectors in feature space. This process is conceptualised in Figure 1.5.

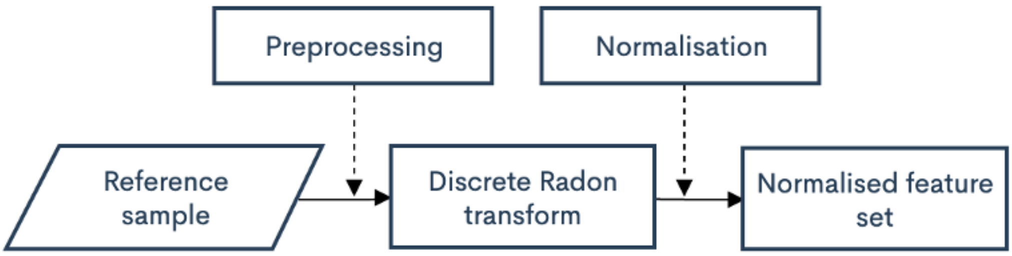


Figure 1.5: A conceptualisation of the proposed SDC-based feature extraction process.

Feature matching

A model is developed by matching all of the reference samples associated with a specific writer to non-reference samples associated with other writers, which results in negative dissimilarity values. Positive dissimilarity values are obtained by matching all of the reference samples associated with a specific writer to the remaining positive test samples associated with the aforementioned writer. The set of dissimilarity values are subsequently normalised using the writer-specific mean and standard deviation. Finally, the set of normalised dissimilarity values is used to obtain an optimal global threshold for classification purposes. This process is conceptualised in Figure 1.6.

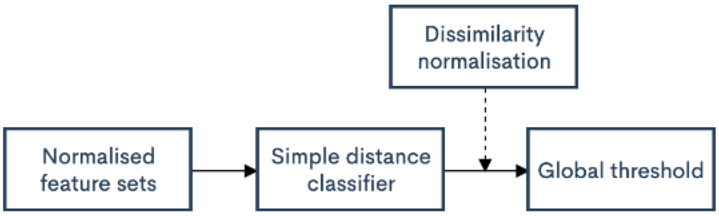


Figure 1.6: A conceptualisation of the proposed SDC-based feature matching process.

Verification

When presented with a questioned sample and claim of ownership, the system first compares the questioned sample to the entire set of reference samples belonging to the claimed writer using the appropriate SDC, yielding a set of dissimilarity values. These values are then normalised using the writer-specific mean and standard deviation. The normalised values are subsequently averaged. A global threshold is finally applied to this average value in order to determine class membership. The process is conceptualised in Figure 1.7.

1.6.2 Machine learning-based classifiers

Dissimilarity vector construction

Similar to the proposed SDC-based systems, the extracted samples undergo several preprocessing procedures. Following this, features are extracted from each image through the discrete Radon transform (DRT), producing a normalised feature set. This process results in a feature set consisting of T, d-dimensional vectors. The extracted feature set is subsequently converted from

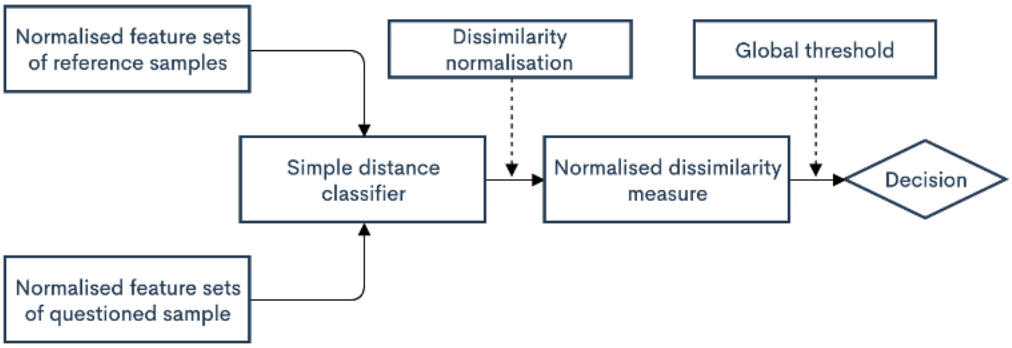


Figure 1.7: A conceptualisation of the verification process within the context of the proposed SDCs.

feature space into dissimilarity space for the purpose of following a writer-independent approach to writer authentication. Each feature set extracted from a training or questioned sample is matched to the feature set extracted from an authentic reference sample by means of a DTW-based dichotomy transformation. This converts the questioned feature set consisting of T, d-dimensional feature vectors into a single T-dimensional *dissimilarity vector* which is represented in dissimilarity space. This process is conceptualised in Figure 1.8.

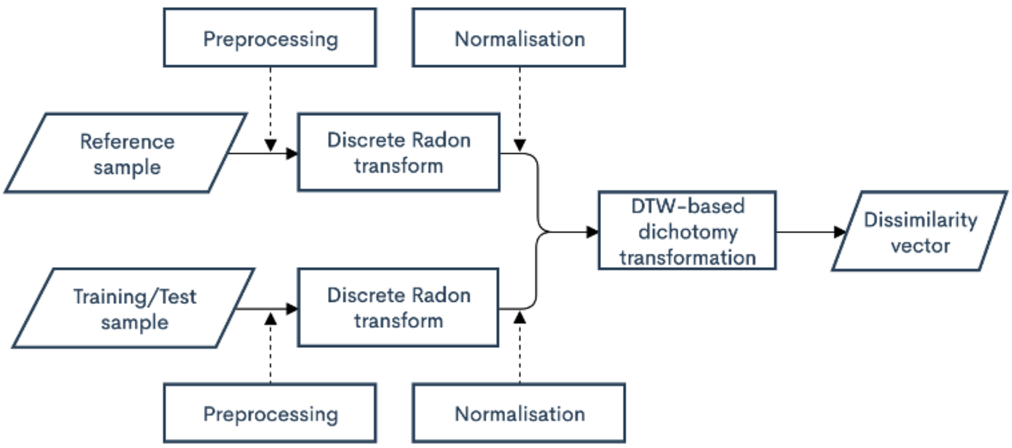


Figure 1.8: A conceptualisation of the proposed machine learning-based sample representation.

Modelling

A *single* writer-independent model is trained for classification purposes. Positive and negative samples associated with a number of so-called guinea pig writers are matched to positive reference samples in order to obtain positive and negative dissimilarity vectors. The dissimilarity vectors are subsequently normalised using writer-specific statistics. These normalised vectors are finally used to train a SVM or QDA-based classifier. This process is conceptualised in Figure 1.9.

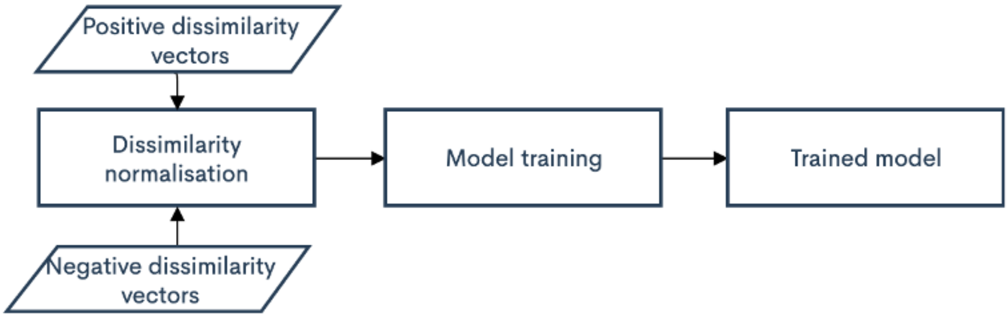


Figure 1.9: A conceptualisation of the proposed protocol for machine learning-based modelling.

Verification

When a questioned sample (that form part of the test set) is claimed to belong to a specific writer (that does not constitute a guinea pig writer), it is first matched to a number of positive reference samples, after which the resulting normalised dissimilarity vectors are presented to the trained SVM or QDA-based classifier. This results in a number of distance measures relative to the decision boundary. These distance measures are converted into partial confidence scores using a conventional logistic function. The partial confidence scores are averaged so that a final score is obtained. An appropriate threshold is finally applied to the aforementioned final score. This process is conceptualised in Figure 1.10.

Although SVMs and QDA-based classifiers typically produce discrete output, this study employs a modified approach that converts the initial output (a distance measure relative to the decision boundary in dissimilarity space) into a confidence score.

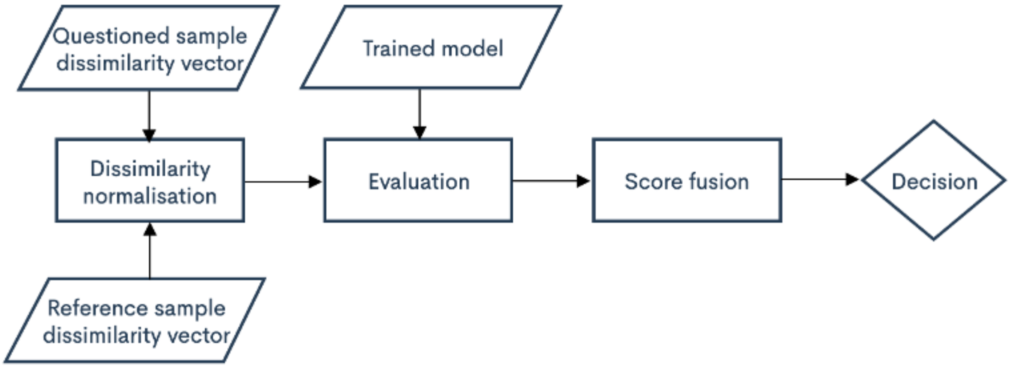


Figure 1.10: A conceptualisation of the proposed protocol for machine learning-based verification.

1.7 Abbreviated results

After rigorous experimentation, the proficiency of the systems developed in this study is summarised in this section.

In the case of the proposed SDC-based systems, the DTW-based system outperforms the Euclidean distance-based system. The DTW-based system achieves an optimal equal error rate (EER) of 20.6%, while the Euclidean distance-based classifier achieves an optimal EER of 32.01%.

In the case of machine learning-based authentication the number of reference samples available for each writer is varied between eight and sixteen, with the number of feature vectors varied between 2 and 1024. The proposed SVM-based system achieves an optimal average AUC and EER of 93% and 14.93% respectively when 16 reference samples and 1024 feature vectors are employed. The acronym AUC denotes the area under the curve in receiver operating characteristic (ROC) space. The proposed QDA-based system achieves an optimal average AUC and EER of 76% and 31.51% respectively when 16 reference samples and 512 feature vectors are employed. The proposed SVM consistently outperforms the proposed QDA-based classifier.

The proficiency of the systems developed in this study is compared to a number of existing systems in Section 6.6.

1.8 Contribution of this study

This study proposes and develops several techniques that are novel within the context of handwriting authentication. These include:

- **A DTW based dichotomy transformation for writer-independent handwriting representation.**
  Feature vectors are matched based on similarity rather than location in

feature space, which minimises the adverse effects associated with intra-class variability and results in a superior classifier. It was shown in this study that a classifier primarily based on a DTW distance outperforms a classifier primarily based on an Euclidean distance.

- **A writer-specific dissimilarity normalisation strategy within the context of writer-independent modelling.**
  The proposed writer-independent system, in conjunction with writer-specific dissimilarity normalisation, improves interclass variability. This results in a superior system

- **A proficient writer-independent handwriting authentication system.**
  When compared to existing systems that make use of similar techniques, the systems developed in this study have been shown to be at least on par and, in some cases, outperform other systems.

1.9 Thesis outline

This thesis is structured in the following manner:

Chapter 2: Literature review. A concise review of existing handwriting authentication and recognition systems are presented. A number of existing methodologies are discussed and compared.

Chapter 3: Image processing and feature extraction. This chapter discusses how the raw images are extracted from a sample document, as well as the preprocessing techniques that allow for optimal feature extraction

Chapter 4: Classifiers. This chapter discusses and compares the classification techniques considered in this study.

Chapter 5: Modelling and verification. This chapter describes a number of additional processing techniques that are performed on the extracted features. Within the context of the proposed SDC-based systems, it describes the relevant feature matching protocols. Within the context of the proposed machine learning-based classifiers, it introduces the concept of a DTW-based dichotomy transformation for dissimilarity vector-based representation. This chapter furthermore discusses the writer-independent approach to writer authentication adopted in this study. It continues to discuss how, with the aid of a suitable classification technique, handwriting authentication models are developed. The verification protocol is finally discussed.

Chapter 6: Experiments. This chapter describes the data set considered in this study and also outlines the adopted experimental protocol. It continues to describe the experiments conducted in this study for the purpose of gauging the proficiency of the classification models developed in this thesis. A detailed analysis of the proficiency of the proposed systems is provided, after which the aforementioned proficiency is compared to those of existing systems

Chapter 7: Conclusion and future work. This chapter provides concluding remarks regarding the work presented in this thesis. In addition to this, it discusses avenues for future research which may result in improved system proficiency.

Chapter 2

Literature review

2.1 Introduction

Writer authentication is a well-established field of research within the context of pattern recognition. In the offline scenario, the objective of writer authentication is to verify whether a questioned sample of handwriting in the form of a digitised (scanned-in) image was produced by the claimed writer. Traditionally, this requires the extraction of a set of features from the questioned handwriting, which is then matched to a statistical model trained with a number of samples known to belong to the claimed writer.

The concept of writer recognition is similar to writer authentication, except for the fact that the objective is to ascertain which writer, amongst a set of candidate writers, most likely produced the questioned handwriting sample. In this chapter existing offline writer authentication and recognition systems are discussed, while emphasising the research progress that has been made to date. Various systems and their contribution to this field of research will therefore be discussed. For a comprehensive review on the current state of the art within the context of writer authentication and recognition, the reader is directed to Rehman *et al.* (2019). Although the focus of the study is on writer authentication, a number of writer recognition systems, for which the methodologies and techniques are applicable to writer authentication, will also be discussed in this chapter to provide context. The systems discussed in this chapter are categorised according to the relevant classification protocol, which include distance-based classifiers, machine learning-based classifiers (that do not employ deep learning) and finally deep learning-based classifiers (neural networks).

Each system will be discussed by detailing the following:

- The relevant authors and the year of publication;
- The employed feature extraction protocol;

- The classification methodology employed during modelling and verification;
- The type of data set(s) being considered; and
- The experimental results.

It is important to note that the systems discussed in this chapter are evaluated by considering a wide range of data sets, which include data sets containing handwriting in languages besides English. It is furthermore important to note that a true benchmark for evaluating writer authentication systems does not exist. For these reasons, the reported proficiency of the systems discussed in this chapter should only serve as a guideline for the reader when gauging the performance of the systems proposed in this study. It will be demonstrated that the proficiency of the best system proposed in this study, which is novel within the context of writer authentication, at least matches those of a significant number of systems discussed in this chapter.

2.2 Distance-based classifiers

Distance-based classifiers employ a distance measure between samples for classification purposes. A smaller distance between samples is typically associated with a higher probability that the samples were authored by the same writer. These approaches do not involve the training of parameters for a suitable statistical model and are therefore less complex than typical machine learning-based approaches. The most widely used distance measures include the Euclidean distance, the Chi-square distance, the Manhattan distance, and the Hamming distance. The systems discussed in this section employ the aforementioned distance measures.

In Fiel and Sablatnig (2012) a writer retrieval and recognition system that makes use of local features is proposed. Only the writer recognition module of this system is discussed here. The proposed system considers the provided sample to be a texture image, which implies that image binarisation is not required. Feature extraction is performed by implementing the scale invariant feature transform (SIFT). The extracted features are clustered through k-means clustering, after which the histogram depicting the occurrences of the cluster centres are used to construct a codebook. Three hundred cluster centres are empirically selected. The chi-square distance χ^2 between histograms is employed as a similarity measure. In order to achieve writer recognition, the histogram associated with the questioned sample is compared to each histogram associated with the candidate writers, after which the one associated with the lowest chi-square distance, that is the closest neighbour, is selected. Two data sets are considered for experimental purposes. The first data set, that is the IAM data set, contains handwriting data from 657 different writers, where each writer produced 59 pages of handwriting, and consists of 1539 images in total. The second data set, that is the TrigraphSlan data set, contains 188 scanned images from 47 different writers. When 300 clusters are selected within the context of writer recognition, and only the closest neighbour is considered, a recognition rate of 90.8% is achieved. When the five closest neighbours are considered, the recognition rate increases to 96.7%, while the recognition rate increases to 97.5% when the ten closest neighbours are considered.

Dhandra and Vijayalaxmi (2015) propose a novel approach to text-dependent writer recognition. The proposed system is based on Kannada handwriting which is predominantly used in Southwestern India. The Radon transform and discrete cosine transform are employed for the purpose of extracting textural and structural features respectively. The aforementioned features are assembled into a single feature vector that, as a result, contains directional multi-resolution spatial features. It is established that when the features associated with two or more words are combined, better results are achieved. For classification purposes, the respective Euclidean distances between the feature vector associated with a questioned word and the feature vectors associated with the same word, but produced by the different candidate writers, are calculated. A nearest neighbour classifier is subsequently employed by assigning the questioned word to the candidate associated with the smallest Euclidean distance. Five-fold cross-validation is employed. A database containing words produced by 25 different writers is considered. An average accuracy of 93.3% is reported when only a single word is considered, while average accuracies of 97.2% and 100% are respectively achieved when combined features from three, and more than three, different words are considered.

Bensefia and Paquet (2016) propose a writer authentication system that makes use of graphemes. A grapheme is defined as an elementary graphical shape resulting from handwriting segmentation. This system considers only one handwritten word which is segmented into a number of graphemes. Ideally, each grapheme should constitute a single character within the word. Each grapheme is typically represented as a string. For classification purposes, the elementary components of two graphemes, that is the respective symbols within the two strings, are compared. This is achieved using the Wagner-Fisher algorithm, which transforms one string into another, and forms the basis of computing the edit distance between the two graphemes. Low values of the transformation cost is associated with a high probability that the word has been produced by the same writer. The system is evaluated on the IAM database that contains handwriting samples produced by 657 different writers. For evaluation purposes, 44 different words produced by 50 different writers, are considered. A true positive rate of 87% and a false negative rate of 27% are reported.

Al-Maadeed *et al.* (2016) proposes the use of novel geometric features for offline writer recognition. The aforementioned features include directional features, curvature-based features and tortuosity-based features. The proposed methodology is applied to both English and Arabic handwriting samples. The images are first binarised using Otsu's thresholding algorithm. The abovementioned features are extracted by employing both quadruple-order chain codes and edge-based directional features. For feature matching, the respective distances between each of the corresponding feature vectors associated with the reference and questioned documents are calculated. Three separate distance measures are employed, that is the χ^2 distance, the L^2 or Euclidean distance, as well as the L^1 distance. In addition to these distances, the "list" distances at element level for each pair of vectors are computed. The aforementioned distances are subsequently combined using kernel discriminant analysis with spectral regression (SR-KDA). The proposed system is evaluated on the IAM and QUWI handwriting databases, and achieves recognition rates of 82% and 87% respectively.

Ghiasi and Safabakhsh (2010) propose an efficient method for text-independent writer recognition. The proposed method considers small fragments of connected components and avoids complex patterns. A codebook is constructed by considering histograms that depict the frequency of occurrence of the different fragments that appear in the handwriting. The authors propose a novel approach for extracting these fragments, which they claim to be more efficient than the so called "Segment on Y-minima" method. The Segment on Y-minima method separates connected components at those points that result in "meaningful" fragments. This results in a loss of certain information pertaining to the shape of the local minima. The χ^2 function is employed for the purpose of measuring the similarity between two handwriting samples, and in so doing, recognising the author of the questioned sample. The proposed methodology is evaluated by considering three varieties of a Farsi database that include texts of short, medium, and large length, resulting in recognition rates of 92.7%, 99.4% and 100% respectively.

Wang (2019) proposes an offline writer recognition system that employs different levels of features. The proposed system combines local directional chain-code features (LDCFs) and global improved texture features (GITFs). During the first stage, LDCFs are extracted and subsequently matched so as to generate a candidate sample set. The candidate sample set is subsequently refined using the GITFs. For the purpose of feature matching two methodologies are employed, namely rough matching and fine matching. Within the context of rough matching, the weighted Euclidean distance between the questioned sample and the samples in the candidate set is calculated for the purpose of conducting k-nearest neighbour classification. The aforementioned k-nearest neighbour classification protocol is subsequently improved upon through the implementation of an iKNN (iterative k-nearest neighbour) algorithm. Within the context of the iKNN algorithm, the data dimension attribute of the nearest feature vector associated with the questioned sample is selected from the known handwriting samples. The class label is specified using the frequency of known samples that correspond to the attribute values associated with the nearest neighbour of each dimension within the context of the questioned sample. This process is implemented using the GITFs and constitutes fine matching. By considering a database that consists of Chinese language samples produced by 203 writers, an optimal recognition rate of 98.4% is achieved through the combination of LDCFs and GITFs.

2.3 Machine learning-based classifiers

Machine learning-based classifiers employ machine learning algorithms for classification purposes. In this section the focus will be on machine learning-based classifiers that do not employ deep learning. Within the context of *writer-specific* supervised learning, different models are typically trained for different writers by considering positive samples known to belong to the writer in question, as well as negative samples known to belong to other writers. A single *writer-independent* model is however proposed in this thesis, as will be discussed at a later stage. Popular algorithms include Bayesian classifiers, support vector machines (SVMs) and k-nearest neighbour classifiers. This section describes existing classification systems based on these algorithms.

Kamal *et al.* (2014) propose a robust local feature-based approach to writer authentication. The proposed system extracts writer-specific characteristics from images of text which are then analysed to authenticate the writer. The process involves subdividing an image of handwritten text into a grid of sub-images. Similarly shaped sub-images are subsequently assigned to a specific class through correlation so as to identify repetitive patterns of a particular piece of handwriting that embodies the writer's handwriting style. For each of the aforementioned classes, only the sub-image that is the closest to all of the other sub-images belonging to said class is retained as representative of that class. Feature vectors are subsequently extracted for each class, and the process is repeated for each document associated with a specific writer so as to create writer-specific models. In order to authenticate the author of a questioned document, the image of the questioned handwriting is subdivided into sub-images and clustered using a similar protocol than the one used for the reference images. Similar feature vectors are extracted for the questioned sample, and a Bayesian classifier is finally employed to authenticate the writer. The system is evaluated on the IAM database and a recognition rate of 94% is reported.

In Pandey and Seeja (2018) a writer recognition system that makes use of projection profile representations of graphemes is proposed. The proposed system uses the Zhang-Suen thinning algorithm to process the images. This results in strokes that are one pixel thick through the middle of the handwriting. The thinned handwriting sample is subsequently divided into graphemes. The graphemes are further processed so as to only retain those foreground pixels with no more than two neighbouring foreground pixels. Further cleaning is conducted for the purpose of eliminating graphemes that are smaller than a predefined threshold. This process is repeated for all the documents. The graphemes are represented as projection profiles using a novel technique that defines a projection profile as the position of the first foreground pixel in each row. This is implemented by inserting all the graphemes into bounding boxes with a fixed size of 43×43 pixels. Each grapheme is then represented by a vector with 43 entries. A dictionary or codebook of the vectors is created using k-means clustering. The proposed system is tested on the IAM handwriting database that contains 1539 scanned pages of handwriting from 657 writers of which 650 were used for creating the dictionary. The system achieves an accuracy within the range of 84% to 88.5%. The accuracy varies depending on the size of the cluster.

Christlein *et al.* (2017) propose a robust strategy for writer recognition. The proposed system employs RootSIFT descriptors. The extracted features constitute Gaussian mixture model (GMM) supervectors that describe the characteristics of each individual writer. The GMM-UBM (universal background model) is adapted for the purpose of modelling the descriptors. An exemplar-SVM is used for classification purposes. The system is evaluated by considering three data sets, that is the ICDAR data set, the CVL data set and the KHATT data set. The highest mean average precision (mAP) reported for these data sets is 89.4%, 99.3% and 98.0% respectively.

Bertolini *et al.* (2016) proposes a writer-independent classifier for writer recognition that is based on dissimilarity representation. The system employs local binary pattern (LBP) and local phase quantization (LPQ) texture descriptors. For model training, the system extracts feature vectors from reference samples and subsequently computes the dissimilarity between these feature vectors using an appropriate dichotomy transformation. To generate positive dissimilarity vectors, genuine samples from the same writer are compared. On the other hand, reference samples from different writers are used to generate negative dissimilarity vectors. For classification purposes, a two-class SVM with a Gaussian kernel is employed. The system is evaluated on the QUWI database, that contains 4 068 handwritten samples from 1 017 different writers. When evaluating the text-dependent approach using LBPs, an equal error rate (ERR) of 31.2% is reported for English samples, while an EER of 25.7% is reported for Arabic samples. When the aforementioned text-dependent approach is evaluated using LPQ descriptors, EERs of 20.3% and 19.4% are reported for English and Arabic samples respectively.

Saranya and Vijaya (2013) propose a text-dependent writer recognition model that employs a SVM. Several distinctive features are extracted from scanned images of handwritten text to form a feature vector describing a writer's writing style. The proposed model associates each writer with a class, and for each class a binary SVM-based classifier is created. The model is subsequently trained using linear, polynomial and radial basis function (RBF) kernels. When presented with a questioned sample, each classifier renders a decision, and the classifier associated with the highest positive decision value is assigned the label. The model is implemented using 1000 data instances of 100 words written by 10 writers. The best results are reported for the polynomial kernel, that is an accuracy of 94.3%, while accuracies of 77.6% and 92.3% are reported for the linear and RBF kernels respectively.

2.4 Deep learning-based classifiers

Neural network-based classifiers employ architectures that are not explicitly based on statistical models and typically require large amounts of training data.

In Mukherjee and Ghosh (2020) a writer recognition scheme based on the individuality of the handwriting in question is proposed. In the proposed experiment, three separate types of features are extracted, namely lined-based features, word-based features, as well as character-based features. The line-based features are extracted by means of a strategy that employs projection profiles, while the word-based features are extracted by first segmenting each word, after which the slant angle is calculated. The aforementioned segmentation is conducted in order to verify whether the word in question is produced in a continuous or discrete way. The connectivity of each word is subsequently measured by means of a *connectivity ratio*. Feature matching is achieved by employing a multilayer perceptron (MLP), which constitutes a type of artificial neural network, as well as the so-called K-STAR method, which is based upon an entropic distance measure. A data set containing English handwritten samples authored by subjects of Bengali origin with a Bengali-medium schooling background is used for the purpose of experimentation. The experiment achieves optimal results of 93.54% and 95.69% using the MLP and the K-STAR protocol respectively.

Vásquez *et al.* (2018) investigate the feasibility of holistic graphometric features for the purpose of writer recognition. A large amount of feature-based graphometric information is extracted from handwritten samples for the purpose of constructing feature vectors. The system employs a neural network and a score fusion block classifier. A majority voting algorithm (MVA) is applied for the purpose of selecting the most frequent solution (mode) to be the correct classification. The system is evaluated by considering a data set associated with 100 different writers, where each writer produced 10 samples. The system achieves an accuracy of 99.8% when 40% holdout cross-validation is used, while an accuracy of 100% is achieved when 50% holdout cross-validation is employed.

Kumar and Kaur (2017) propose a handwriting recognition system that employs both neural network and SVM classifiers. The performance of the proposed model is reportedly similar to that of existing state-of-the-art writer recognition systems. The extracted features are based on slant, as well as slant energy, skew, pixel distribution, curvature, and entropy. Directional features are initially extracted, followed by principal component analysis (PCA) for the purpose of dimension reduction. Fisher's linear discriminant analysis (LDA) is finally employed for the purpose of extracting the most discriminative features. The system is implemented on the IAM data set and an accuracy of 95% is achieved.

2.5 Concluding remarks

In this chapter various handwriting authentication and recognition systems were discussed. As mentioned earlier, there is no standard protocol for evaluating handwriting authentication systems. Furthermore, the systems discussed in this chapter were evaluated using different data sets, which makes it impossible to compare the efficacy of the different models. However, a number of the systems discussed in this chapter employ techniques that are relevant to this study. The systems proposed by Bensefia and Paquet (2016) and Saranya and Vijaya (2013) are relevant to this study since they also adopt a *text-dependent* approach to handwriting modelling.

The protocol proposed by Bertolini *et al.* (2016) is also relevant to this study, since it is also based on the extraction of dissimilarity vectors for the purpose of implementing a *writer-independent* approach to handwriting modelling. The aforementioned writer-independent approach to handwriting modelling is crucial, since it allows for the authentication of writers that have not been part of the original training set. In addition to this, the use of an Euclidean distance measure by Dhandra and Vijayalaxmi (2015) and SVM classifiers by Saranya and Vijaya (2013) and Bertolini *et al.* (2016) are relevant to this study. The proficiency of the systems developed in this study is compared to a number of existing systems in Section 6.6.

The next chapter discusses the image acquisition, pre-processing and feature extraction techniques utilised in this study.

Chapter 3

Image processing and feature extraction

3.1 Introduction

This chapter describes the image processing and feature extraction techniques adopted in this study. Firstly, the methods used to source samples of handwritten text are discussed, followed by the preprocessing procedures performed to ensure an accurate representation of each handwritten sample. The feature extraction method, which is based on the calculation of the discrete Radon transform of the preprocessed raw sample images, is subsequently discussed. The feature vectors are finally normalised in such a way that each feature set constitutes a scale-invariant representation of the sample in question. The aforementioned protocol is depicted graphically in Figure 3.1.

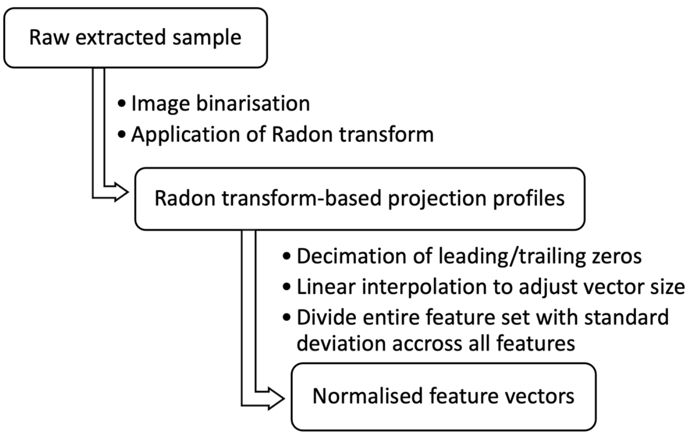
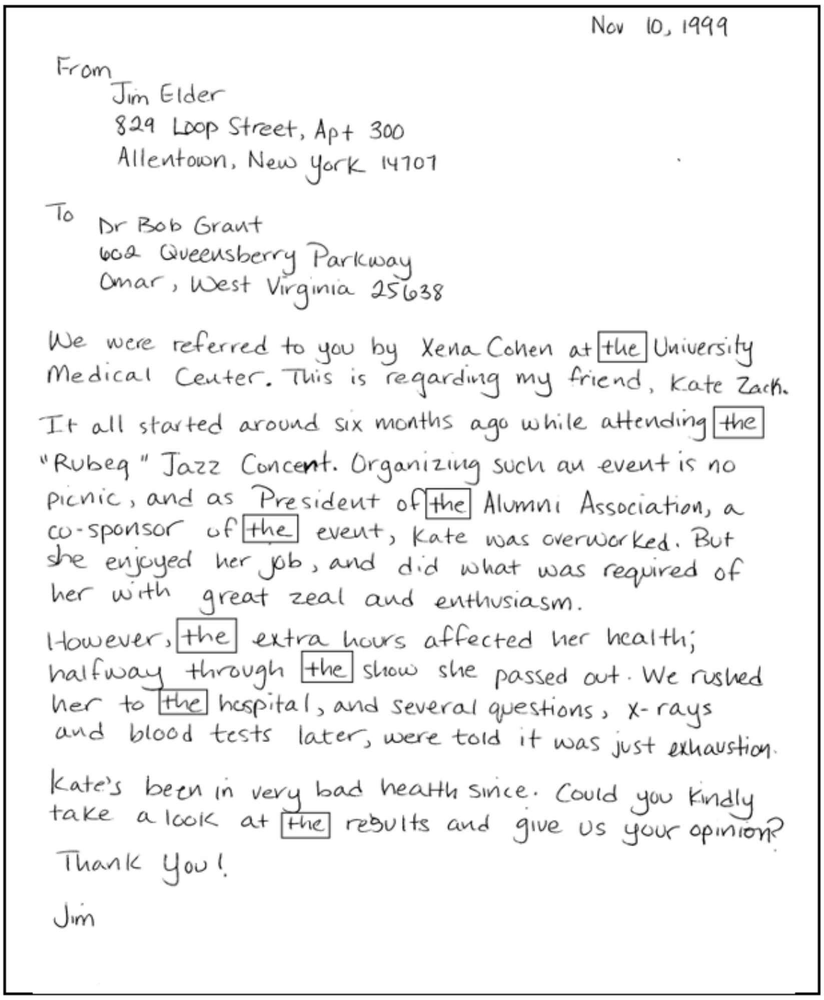


Figure 3.1: Conceptualisation of the image preprocessing and feature extraction protocols proposed in this study.

3.2 Image processing

In this study, a *text-dependent* approach to writer authentication is adopted. The selected word for this purpose is the word "the", which is extracted from the handwritten text associated with each of the individual writer's handwritten samples. A typical handwritten sample is depicted in Figure 3.2.



Nov 10, 1999

From
Jim Elder
829 Loop Street, Apt 300
Allentown, New York 14707

To
Dr Bob Grant
602 Queensberry Parkway
Omar, West Virginia 25638

We were referred to you by Xena Cohen at the University Medical Center. This is regarding my friend, Kate Zach.

It all started around six months ago while attending the "Rubeq" Jazz Concert. Organizing such an event is no picnic, and as President of the Alumni Association, a co-sponsor of the event, Kate was overworked. But she enjoyed her job, and did what was required of her with great zeal and enthusiasm.

However, the extra hours affected her health; halfway through the show she passed out. We rushed her to the hospital, and several questions, x-rays and blood tests later, were told it was just exhaustion.

Kate's been in very bad health since. Could you kindly take a look at the results and give us your opinion?

Thank You!

Jim

Figure 3.2: An example of a sample document produced by a writer with the selected word "the" boxed.

Each writer produced the *same* handwritten text in his/her *own* handwriting and it is clear that eight instances of the word "the" are present per sample (boxed in Figure 3.2). The proposed system is *semi-automated* in the sense that the instances of the word "the" are *manually* selected (extracted). Within the context of a fully automated system, handwritten word-spotting (Fischer *et al.* (2012), Rodríguez-Serrano and Perronnin (2009), Manmatha *et al.* (1996)) may be employed for the purpose of determining the locations of the targeted word. This however falls outside the scope of this research.

The extraction process results in raw, grey-scale images with dimensions $m \times n$ pixels where the respective values of m and n are determined by the size of the extracted sample text. From Figure 3.2 it is clear that the different instances of the relevant word "the" (boxed) may vary in size. It is ensured that the selected word "the" is manually extracted in such a way that it does not contain any text associated with the surrounding words. In order to provide the reader with some context, the first instance of the boxed word "the" in Figure 3.2 has dimensions of 141×82 pixels. The resolution of the manually extracted images is therefore typically low. At this stage, the image contains only the required handwritten text, that is the word "the", while no other handwritten or machine-printed text is present. Figure 3.3 shows three examples of the aforementioned images produced by three different writers.

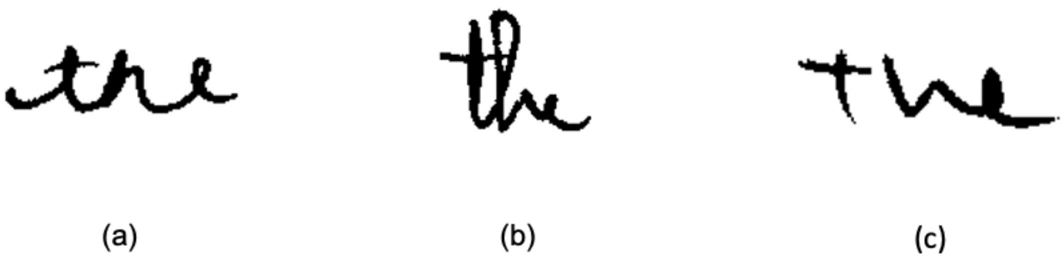


Figure 3.3: Examples of extracted sample images of the word “the” produced by three different writers.

Preprocessing is required to render the raw images suitable for feature extraction. Within the context of this study, the relevant preprocessing techniques involve image segmentation and image binarisation. Due to the fact that the images are very small and of low resolution, noise reduction through the implementation of, for example, a median filter is not advisable, since it inevitably leads to the decimation of important pen strokes.

3.2.1 Boundary box constraint

Image segmentation is achieved by enforcing a boundary box constraint. The boundary box of a sample constitutes the smallest rectangular box that completely contains the sample in question. The dimensions of a particular bounding box depend on the size of the sample. Figure 3.4 illustrates this concept.

Figure 3.4: Illustration of the application of the bounding box constraint. (a) The raw extracted image. (b) The image after the bounding box constraint has been applied. Note that the bounding box is rendered black for visualisation purposes only and is in reality the same colour as the background (white).

3.2.2 Image binarisation

The original grey-scale image is subsequently subjected to binarisation, that is a process which converts it into binary format. This process differentiates between two different pixel types, that is those pixels associated with pen strokes and those associated with the background, and allows for efficient shape analysis. The input image I consists of pixels with a grey level intensity within the range [0, 255], where 0 corresponds to the darkest pixels (black) and 255 corresponds to the lightest pixels (white). The original grey-scale image I is converted into a binary image I^B by applying a global threshold value of $p \in [0, 255]$ to each pixel $i = 1, \dots, mn$ as follows,

$$I_i^B = \begin{cases} 1, & \text{if } I_i < p \\ 0, & \text{if } I_i \geq p \end{cases}. \tag{3.1}$$

This results in an image where a pixel value of 1 is associated with a pen stroke and a pixel value of 0 is associated with the background (see Figure 3.5).

Figure 3.5: An example of a binarised sample image containing the word "the". The pen strokes are associated with a binary value of 1 and the background with a binary value of 0.

3.3 Feature extraction and scale normalisation

This study employs the Radon transform, first proposed within the context of *continuous* functions in Radon (1917), for the purpose of feature extraction. When the *discrete* Radon transform (DRT) is applied to an image for a given set of angles, the projection profiles of the aforementioned image are calculated in the corresponding directions. Each projection profile is therefore obtained by calculating the sum of the pixel intensities along a number of parallel non-overlapping beams in a specific direction. It is sufficient to generate these projection profiles from equally distributed angles in the interval $\theta \in [0°, 180°)$. Each projection profile may be considered to be an unnormalised feature vector, while the entire set of projection profiles constitutes a feature set (observation sequence) in feature space. In this study, a high narrow peak within the projection profile associated with an angle θ denotes a prominent pen stroke perpendicular to the x' axis (see Figure 3.6).

Coetzer (2005) demonstrated that the DRT is well suited for representing a single word of handwritten text within the context of offline signature verification. Since this study employs a text-dependent approach to offline writer authentication, and specifically considers instances of the single word "the", the DRT is also deemed appropriate within this context.

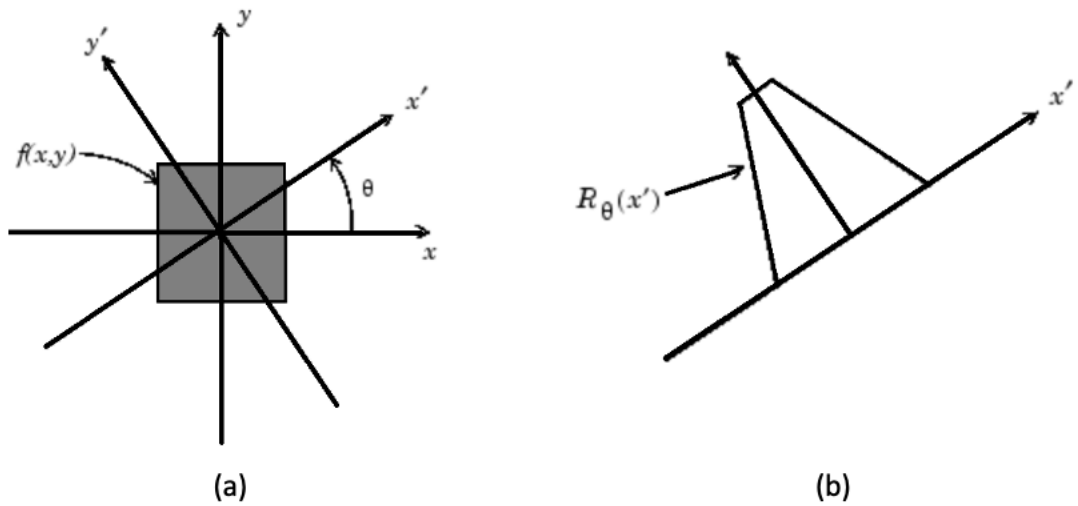


Figure 3.6: A graphical conceptualisation of the *continuous* Radon transform $R_\theta(x')$ of a function $f(x, y)$ for a specific angle θ.

A more formal definition of the Radon transform is now provided. The *continuous* Radon transform $R_\theta(x')$ constitutes an integral transform which takes the line integral of a two-dimensional function, $f(x, y)$, parallel to the y' axis (see Figure 3.6). The y' axis is perpendicular to the x' axis, where the x'

axis constitutes a rotated version of the horizontal x axis through an angle θ (counterclockwise). The continuous Radon transform is represented as

$$R_\theta(x') = \int_{-\infty}^{\infty} f(x' \cos\theta - y' \sin\theta, x' \sin\theta + y' \cos\theta)\mathrm{d}y' \tag{3.2}$$

where

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.3}$$

The *discrete* Radon transform (DRT) is a discretisation of the above (see Figure 3.7) and its application within the context of this study is now outlined.

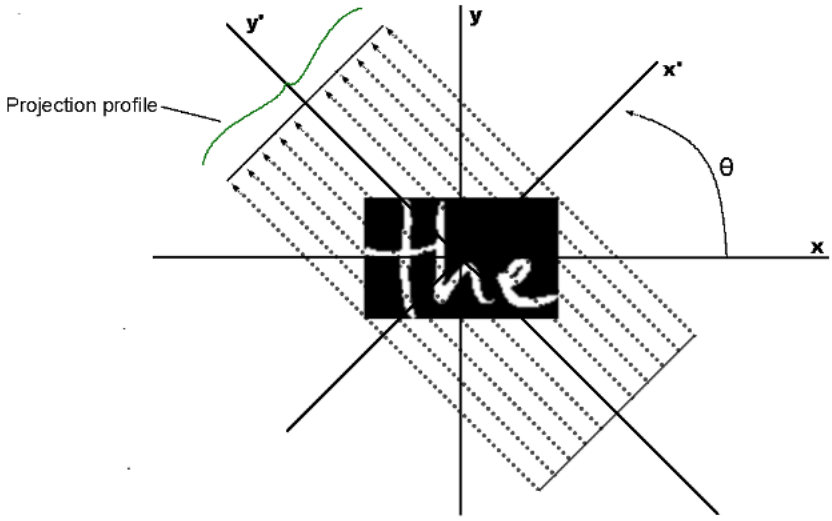


Figure 3.7: Conceptualisation of the *discrete* Radon transform being applied to a binary image containing the word "the" for a specific angle θ.

The classical application of the DRT is within the field of x-ray tomography. The DRT simulates the acquisition of CT scans ("projection profiles") from a number of angles, while the original image is estimated (reconstructed) through the application of the discrete inverse Radon transform (DIRT). Note that the DIRT is not calculated at any stage throughout the course of this study. An array of "sources" for the projection beams is rotated around the image $f(x, y)$. The set of beams originating from the sources are equidistant, that is 1 pixel unit apart of each other and non-overlapping. The DRT is calculated for T equally distributed angles in the interval $\theta \in [0°, 180°)$, while N beams are utilised for each angle, where N is determined by the dimensions of the image. For each angle θ, the "density of the matter" (pixel values) through which the beams from the source pass is "accumulated" (summed) and "detected" at the "sensor". Figure 3.7 conceptualises the DRT being applied to a sample image containing the word "the". The number of angles T and the number of beams per angle N determine the overall accuracy (resolution) with which the features are extracted by means of the DRT. Figure 3.8 (a) depicts a binary image containing the word "the", while its DRT is depicted as a grey-scale image in Figure 3.8 (b). Each column of the image depicted in Figure 3.8 (b) constitutes a projection profile (unnormalised feature vector) generated from a specific angle. For more information relating to the implementation of the DRT, the reader is referred to Beylkin (1987) and Toft (1996).

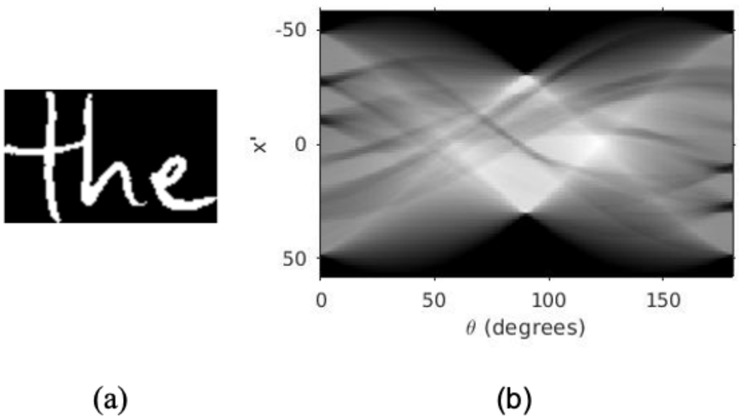


Figure 3.8: (a) A binary image of the selected word "the". (b) The DRT-based representation where $T = 180$. Each column represents a projection profile (unnormalised feature vector).

Each projection profile (unnormalised feature vector) that results from the application of the DRT has to be *normalised* in order to render the resulting set of normalised feature vectors (normalised feature set or normalised observation sequence) a scale invariant representation of the image in question. This normalisation process is now discussed in detail.

It is clear from the colunms of Figure 3.8 (b) that each projection profile (unnormalised feature vector) contains zero-valued entries (black pixels) at both ends. The leading and trailing zero-valued entries contained in each feature vector are removed to reduce external influences on the spatial variations of the extracted features. This process is referred to as the *decimation* of leading and trailing zeros. The resulting feature vectors therefore only consist of the remaining non-zero components. A conceptualisation of the resulting output is shown in Figure 3.9. Note that, after the aforementioned procedure,

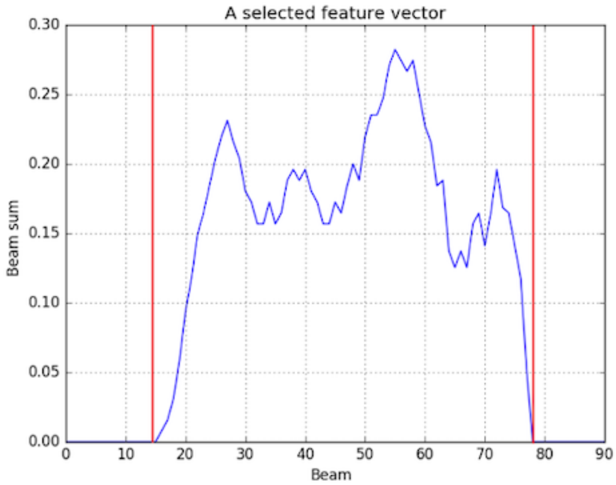


Figure 3.9: A visualisation of the removal (decimation) of the leading and trailing zeros of a feature vector. The red lines indicate where the respective leading and trailing zeros are cut off.

the respective dimensions of the feature vectors (within a feature set) will vary. In order to ensure that all the feature vectors has the *same* dimension, linear interpolation is employed for the purpose of stretching or shrinking each feature vector to a predetermined fixed dimension, d. In order to ensure scale invariance, each feature set (feature matrix), in which each feature vector (column) has a dimension of d, is finally divided by the standard deviation across all the matrix entries. The feature normalisation process is visualised in Figure 3.10.

Each of the T normalised d-dimensional feature vectors $\mathbf{x}_i, i = 1, \ldots, T$ are packed as a column into a matrix $\mathbf{X}$, as follows,

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_T\} \tag{3.4}$$

or

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1T} \\ x_{21} & x_{22} & \ldots & x_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{d1} & x_{d2} & \ldots & x_{dT} \end{bmatrix} \tag{3.5}$$

where $\mathbf{X}$ constitutes the entire normalised feature set (observation sequence), $\mathbf{x}_i$ an individual observation, and x_{ij} an individual feature.

The normalisation process can be summarised as follows:

1. **Remove zeros** - Remove all the leading and trailing zeros from each of the feature vectors.
2. **Adjust length of feature vectors** - The dimension of each feature vector is adjusted through linear interpolation.
3. **Divide feature set by standard deviation** - Each feature is divided by the standard deviation across the entire feature set.
4. **Create a matrix that depicts the entire feature set** - The normalised feature vectors are packed as columns into a matrix $\mathbf{X}$.

3.4 Concluding remarks

In this chapter, the preprocessing techniques required to manipulate the extracted images of the handwritten samples were discussed. It is explained that an effective preprocessing protocol is essential for the purpose of enhancing the quality of the image data before they are converted into binary format. The suitability of the discrete Radon transform for the purpose of extracting features from the aforementioned binary images is subsequently outlined. Finally, the process of normalising the feature vectors for the purpose of obtaining scale-invariant representations of the original binary images containing the word "the" was discussed in detail.

The next chapter discusses the classification techniques that are used in this study for the purpose of handwriting modelling. These protocols include simple distance classifiers (SDCs) which are based on the Euclidean distance and a dynamic time warping-based (DTW-based) distance respectively, support vector machines (SVMs) and quadratic discriminant analysis (QDA).

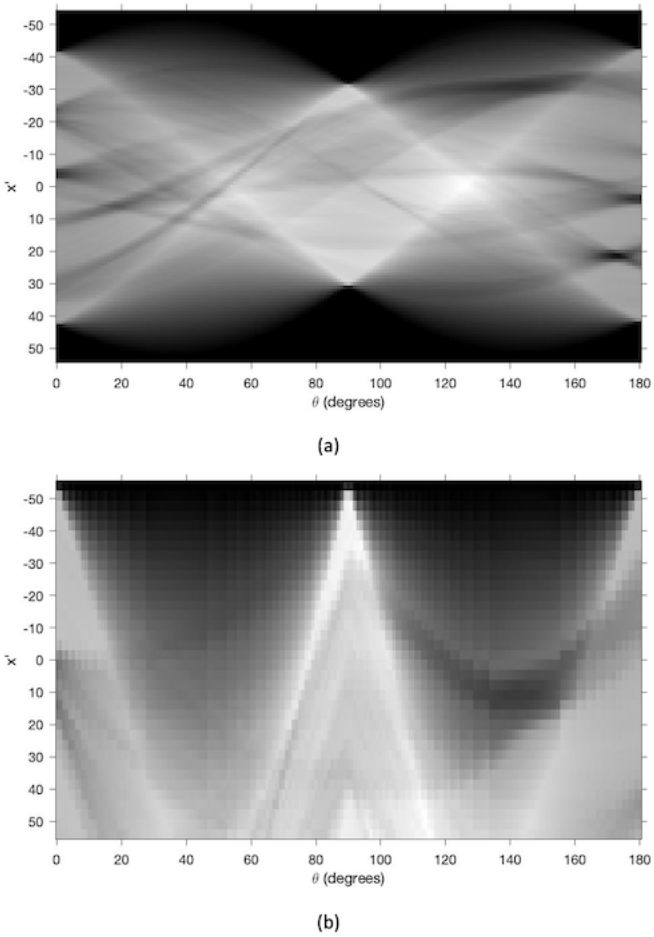


Figure 3.10: Conceptualisation of feature normalisation. (a) An unnormalised feature set. Each column of the depicted matrix represents an unnormalised feature vector. (b) The corresponding normalised feature set. Each column represents a normalised feature vector. Note that, for each column, the leading and trailing zeros have been decimated, after which each decimated vector has been stretched or shrunk to a fixed dimension of $d = 180$ through linear interpolation. Each matrix entry has finally been divided by the standard deviation across all the matrix entries.

Chapter 4

Classifiers

4.1 Introduction

The protocol for extracting a scale-invariant feature set from each instance of the manually selected word "the" within a handwritten sample was outlined in the previous chapter. This chapter introduces the theoretical concepts pertaining to the classifier modelling and feature matching techniques adopted in this study.

Template matching techniques that involve simple distance classifiers (SDCs) based on the Euclidean distance and a dynamic time warping-based (DTW-based) distance respectively, are first discussed in Section 4.2. Machine learning-based classifiers, in particular support vector machines (SVMs) and classifiers that employ discriminant analysis, are subsequently introduced in Section 4.3.

4.2 Simple distance classifiers

Within the context of this study, simple distance classifiers (SDCs) constitute template matching techniques that calculate the average distance between the corresponding feature vectors associated with a test (questioned) and reference feature set. The Euclidean distance and a dynamic time warping-based (DTW-based) distance are considered for this purpose.

4.2.1 Euclidean distance

When $\mathbf{x}^{\text{Ref}}$ represents a feature vector associated with an authentic (positive) reference sample of a particular writer, and $\mathbf{x}^{\text{Test}}$ represents a feature vector associated with a test (questioned) sample, the Euclidean distance between these feature vectors is calculated as follows,

$$D_{\text{Eucl}} = \sqrt{(\mathbf{x}^{\text{Test}} - \mathbf{x}^{\text{Ref}})'(\mathbf{x}^{\text{Test}} - \mathbf{x}^{\text{Ref}})}, \tag{4.1}$$

where it is assumed that $\mathbf{x}^{\text{Test}}$ and $\mathbf{x}^{\text{Ref}}$ represent column vectors and $'$ denotes the transpose.

The dissimilarity between the normalised feature sets $\mathbf{X}^{\text{Ref}}$ and $\mathbf{X}^{\text{Test}}$ is denoted by $D_{\text{Eucl}}(\mathbf{X}^{\text{Test}}, \mathbf{X}^{\text{Ref}})$ and constitutes the *average* of the respective Euclidean distances between the corresponding columns (normalised feature vectors) of the aforementioned matrices (feature sets). The Euclidean distance-based comparison of two normalised feature vectors $\mathbf{x}^{\text{Ref}}$ and $\mathbf{x}^{\text{Test}}$ is conceptualised in Figure 4.1.

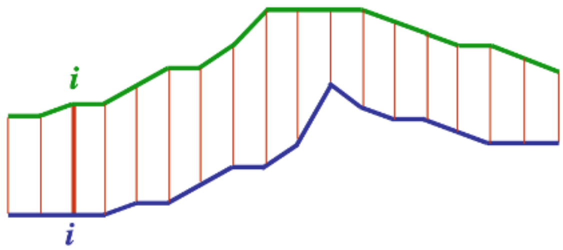


Figure 4.1: Conceptualisation of two normalised feature vectors $\mathbf{x}^{\text{Ref}}$ (green) and $\mathbf{x}^{\text{Test}}$ (blue) being compared. The i^{th} point of $\mathbf{x}^{\text{Ref}}$ is matched to the i^{th} point of $\mathbf{x}^{\text{Test}}$ (Tsiporkova (2012)).

The proposed Euclidean distance-based classifier aims to establish whether $\mathbf{X}^{\text{Test}}$ is sufficiently close to $\mathbf{X}^{\text{Ref}}$ in order to conclude that they belong to the same class. A global threshold τ is employed for this purpose. A questioned feature set $\mathbf{X}^{\text{Test}}$ is deemed to be authentic if and only if

$$D_{\text{Eucl}}(\mathbf{X}^{\text{Test}}, \mathbf{X}^{\text{Ref}}) \leq \tau,$$

and is deemed fraudulent otherwise. The Euclidean distance-based classifier is discussed in more detail in Chapter 5 where, amongst other things, a writer-specific dissimilarity normalisation protocol is proposed.

4.2.2 Dynamic time warping

A DTW-based distance measure may also be used to calculate the dissimilarity between feature sets $\mathbf{X}^{\text{Ref}}$ and $\mathbf{X}^{\text{Test}}$. The DTW algorithm non-linearly aligns feature vectors according to their feature values so as to obtain an optimal path in the two-dimensional space that relates similar prominent features, like peaks and valleys (see Figure 4.2). The alignment is deemed optimal if the Euclidean distance between the *aligned* feature vectors is a *minimum*. The search space is usually restricted within a band (with a predetermined band width) around the diagonal. If the aforementioned band width is specified to be zero and the search space is restricted to the diagonal, the optimal path will inevitably coincide with the diagonal, in which case the DTW-based distance is equivalent to the Euclidean distance. When the feature vectors have the same dimension (as is the case in this study), a band width of approximately one tenth of this dimension is typically employed. The DTW-based dissimilarity between feature sets $\mathbf{X}^{\text{ref}}$ and $\mathbf{X}^{\text{test}}$, denoted by $D_{\text{DTW}}(\mathbf{X}^{\text{Test}}, \mathbf{X}^{\text{Ref}})$, is the average DTW-based distance between the corresponding normalised feature vectors $\mathbf{x}_i^{\text{Ref}}$ and $\mathbf{x}_i^{\text{Test}}$, where $i = 1, \ldots, T$.

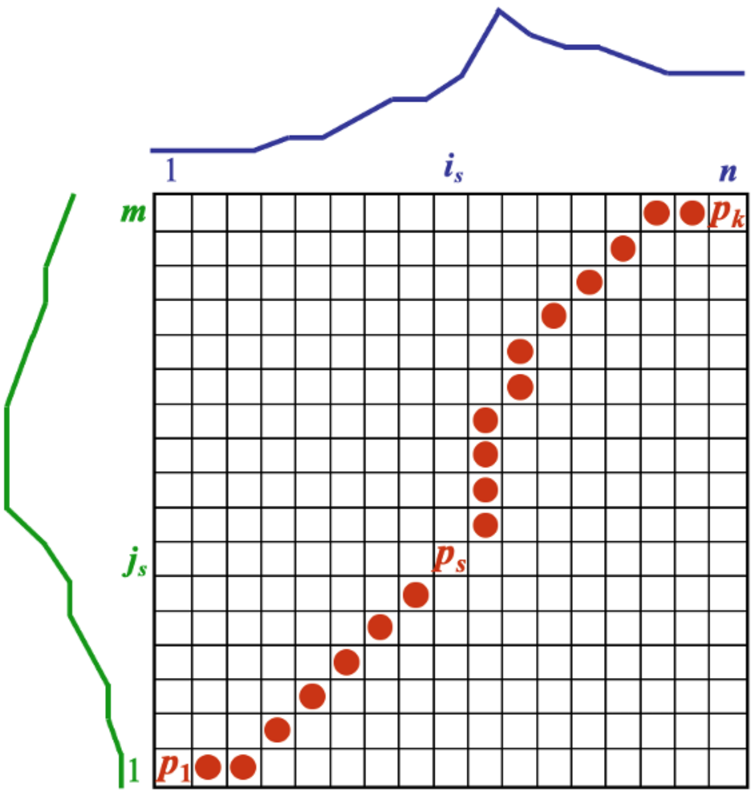


Figure 4.2: A graphical representation of the DTW algorithm used to calculate the distance between a reference feature vector $\mathbf{x}^{\text{Ref}}$ (green) and a test feature vector $\mathbf{x}^{\text{Test}}$ (blue). The algorithm identifies similar elements within these vectors and constructs an optimal path ($P = p_1, p_2, \ldots, p_n$) between them. In this example (and in this study), the feature vectors in question have the same dimension (Tsiporkova (2012)).

Similar to the Euclidean distance-based measure, class membership is determined by means of a global threshold τ such that $\mathbf{X}^{\text{Test}}$ is deemed authentic if and only if

$$D_{\text{DTW}}(\mathbf{X}^{\text{Test}}, \mathbf{X}^{\text{Ref}}) \leq \tau.$$

The DTW distance-based classifier is discussed in more detail in Chapter 5 where, amongst other things, a writer-specific dissimilarity normalisation protocol is proposed.

4.2.3 Euclidean distance versus DTW-based distance

Intra-class variability refers to the slight variations that may occur within a number of samples produced by the same writer. The Euclidean distance-based approach is sensitive to these variations. The variations typically manifest as lateral shifts in the local maxima/minima of two DRT-based feature vectors associated with different samples, but produced by the same writer, as conceptualised in Figure 4.3. The Euclidean distance's inability to compensate for the aforementioned variability is a significant weakness.

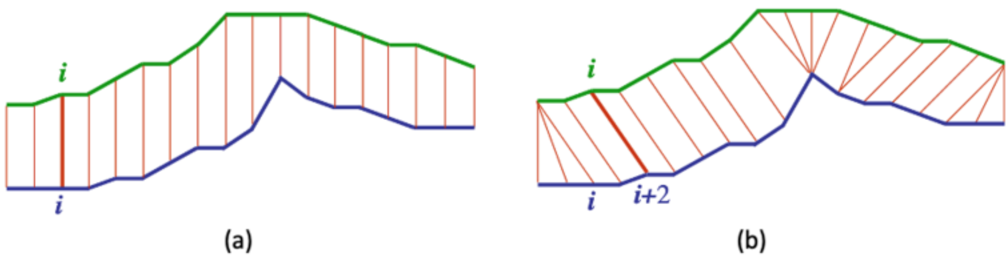


Figure 4.3: Comparison of SDC techniques within the context of two normalised feature vectors $\mathbf{x}^{\mathrm{Ref}}$ (green) and $\mathbf{x}^{\mathrm{Test}}$ (blue). It is clear that in the case of (a), that is the Euclidean distance, the corresponding points of the normalised feature vectors are directly aligned (the i^{th} feature of $\mathbf{x}^{\mathrm{Ref}}$ is matched to the the i^{th} feature of $\mathbf{x}^{\mathrm{Test}}$), whereas, in the case of (b), that is the DTW-based distance, each point is optimally aligned according to the prominent features, which results in the i^{th} feature of $\mathbf{x}^{\mathrm{Ref}}$ being matched to the $(i+2)^{\mathrm{th}}$ feature of $\mathbf{x}^{\mathrm{Test}}$ (Tsiporkova (2012)).

The DTW-based distance is superior to the Euclidean distance, since it is able to detect (and compensate for) misaligned features and therefore reduces intra-class variability. It is however important to note that DTW also reduces inter-class variability, but typically to a lesser extent than is the case for intra-class variability.

4.3 Machine learning-based classifiers

Machine learning (ML) algorithms are able to learn from observed or training data without the need of being explicitly programmed. This results in the construction of a model that is able to identify patterns or structures learnt form the training data, enabling it to predict certain outcomes. Within the context of this study, appropriate ML algorithms are developed for the purpose of classification. The adopted algorithms are support vector machines (SVMs), as well as classifiers that employ linear and quadratic discriminant analysis.

4.3.1 Support vector machines

In order to illustrate the functionality of SVMs, the training set $\mathbf{Z}$ that consists of a set of positive vectors, $\mathbf{Z}_{\text{pos}}$, and a set of negative vectors, $\mathbf{Z}_{\text{neg}}$, is considered. A SVM aims to construct a decision boundary that assigns a questioned vector to one of two different classes (negative and positive). The decision boundary is in the form of a hyperplane in feature space that is obtained through the *maximum margin*, that is, the distance between data points of both sets closest to the hyperplane. The larger the margin, the lower the error of the classifier.

An infinite number of hyperplanes can be created between two linearly separable sets. Figure 4.4 demonstrates how different hyperplanes can be generated for the same sample data. A SVM aims to generate the optimum hyperplane that maximises the margin while ensuring data points of both sets are classified correctly. Figure 4.5 demonstrates the principle of an optimal margin. Support vectors are the data points within the dataset which are the nearest to the hyperplane, which would otherwise alter the position of the hyperplane if removed. They are considered critical elements of the dataset.

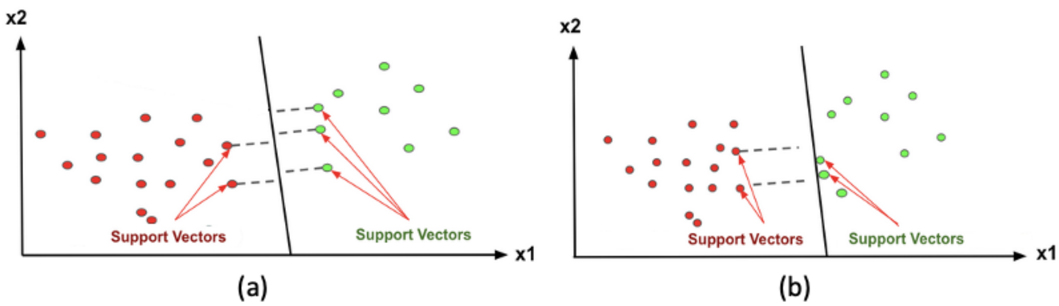


Figure 4.4: A conceptualisation of two hyperplanes separating the same dataset into two classes. (a) An example of an ideal hyperplane where the support vectors have the same distance (margin) from both classes. (b) A less ideal hyperplane where the distance (margin) associated with the positive support vectors (green dots) is much smaller than that of the negative class (red dots) (Liu (2020)).

The hyperplane membership function is represented as

$$f(\mathbf{z}) = \mathbf{w}'\mathbf{z} + b, \tag{4.2}$$

where $\mathbf{w}$ is the weight vector and b is the bias. Given the training set $\mathbf{Z}$, using a *hard margin*, $\mathbf{z} \in \mathbf{Z}$ is assigned to $\mathbf{Z}^{\text{pos}}$ if $f(\mathbf{z}) \geq 0$ and assigned to $\mathbf{Z}^{\text{neg}}$ if $f(\mathbf{z}) < 0$.

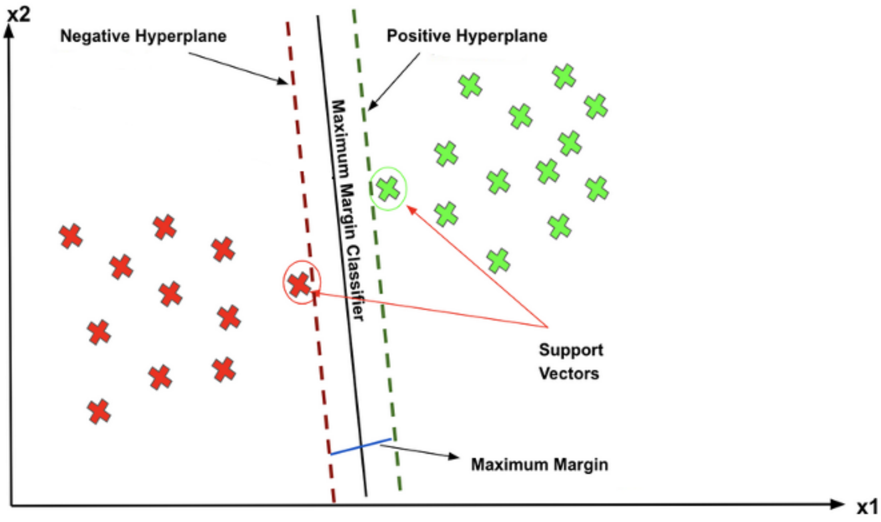


Figure 4.5: An illustration of a hyperplane and its associated optimal margin represented in a two-dimensional feature space (Liu (2020)).

Soft and hard margins

Ideally, the data points are linearly separable such that two parallel hyperplanes can be selected to separate the two sets while ensuring the largest possible margin, known as a *hard margin*. In most cases though, data points are not linearly separable, and a *soft margin* is required. While condcting a *soft margin* approach, margin violation may occur to allow some of the data points to be on the incorrect side of the hyperplane (misclassified). A SVM attempts to develop a hyperplane that minimises misclassification while maximising the margin. The degree of tolerance when finding the decision boundary is determined by the regularisation hyper-parameter C. The higher the value of C, the higher the penalty of misclassification. The hyperplane membership function is thus modified as follows

$$f(\mathbf{z}) = \mathbf{w}'\mathbf{z} + b + C\sum_{i=1}^{N}\mathcal{E}_i, \tag{4.3}$$

where $\mathcal{E}$ is a set of N slack variables, allowing for margin errors and misclassifications. Figure 4.6 draws a distinction between the application of a *hard margin* and a *soft margin* approach.

A smaller margin hyperplane is constructed for larger values of C if that hyperplane is more sufficiently suited for classifying the data correctly. Conversely, smaller values of C cause a larger margin hyperplane, often resulting in misclassification as shown in Figure 4.7.

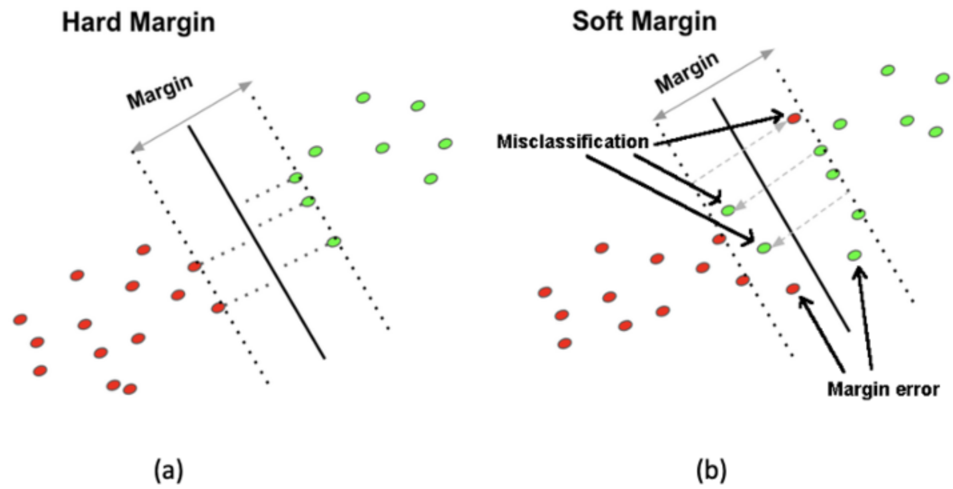


Figure 4.6: A conceptualisation of separating hyperplanes in two-dimensional feature space. (a) The dataset is linearly separable and demonstrates a *hard margin*. (b) The dataset is linearly non-separable and demonstrates a *soft margin* where some data points are misclassified (Liu (2020)).

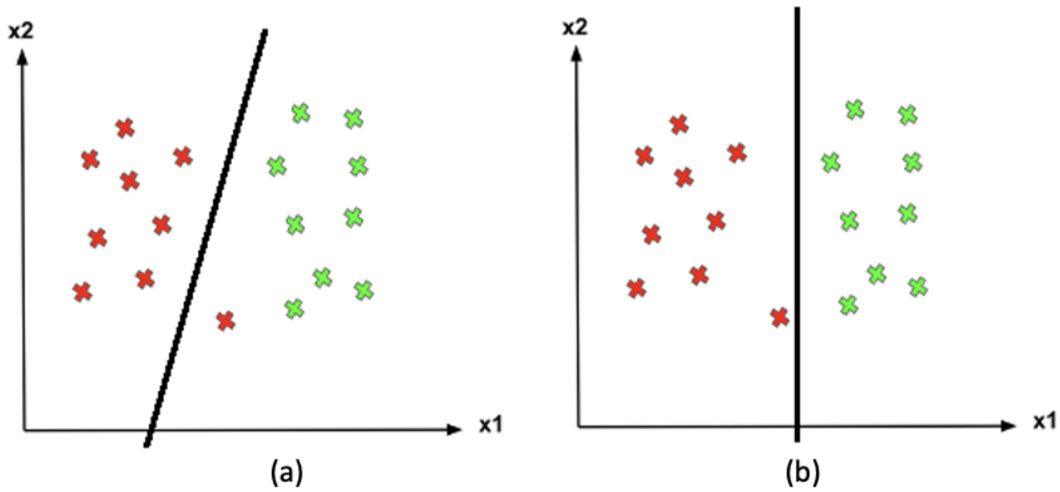


Figure 4.7: The effect of different sizes of the regularisation parameter C on the hyperplane selection. (a) A hyperplane constructed with a low value of C resulting in misclassification (b) A hyperplane constructed with a large value of C (Liu (2020)).

Kernel trick

In order to classify non-linearly separable data, SVMs employ the so-called kernel trick. The kernel trick transforms a low dimensional input space into a higher dimensional space. Figure 4.8 illustrates the application of the kernel trick on non-linearly separable data. The membership function of the non-linear decision boundary (in the original feature space) is represented as

$$f(z) = \mathbf{w}'\phi(\mathbf{z}) + b + C\sum_{i=1}^{N}\mathcal{E}_i, \tag{4.4}$$

where $\phi(\mathbf{z})$ denotes a suitable mapping function.

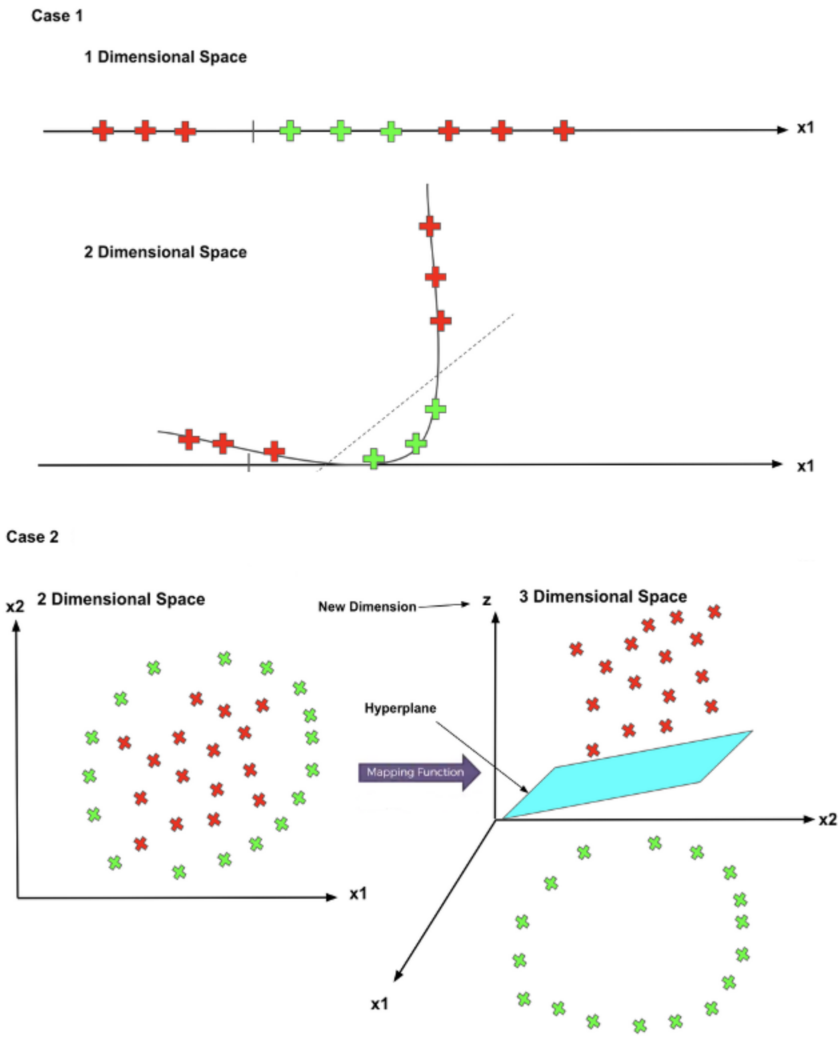


Figure 4.8: A conceptualisation of the application of the kernel trick on non-linearly separable data. **Case 1** shows two classes of data in one-dimensional space. After the kernel trick is applied, the data is represented in a hypothetical two-dimensional kernel space where it is now possible to obtain a separating hyperplane represented by the dashed line. **Case 2** shows two classes of data in two-dimensional feature space. After the kernel trick is applied, the data is represented in a hypothetical three-dimensional kernel space where it is now possible to obtain a separating hyperplane (Liu (2020)).

Some of the frequently used kernels include the radial basis function (RBF) kernel, the sigmoid kernel and the polynomial kernel. Since the RBF kernel is employed in this study, it is discussed in more detail. Figure 4.9 draws a comparison between a linear decision boundary and a non-linear RBF kernel-based decision boundary.

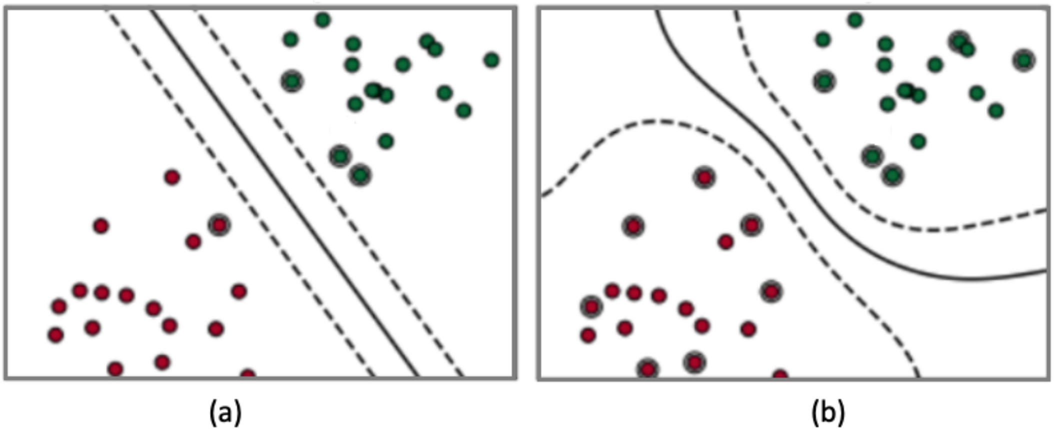


Figure 4.9: A comparison between (a) an linear decision boundary and (b) an RBF-based decision boundary (Chen (2019)).

The RBF kernel is a transformer that generates new features by measuring the distance between the centres of the data points as follows,

$$K(\mathbf{z}, \mathbf{z}') = \exp\left(-\frac{||\mathbf{z} - \mathbf{z}'||^2}{2\sigma^2}\right), \tag{4.5}$$

where $||\mathbf{z} - \mathbf{z}'||^2$ is the squared Euclidean distance, and σ is a free parameter that determines the kernel width, that is the region of influence (decision boundary) of each data sample. The effect that increasing the value of σ has on a hyperplane is demonstrated in Figure 4.10.

4.3.2 Discriminant analysis

Discriminant analysis (DA) is a supervised classification method that determines a set of prediction equations based on independent variables for the purpose of assigning a questioned object to one of a number of classes. Assuming that $\mathbf{z}$ belongs to one of κ classes with a class-specific probability density $f(\mathbf{z})$, DA attempts to divide the dataspace into κ disjoint regions, with each region representing a class. In order to classify an object $\mathbf{z}$ by means of DA, the probability that the object belongs to each class has to be estimated. The class associated with the highest probability is deemed the desired output. For

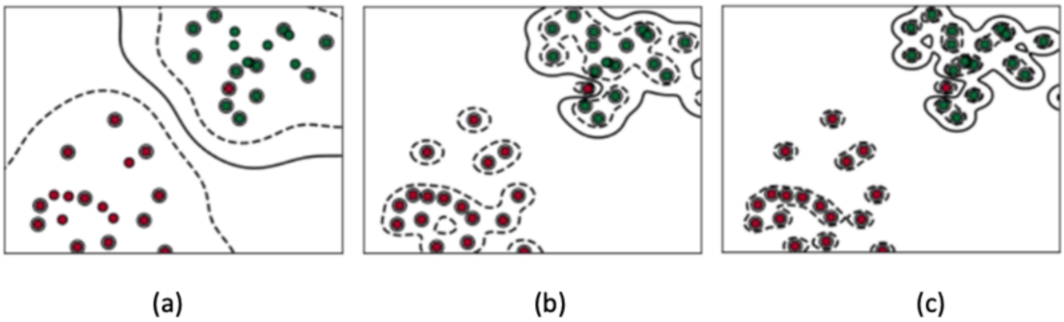


Figure 4.10: The effect of an increased value of σ when implementing the RBF kernel to generate a decision boundary for non-linearly separable data. (a) A relatively low value of σ. (b) A larger value of σ. (c) A relatively high value of σ (Chen (2019)).

multi-class classification, DA uses Bayes' theorem, which relies on the posterior probability $P(k_i|\mathbf{z})$ defined as

$$P(k_i|\mathbf{z}) = \frac{P(\mathbf{z}|k_i)P(k_i)}{\sum_{j=1}^{\kappa} P(\mathbf{z}|k_j)P(k_j)}, \tag{4.6}$$

where $P(k_i)$ is the pior probability of class k_i and $P(\mathbf{z}|k_i)$ is the class conditional density of class k_i represented as

$$P(\mathbf{z}|k) = f_k(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{p}{2}}|\mathbf{C}_k|^{\frac{1}{2}}} \exp^{-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu}_k)'\mathbf{C}_k^{-1}(\mathbf{z}-\boldsymbol{\mu}_k)}, \tag{4.7}$$

where p is the number of predictors, $\boldsymbol{\mu}_k$ is the mean vector and $\mathbf{C}_k$ is the covariance matrix for each class. These statistics are defined as

$$\boldsymbol{\mu}_k = \frac{1}{N_k}\sum_{i=1}^{N_k} \mathbf{z}_i, \tag{4.8}$$

$$\mathbf{C}_k = \frac{1}{N_k - 1}\sum_{i=1}^{N_k} (\mathbf{z}_i - \boldsymbol{\mu}_k)(\mathbf{z}_i - \boldsymbol{\mu}_k)', \tag{4.9}$$

where N_k is the number of training observations in the k^{th} class. These statistics uniquely identify class k_i.

In this study a two class problem is being investigated with a positive (authentic) class k_1 and a negative (fraudulent) class k_2. The class prior probabilities are assumed to be equal, that is $P(k_1) = P(k_2)$. The class membership function is defined as

$$f_{\text{DA}}(\mathbf{z}) = f_{k_1}(\mathbf{z}) - f_{k_2}(\mathbf{z}) \tag{4.10}$$

such that $\mathbf{z}$ is classified as belonging to class k_1 (authentic) if $f_{\mathrm{DA}}(\mathbf{z}) \geq 0$. Similarly, $\mathbf{z}$ is classified as belonging to class k_2 (fraudulent) if $f_{\mathrm{DA}}(\mathbf{z}) < 0$.

The above-mentioned model is applied to either *linear discriminant analysis* (LDA) or *quadratic discriminant analysis* (QDA) depending on the choice of covariance matrix of each class. LDA makes use of a pooled covariance estimate where $\mathbf{C}_k = \mathbf{C}$ for all k, which implies that the variance is shared across all classes. QDA makes use of class-specific covariance estimates resulting in a non-linear decision boundary (see Figure 4.11). As a result, when using QDA, the similarity between a questioned sample and each class-specific PDF is based on a Mahalanobis distance measure.

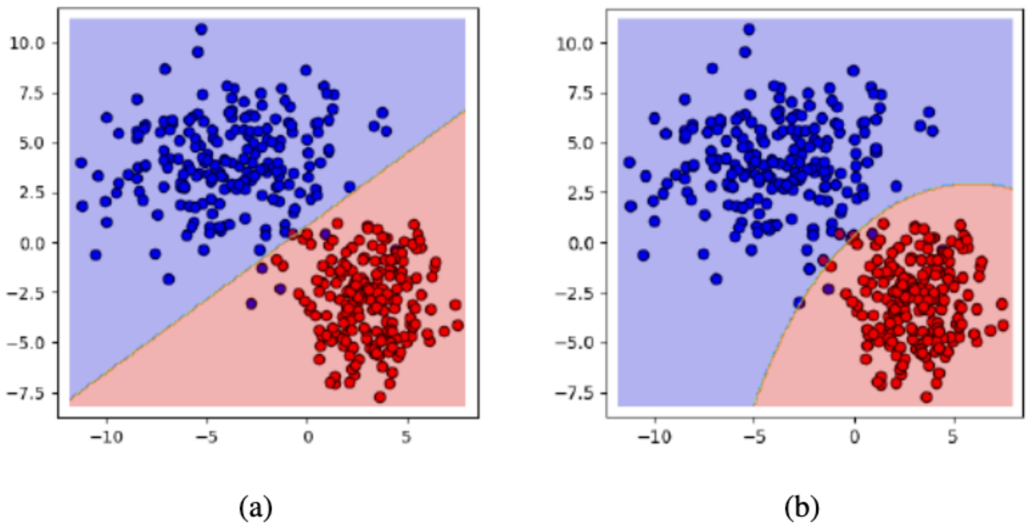


Figure 4.11: An conceptualisation of (a) LDA and (b) QDA applied on the same dataset (Ghojogh and Crowley (2019)).

4.4 Concluding remarks

This chapter presented an overview of the theoretical aspects relating to the classification techniques employed in this study. The relevant simple distance classifiers (SDCs) were first discussed, followed by support vector machines (SVMs) and classifiers based on discriminant analysis (DA).

The next chapter is reserved for a discussion on how models for handwriting authentication can be constructed by considering the classifiers introduced in this chapter.

Chapter 5

Modelling and verification

5.1 Introduction

In the previous chapter simple distance classifiers (SDCs) that enable the authentication of a questioned handwritten sample by means of template (feature) matching, as well as the relevant machine learning-based classification techniques employed in this study, were discussed

In this chapter the aforementioned template matching-based authentication protocol is discussed in more detail by, amongst other things, proposing a suitable writer-specific dissimilarity normalisation protocol.

Within the context of the machine learning-based approaches, a DTW-based dichotomy transformation for the purpose of generating dissimilarity vectors are proposed. This facilitates a writer-independent representation in dissimilarity space. A writer-specific normalisation strategy is subsequently employed so as to render the classes more separable in dissimilarity space.

5.2 Simple distance classifiers

Within the context of SDCs, Euclidean and DTW-based distance measures are employed for the purpose of authenticating a questioned document. A writer-specific dissimilarity normalisation protocol is proposed.

5.2.1 Feature matching

It is assumed that a reference set $\underline{\mathbf{X}}_r^{\mathrm{Ref}} = \left\{ \left(\left(\mathbf{X}_r^{\mathrm{Ref}}\right)_{(1)}, \left(\mathbf{X}_r^{\mathrm{Ref}}\right)_{(2)}, \ldots, \left(\mathbf{X}_r^{\mathrm{Ref}}\right)_{(K)} \right) \right\}$ that contains K positive samples is available for each writer r.

A typical test set for writer r contains an additional $N^{(+)}$ positive samples as well as a large number of negative samples $N^{(-)}$ (belonging to other writers) that are claimed to belong to writer r. The SDC-based dissimilarity between two normalised feature sets was defined in Sections 4.2.1 and 4.2.2. This dissimilarity can be based on the Euclidean distance or a DTW-based distance and is now denoted by $D_{\mathrm{SDC}}(\mathbf{X}_r^{\mathrm{Test}}, \mathbf{X}_r^{\mathrm{Ref}})$ for brevity. The dissimilarity between a questioned sample and a claimed writer may be represented by the *average* SDC-based dissimilarity between the normalised questioned feature set and the feature sets associated with the K reference samples. However, when any questioned feature set is compared with a reference feature set, the resulting dissimilarity value is *normalised* as is explained in the next section.

5.2.2 Normalisation

As was eluded to earlier, once $D_{\mathrm{SDC}}(\mathbf{X}_r^{\mathrm{Test}}, \mathbf{X}_r^{\mathrm{Ref}})$ is obtained, it is subjected to writer-specific dissimilarity normalisation. This is done by calculating the writer-specific statistic μ_r, that is the average SDC-based disimilarity between each of the *reference* samples associated with writer r, as follows,

$$\mu_r = \frac{1}{\frac{K(K-1)}{2}} \sum_{i=2}^{K} \sum_{j=1}^{i-1} D_{\mathrm{SDC}}\left(\left(\mathbf{X}_r^{\mathrm{Ref}}\right)_{(i)}, \left(\mathbf{X}_r^{\mathrm{Ref}}\right)_{(j)} \right). \tag{5.1}$$

This results in the following normalisation function,

$$D_{\mathrm{Norm}}(\mathbf{X}_r^{\mathrm{Test}}, \mathbf{X}_r^{\mathrm{Ref}}) = \frac{D_{\mathrm{SDC}}(\mathbf{X}_r^{\mathrm{Test}}, \mathbf{X}_r^{\mathrm{Ref}}) - \mu_r}{\mu_r}. \tag{5.2}$$

This process is followed for each writer. The writer specific statistic μ_r is subsequently used to normalise any dissimilarity value associated with writer r.

5.2.3 Verification

Each questioned feature set (positive or negative) claimed to belong to writer r is compared to the K reference feature sets associated with writer r, after which each of the K dissimilarity values are normalised according to Equation 5.2 and averaged.

Following this procedure, a set of predetermined thresholds $\tau_i \in [0, \infty)$ are considered for the purpose of determining an optimal global threshold τ so that a questioned sample $\mathbf{X}_r^{\mathrm{Test}}$ is deemed to be authentic if and only if the *average* normalised distance between $\mathbf{X}_r^{\mathrm{Test}}$ and the reference samples of writer r is less than or equal to said threshold τ. The optimal global threshold τ is obtained by computing the false acceptance rate (FAR), that is the fraction of negative samples falsely accepted and the false rejection rate (FRR), that is the fraction of positive samples falsely rejected, for each sliding threshold τ_i by considering the entire test set (for all writers). The threshold associated with the *equal error rate* (EER), that is where FAR = FRR, is deemed optimal.

Following the successful computation of the optimal threshold, this threshold may now be used to authenticate *any* questioned sample. When a questioned sample is presented, the verification process is therefore as follows:

1. First, K reference samples of the word "the" is extracted from a document that is known to belong to the claimed writer and converted to appropriate feature sets, after which each feature set is normalised according to the protocol introduced in Section 3.3.

2. The normalised feature set associated with the questioned sample is compared to each of the K normalised reference feature sets of the claimed writer by employing either an Euclidean distance-based measure or a DTW distance-based measure. The writer specific statistic μ is subsequently obtained from the claimed writer's reference samples (Equation 5.1), after which the dissimilarity values are normalised (Equation 5.2).

3. The average of the normalised dissimilarity values obtained by comparing the questioned sample to all the reference samples of the claimed writer is computed.

4. If this average is less than or equal to the optimal EER-based global threshold value τ, the claim that the questioned sample belongs to the claimed writer is accepted, and is otherwise rejected.

5.3 Machine learning-based classifiers

In this section a dissimilarity vector-based representation in dissimilarity space is proposed for the purpose of implementing a writer-independent approach to authentication. In addition to this, a writer-specific dissimilarity vector normalisation protocol is proposed. Within the context of machine learning, an SVM-based and a QDA-based authentication protocol are proposed.

5.3.1 Dissimilarity representation

The previously described feature sets $\mathbf{X}^{\text{Test}}$ and $\mathbf{X}^{\text{Ref}}$ are suitable for a *writer-dependent* strategy towards authentication. Such a strategy is however limited in the sense that only samples claimed to belong to writers that form part of the training set can be presented for authentication. Within the context of the writer-dependent approach, *separate* models are developed for each individual writer for the purpose of authenticating a questioned sample. This implies that, whenever a new writer is to be enrolled into the system, a new model needs to be constructed for said writer. In addition to this, a writer-dependent strategy requires a large number of training samples for *each* individual writer in order to construct the writer-specific model in question. In order to overcome the aforementioned obstacles, the use of a *writer-independent* framework is adopted in this study. A *single* writer-independent model is trained for *all* writers and is able to assign a questioned sample to either a generic positive class or a generic negative class. Whenever a new writer is enrolled into a writer-independent system, only a set of K reference samples is required for this writer. Within the context of the writer-independent approach, the traditional feature vector-based representation in feature space is not appropriate and a dissimilarity vector-based approach in dissimilarity space is adopted as proposed by Pekalska *et al.* (2001)

It is therefore necessary to convert the entire feature sets belonging to all the writers from feature space into a a dissimilarity vector-based representation in dissimilarity space. Whenever a reference feature set $\mathbf{X}^{\text{Ref}}$ for a specific writer, and a questioned feature set $\mathbf{X}^{\text{Test}}$ claimed to belong to this writer, are considered, the aforementioned feature sets are converted into a writer-independent dissimilarity vector $\mathbf{z}^{\text{Ref,Test}}$. This conversion can be achieved by either using an Euclidean distance-based or a DTW distance-based measure for the purpose of computing the dissimilarity between the feature sets in question. Recall that the DTW-based distance is often more proficient than the Euclidean distance within the context of pattern recognition as was explained in Section 4.2.3. The DTW-based distance is therefore used for dissimilarity vector generation in this study, as is now explained. Using the DTW algorithm, the dissimilarity vector $\mathbf{z}^{\text{Ref,Test}}$ is calculated as follows,

$$\mathbf{z}^{\text{Ref,Test}} = \bigcup_{t=1}^{T} D_{\text{DTW}}(\mathbf{x}_t^{\text{Test}}, \mathbf{x}_t^{\text{Ref}}), \tag{5.3}$$

where $\mathbf{x}_t$ represents a d-dimensional feature vector (column) within the feature set (matrix) depicted in Equation 3.5. This results in a single T-dimensional dissimilarity vector $\mathbf{z}^{\text{Ref,Test}}$.

For the purpose of machine learning-based model construction, training data is required. It is assumed that a set of R so-called *guinea pig writers* are available and that there are K reference samples for each of these writers. It is furthermore assumed that N additional training samples consisting of an equal number of positive and negative samples for each writer is also available. A set of dissimilarity vectors is generated for each writer by computing $\mathbf{z}^{(n,k)}$ for $n = \{1, 2, \ldots, N\}$ and $k = \{1, 2, \ldots, K\}$, which shall now be denoted by the simplified notation $\mathbf{z}^{(i)}$. Let $\mathbf{Z}_{\text{pos}} = \{\mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \ldots, \mathbf{z}^{(M+)}\}$ denote the set that contains $M^+ = \frac{NKR}{2}$ positive dissimilarity vectors for all guinea pig writers and, similarly, let $\mathbf{Z}_{\text{neg}}$ denote the set of negative dissimilarity vectors obtained for all guinea pig writers.

5.3.2 Normalisation

At this stage, there still exists a large degree of overlap between the positive and negative classes in dissimilarity space. In order to maximally sperate $\mathbf{Z}_{\text{pos}}$ and $\mathbf{Z}_{\text{neg}}$, an appropriate normalisation strategy is required. A wide variety of suitable normalisation functions exist as documented in Jain *et al.* (2005) and Snelick *et al.* (2005). A well-known normalisation strategy involves the utilisation of sigmoidal normalisation functions, such as the logistic function. This study makes use of the modified logistic function conceptualised in Figure 5.1 and defined as

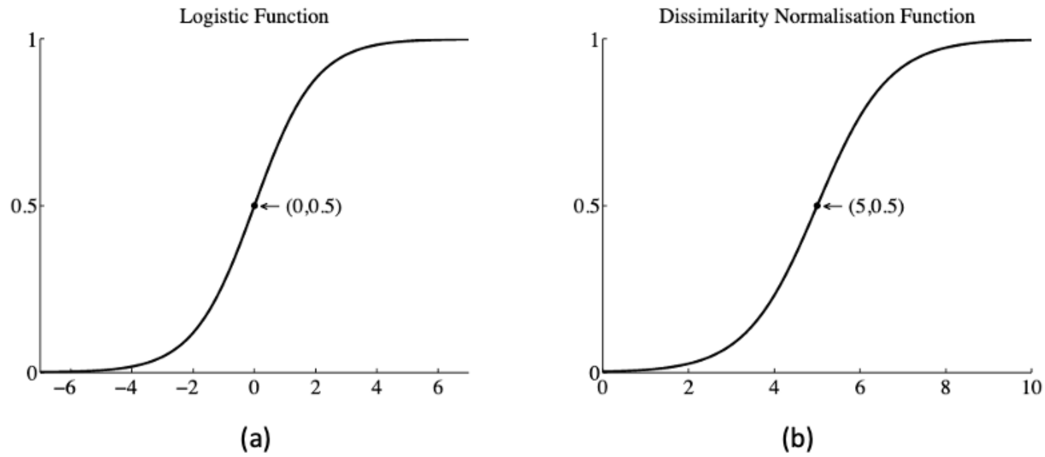


Figure 5.1: Conceptualisation of (a) the conventional logistic function and (b) the normalisation function utilised in this study (Swanepoel, 2015).

$$\eta(z,c) = \left[1 + \exp\left(c - \frac{6z}{c}\right)\right]^{-1}, \tag{5.4}$$

where $c \in \mathbb{R}$ represents the predefined scaling factor and $z \in [0, \infty)$. This normalisation function may subsequently be used to convert any T-dimensional dissimilarity vector $\mathbf{z} = \{z_1, z_2, \ldots, z_T\}$ into a normalised dissimilarity vector $\bar{\mathbf{z}}$ in the following way

$$\bar{\mathbf{z}} = \bigcup_{t=1}^{T} \eta(z_t, \mu_t + \sigma_t), \tag{5.5}$$

$$\mu_t = \frac{1}{M^+} \sum_{i=1}^{M^+} z_t^{(i)}, \tag{5.6}$$

$$\sigma_t = \sqrt{\frac{1}{M^+ - 1} \sum_{i=1}^{M^+} \left(z_t^{(i)} - \mu_t\right)^2}. \tag{5.7}$$

This strategy rescales any dissimilarity value to the interval [0, 1]. However the strategy is monotonic and unsuccessful to address the overlap between $\mathbf{Z}_{\text{pos}}$ and $\mathbf{Z}_{\text{neg}}$ as shown in Figure 5.2.

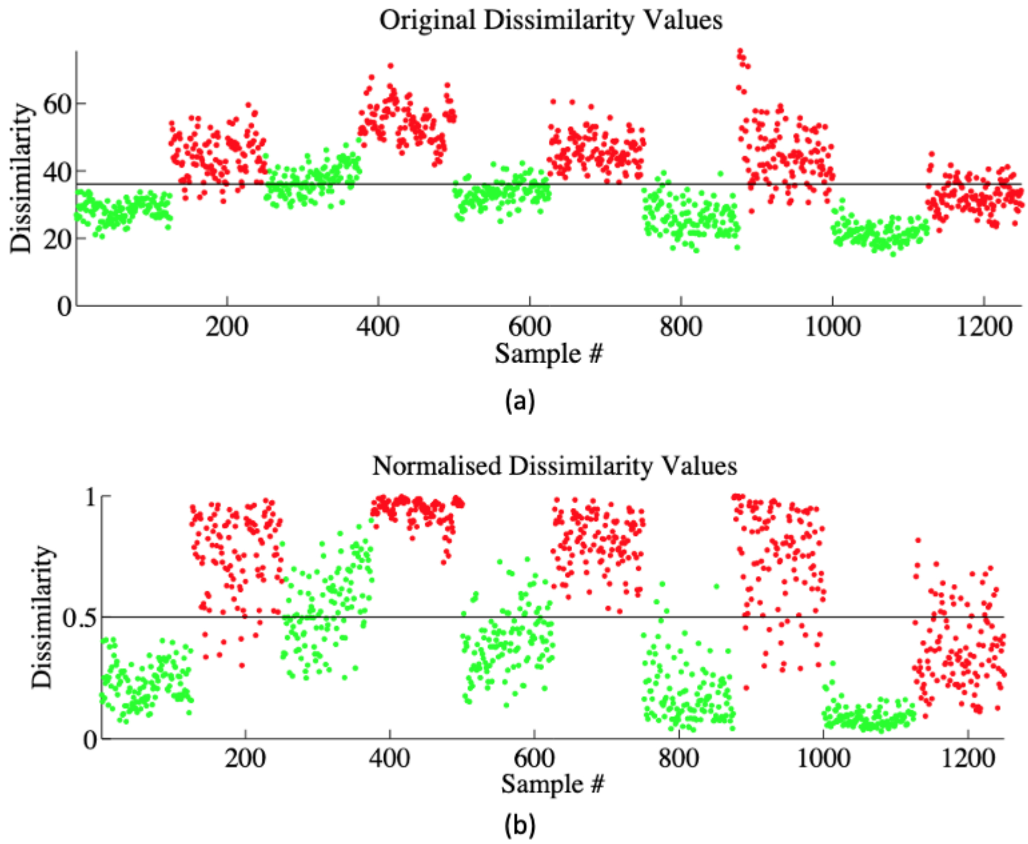


Figure 5.2: A visual representation of the application of the *global* dissimilarity normalisation strategy within the context of offline signature verification (Swanepoel, 2015). (a) Dissimilarity values of both negative (red) and positive (green) samples obtained from 5 writers. (b) Normalised dissimilarity values obtained using a *global* normalisation strategy $\eta(z, \mu + \sigma)$

In order to address this, the introduction of writer-specific dissimilarity statistics is proposed. For each writer r, the statistics $\mu^{(r)}$ and $\sigma^{(r)}$ are estimated. Any sample belonging to writer r (or claimed to belong to writer r) is subsequently normalised using the writer-specific normalisation function

$$\bar{\mathbf{z}} = \bigcup_{t=1}^{T} \eta\left(z_t, \mu_t^{(r)} + \sigma_t^{(r)}\right), \tag{5.8}$$

$$\mu_t^{(r)} = \frac{1}{M^{(r)}} \sum_{\substack{k=1 \\ j>k}}^{K} z_t^{(k,j)}, \tag{5.9}$$

$$\sigma_t^{(r)} = \sqrt{\frac{1}{M^{(r)}-1} \sum_{\substack{k=1 \\ j>k}}^{K} \left(z_t^{(k,j)} - \mu_t^{(r)}\right)^2}, \tag{5.10}$$

where $M^{(r)} = \frac{K^2-K}{2}$, that is the number of unique dissimilarity vectors generated when every reference sample belonging to writer r is compared to every *other* reference sample belonging to the same writer. A conceptualisation of this writer-specific normalisation strategy is shown in Figure 5.3.

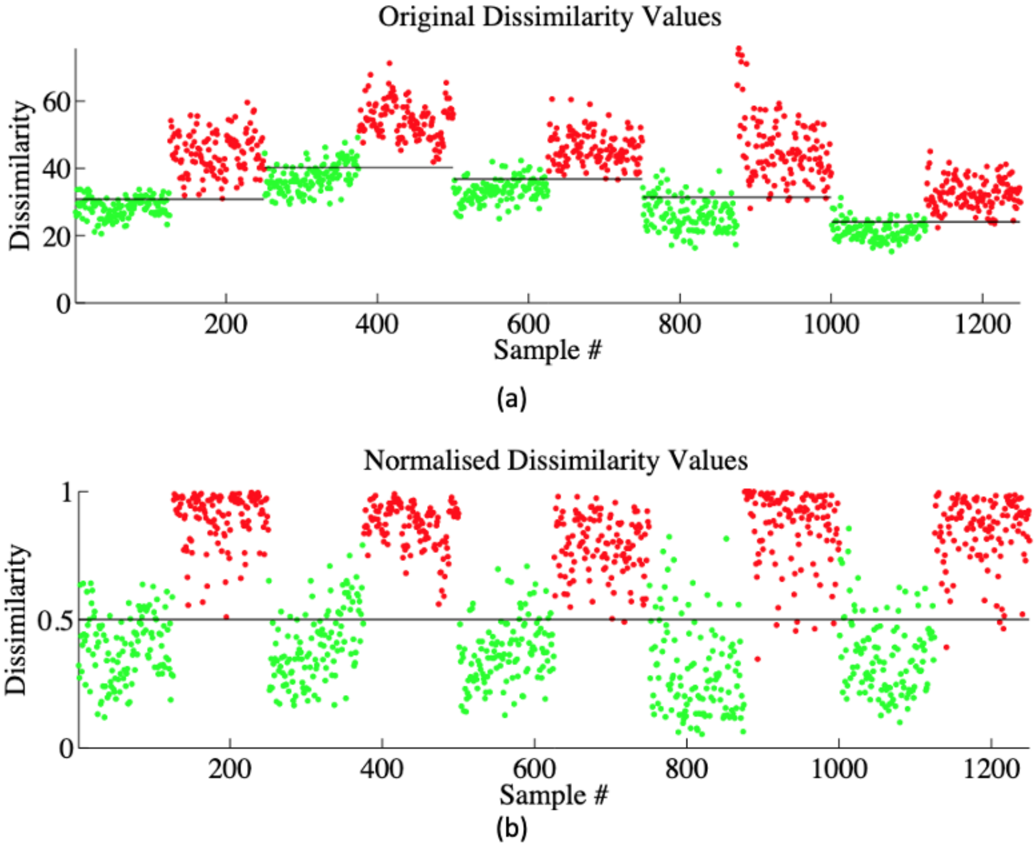


Figure 5.3: A visual representation of the application of the *writer-specific* dissimilarity normalisation strategy within the context of offline signature verification (Swanepoel, 2015). (a) Dissimilarity values of both negative (red) and positive (green) samples obtained from 5 writers. (b) Normalised dissimilarity values obtained using the proposed *writer-specific* normalisation strategy $\eta(z, \mu^{(r)} + \sigma^{(r)})$.

A comparison between the global dissimilarity normalisation strategy and the writer-specific dissimilarity normalisation strategy within the context of offline signature verification (Swanepoel (2015)) is drawn in Figure 5.4. It is clear that better separation in dissimilarity space between the positive and negative classes is achievable when a writer-specific dissimilarity normalisation strategy is employed.

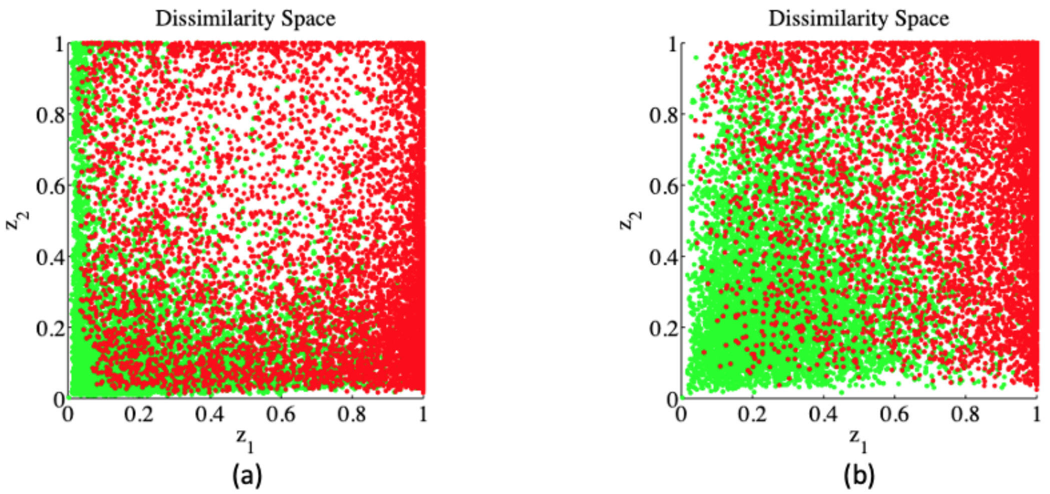


Figure 5.4: A visual comparison of the (a) global dissimilarity normalisation strategy and (b) the writer specific normalisation strategy being applied within the context of offline signature verification (Swanepoel, 2015).

5.3.3 Model training

Upon completion of the dissimilarity vector normalisation protocol, $\bar{\mathbf{Z}}_{\text{pos}}$ and $\bar{\mathbf{Z}}_{\text{neg}}$ can be used as training data for the proposed writer-independent authentication model. Two separate models are developed using an SVM and a QDA-based classifier respectively for the purpose of obtaining the optimal decision boundary.

SVM

As previously stated, this study makes use of radial basis function kernel support vector machines, which shall henceforth be denoted by RBK-SVMs. Within the context of these SVMs, a number of parameters have to be trained and hyperparameters have to be specified. As discussed in the previous chapter, the weight vector $\mathbf{w}$ and bias b are estimated from the training data. However the regularisation parameter C and the kernel width σ need to be specified prior to training. Selecting appropriate values of C and σ is imperative since this significantly influences the decision boundary and, ultimately, the accuracy of the RBK-SVMs. To obtain the optimal values of σ and C, all the normalised training samples are compared to every other normalised training sample by means of the Euclidean distance. The parameter σ is then assigned to the *median* value while C is assigned the *maximum* value of the computed distances. Swanepoel (2015) demonstrates that this method is sufficiently accurate in determining the values of C and σ. This yields the following model for the RBK-SVM,

$$\lambda_{\text{SVM}} = \{\mathbf{w}, b, C, \sigma\}.$$

QDA

The quadratic discriminant analysis (QDA) class membership function is described by the mean vectors and covariance matrices that represent the positive and negative classes. These are obtained using equations 4.8 and 4.9 respectively. This facilitates the construction of the positive and negative mean vectors $\boldsymbol{\mu}^+$ and $\boldsymbol{\mu}^-$, as well as the class-specific covariance matrices, $\mathbf{C}^+$ and $\mathbf{C}^-$, for $\bar{\mathbf{Z}}_{\text{pos}}$ and $\bar{\mathbf{Z}}_{\text{neg}}$ respectively. This yields the following model for the QDA-based classifier,

$$\lambda_{\text{QDA}} = \{\boldsymbol{\mu}^+, \boldsymbol{\mu}^-, \mathbf{C}^+, \mathbf{C}^-\}.$$

5.3.4 Verification

Following the successful training of the RBK-SVM-based and QDA-based models, a questioned sample can now be authenticated. Given a questioned handwritten sample and a claim that the sample belongs to a specific writer r, the system will attempt to verify this claim in the following manner:

1. Similar to the protocol for SDC-based verification, K reference samples of the word "the" are extracted from a document that is known to belong to the claimed writer and converted to appropriate feature sets, after which each feature set is normalised.

2. The normalised feature set associated with the questioned sample is compared to each of the K normalised reference features sets associated with the claimed writer r using the proposed DTW-based dichotomy transformation. This results in a set of dissimilarity vectors in dissimilarity space. The writer-specific statistics $\mu^{(r)}$ and $\sigma^{(r)}$ are subsequently obtained from the claimed writer's reference samples and used to normalise the generated dissimilarity vectors.

3. The normalised dissimilarity vectors are then individually presented to the trained model, each resulting in a distance measure L. This distance measure constitutes the dissimilarity vector's distance from the decision boundary for either the QDA-based or RBK-SVM-based membership function.

4. Said distance measures are then converted into partial confidence scores s_i for each of the K distance measures using the function

$$s_i = [1 + \exp(-L_i)]^{-1}.$$

This set of confidence scores is then averaged to produce a *final* confidence score s as follows

$$s = \frac{1}{K}\sum_{i=1}^{K} s_i.$$

5. A global threshold τ is imposed such that the questioned sample is accepted as belonging to the claimed writer if and only if $s > \tau$.

5.4 Concluding remarks

This chapter discussed SDC-based modelling approaches, as well as a writer-specific normalisation strategy, for the purpose of developing a verification system that is able to authenticate a questioned sample by employing a global threshold. Within the context of machine learning-based approaches, the use of a DTW-based dichotomy transformation for the purpose of obtaining dissimilarity vector-based representations in dissimilarity space for the purpose of constructing a writer-independent model were discussed. A writer-specific normalisation strategy allowing for greater separability between classes, after which RBK-SVM based and QDA-based writer authentication models were trained, was discussed. Finally, the verification protocols of the aforementioned systems were discussed.

In the next chapter the proficiency of the proposed systems within the context of offline writer authentication is experimentally investigated and analysed.

Chapter 6

Experiments

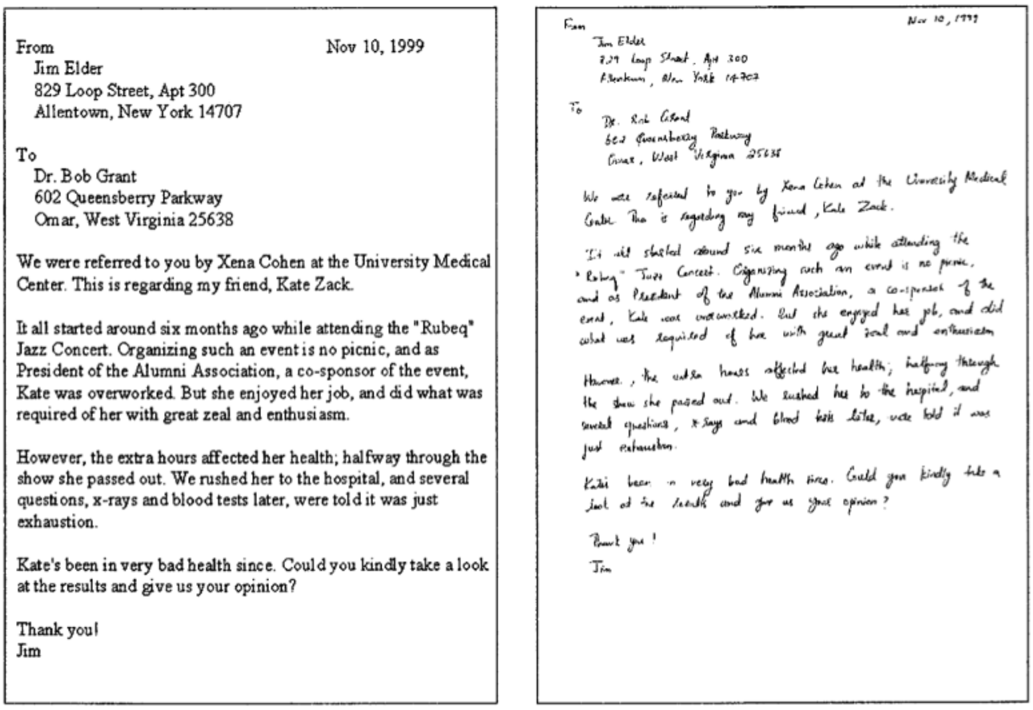
6.1 Introduction

In this chapter a number of rigorous experiments are conducted for the purpose of estimating the proficiency of the proposed handwriting authentication models presented in Chapters 3 to 5.

The data set considered for experimental purposes is described in Section 6.2. Section 6.3 introduces the statistical performance measures that are considered for the purpose of estimating the proficiency of each of the classifiers proposed in this study. Section 6.4 outlines the experimental protocol for each of the individual experiments conducted in this study. The experimental results, as well as a detailed analysis of these results, are presented in Section 6.5. The results achieved in this study are finally placed into context in Section 6.6 by comparing the proficiency of the classifiers proposed in this study to those of existing systems reported in the literature.

6.2 Data

The data utilised in this study was extracted from the CEDAR-LETTER data set that contains handwritten samples of 1500 writers. This data set was acquired as part of a study by Srihari *et al.* (2002). Each writer was given the same typed master document that had to be reproduced in handwriting on three separate occasions which resulted in three sample documents associated with each writer. In Figure 6.1 the typed master document, as well as a handwritten document authored by one of the writers, is depicted. This study employs a *subset* of the CEDAR-LETTER database and consists of handwriting authored by 100 different writers. The writers contained in this subset were selected in the order of their appearance in the CEDAR-LETTER database. For quality control reasons, certain writers were not selected since the handwritten sample documents in question were deemed not to be of appropriate standard. Some of these handwritten sample documents were excluded due to the fact that the writer erroneously omitted a word or the fact that a number of words were not adequately separated. Only 100 writers were selected due to the fact that the *manual* selection of each of the eight occurrences of the word "the" in each of the 300 handwritten samples is extremely labour intensive.



From Nov 10, 1999
Jim Elder
829 Loop Street, Apt 300
Allentown, New York 14707

To
Dr. Bob Grant
602 Queensberry Parkway
Omar, West Virginia 25638

We were referred to you by Xena Cohen at the University Medical Center. This is regarding my friend, Kate Zack.

It all started around six months ago while attending the "Rubeq" Jazz Concert. Organizing such an event is no picnic, and as President of the Alumni Association, a co-sponsor of the event, Kate was overworked. But she enjoyed her job, and did what was required of her with great zeal and enthusiasm.

However, the extra hours affected her health; halfway through the show she passed out. We rushed her to the hospital, and several questions, x-rays and blood tests later, were told it was just exhaustion.

Kate's been in very bad health since. Could you kindly take a look at the results and give us your opinion?

Thank you!
Jim

(a) (b)

Figure 6.1: (a) The typed master document considered by each writer. (b) A digitally scanned handwritten sample authored by one of the writers (Srihari *et al.*, 2002).

Srihari *et al.* (2002) selected writers from a sample population deemed to be representative of the U.S. population. The master document was typed in English and comprises of 156 words. The document is considered to be "complete" in the sense that it contains all the characters (letters and numerals) of the English language, as well as certain character combinations. Since this study employs a *text-dependent* approach to handwriting authentication, the experiments focus on the selected word "the" contained in each document. The specified word "the" occurs eight times in each of the handwritten sample documents. Figure 6.2 shows samples of the extracted word "the" associated with a number of *different* writers.

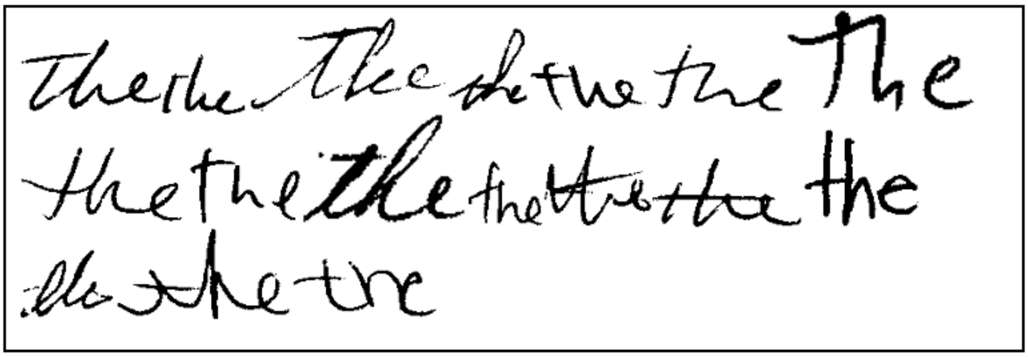


Figure 6.2: Sample images of the extracted word "the" authored by a number of *different* writers.

6.3 Statistical performance measures

This study employs a binary classifier for the purpose of distinguishing between a valid and fraudulent claim that a questioned handwritten sample was authored by a specific writer. All the data instances are therefore predicted to be either positive or negative. In order to estimate the proficiency of the systems proposed in this study, a number of statistical performance measures are considered, namely the false acceptance rate (FAR), the false rejection rate (FRR), and the equal error rate (EER). The aforementioned performance measures are defined in Table 6.1.

Performance measure	**Definition**
False acceptance rate (FAR)	$\frac{\text{FP}}{\text{FP+TN}}$
False rejection rate (FRR)	$\frac{\text{FN}}{\text{FN+TP}}$
Equal error rate (EER)	$\frac{\text{FP+FN}}{\text{TP+TN+FN+FP}}$ (FAR $\approx$ FRR)

Table 6.1: The statistical performance measures employed in this study.

The performance measures depicted in Table 6.1 are based on the following definitions:

- TP - The number of true positives, that is the number of positive instances correctly classified.
- FP - The number of false positives, that is the number of negative instances erroneously classified.

- TN - The number of true negatives, that is the number of negative instances correctly classified.

- FN - The number of false negatives, that is the number of positive instances erroneously classified.

Within the context of the machine learning-based classification protocols proposed in this study, an additional performance measure, the area under the curve (AUC) is employed for the purpose of further quantifying system proficiency. The AUC constitutes a single measure of the receiver operating characteristics (ROC) curve that quantifies the diagnostic proficiency of a binary classification system as the discrimination threshold is varied. The ROC curve plots the true positive rate (1-FRR) as a function of the false positive rate (FAR) for different threshold values. The AUC, that constitutes the area under the ROC curve, ranges from 0 to 1, where a value of 1 indicates a perfect predictor system. The ROC curve is a popular measure for evaluating classifier performance.

6.4 Experimental protocol

The aim of the experiments conducted in this study is to determine the proficiency of the proposed handwriting authentication systems within the context of authenticating a questioned handwritten sample. The experiments are categorised according to the classification technique employed by each of the proposed systems. Experiment 1 evaluates the performance of the SDC-based (Euclidean distance and DTW distance) classification techniques proposed in this study. Experiment 2 evaluates the proficiency of the machine learning-based (SVM and QDA) classification techniques proposed in this study, which are based on a dissimilarity vector representation.

6.4.1 Experiment 1

This experiment is further subdivided according to the employed template matching technique, that is a classifier based on the Euclidean distance (Experiment 1A) and a classifier based on a DTW distance (Experiment 1B):

- **Experiment 1A :** Euclidean distance-based classifier - This experiment investigates the proficiency of the proposed Euclidean distance-based classification technique for the purpose of authenticating a questioned handwritten sample.

- **Experiment 1B :** DTW distance-based classifier - This experiment investigates the proficiency of the proposed DTW distance-based classification technique for the purpose of authenticating a questioned handwritten sample.

Experiment 1 is conducted on a dataset with $\Psi = 100$ writers. For a specific writer r, *one* sample document, that contains $K = 8$ instances of the word "the" known to belong to the aforementioned writer, is used for the purpose of creating a reference set. The remaining *two* sample documents associated with the aforementioned writer contains $N^{(+)} = 2 \times 8 = 16$ instances of the word "the" and is used for the purpose of creating a positive test set. A total of $N^{(-)} = 99 \times 16 = 1584$ non-reference instances of the word "the" that have been authored by other writers, but are *claimed* to belong to the aforementioned writer are used for the purpose of constructing a negative test set. This is represented in Table 6.2.

Ψ	K	$N^{(+)}$	$N^{(-)}$
100	8	16	1584

Table 6.2: A representation of the data utilised in Experiment 1. The number of writers is denoted by Ψ, while the number of reference samples associated with a specific writer is denoted by K. The number of positive and negative test samples that are claimed to belong to a specific writer is denoted by $N^{(+)}$ and $N^{(-)}$ respectively.

Positive and negative dissimilarity values associated with each writer are required for the purpose of determining a global threshold that is (in turn) utilised to determine whether a questioned sample is authentic or not. For the purpose of constructing positive dissimilarity values, each of the 8 reference samples associated with a specific writer is compared to all positive non-reference samples claimed to belong to the aforementioned writer. This results in $M^{(+)} = 8\times(2\times8) = 128$ positive dissimilarity values. In order to obtain negative dissimilarity values, each of the 8 reference samples associated with the aforementioned writer is compared to each of the negative non-reference samples claimed to belong to the writer, resulting in $M^{(-)} = 8\times(2\times8\times99) = 12\,672$ negative dissimilarity values. This is represented in Table 6.3.

Ψ	$M^{(+)}$	$M^{(-)}$
100	128	12 672

Table 6.3: The number of dissimilarity values utilised for each writer in Experiment 1. The number of writers is denoted by Ψ. The number of positive dissimilarity values available for *each* writer is denoted by $M^{(+)}$, while $M^{(-)}$ denotes the number of negative dissimilarity values available for *each* writer.

6.4.2 Experiment 2

The objective of Experiment 2 is to ascertain the proficiency of the machine learning-based classifiers proposed in this study, that is SVMs (Experiment 2A) and QDA-based classifiers (Experiment 2B). Proficiency is determined by the systems ability to authenticate a questioned handwritten sample when it is represented by a dissimilarity vector in dissimilarity space. Experiment 2 is therefore subdivided as follows:

- **Experiment 2A :** SVM-based classifier - This experiment investigates the proficiency of the proposed SVM classification technique that makes use of a dissimilarity vector-based representation for the purpose of authenticating a questioned handwritten sample.

- **Experiment 2B :** QDA-based classifier – This experiment investigates the proficiency of the proposed QDA-based classification technique that makes use of a dissimilarity vector-based representation for the purpose of authenticating a questioned handwritten sample.

Experiment 2 is conducted on a data set containing $\Psi = 100$ writers, with Ψ_{T} guinea pig writers used for the purpose of training and Ψ_{E} other writers used for the purpose of evaluation. In order to ensure that the results reported in this study are both comprehensive and unbiased, k-fold cross validation, as well as n-fold data randomisation, is employed in the following manner.

- **k-fold cross validation**
  The data set containing Ψ writers is initially partitioned into k equal subsets. Each of these subsets is then in turn used as an evaluation set representative of $\frac{\Psi}{k}$ writers while the remaining $\frac{\Psi(k-1)}{k}$ writers constitute the training set. Each one of these individual evaluations is known as a run and the entire set of k runs constitutes a trial. This ensures that handwritten samples from different writers are used for the purpose of model training and evaluation so as to avoid the potential of overfitting. A total of $n = 100$ trials are conducted for each of the systems considered in this experiment.

- **n-fold data randomisation**
  The k-fold cross-validation process described above is repeated n times. In the case of each trial, the order of the writers is randomised prior to data partitioning. This procedure serves to nullify the influence of outliers, that is, writers whose inclusion into the evaluation set results in atypical high or low performance estimates. Each of the results reported in this experiment involve $n = 100$ trials.

In order to obtain positive dissimilarity vectors, each of the K reference samples associated with each writer is compared to all $N^{(+)}$ positive non-reference samples claimed to belong to that writer. In order to obtain negative dissimilarity vectors, each of the K reference samples associated with each writer is compared to all $N^{(-)}$ non-reference negative samples authored by *one* other writer. The refrence samples are compared to the negative samples authored by only *one* writer so as to ensure that there is an equal number of both positive and negative dissimilarity vectors for each writer. It is important to note that for the training set only Ψ_{T} guinea pig writers are considered. Similarly, for the evaluation set, only Ψ_{E} other writers (which do not constitute guinea pig writers) are considered.

The aforementioned data partitioning protocol is depicted in Table 6.4. This protocol constitutes a single run within a single trial and is conceptualised in Figure 6.3.

Ψ	Ψ_{T}	Ψ_{E}	$N^{(+)}$	$N^{(-)}$
100	67	33	128	128

Table 6.4: The data partitioning protocol utilised for a single run of a single trial in Experiment 2. The total number of writers is denoted by Ψ, while the number of writers in the training and evaluation sets are denoted by Ψ_{T} and Ψ_{E} respectively. The number of positive and negative test samples per writer is denoted by $N^{(+)}$ and $N^{(-)}$ respectively.

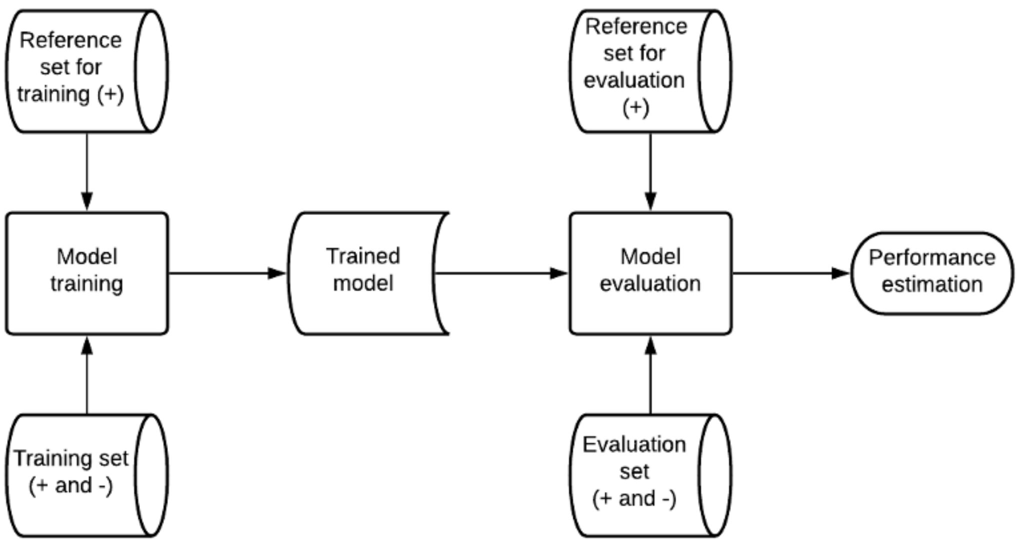


Figure 6.3: Conceptualisation of the protocol undertaken to estimate system performance for a single run of a single trial for Experiment 2.

6.5 Results

In this section, a comprehensive analysis of the results, pertaining to the experimental protocol outlined in the previous section, is presented. The proficiency of the proposed systems is gauged by estimating the EER, and in the case of the machine learning-based approaches, also the AUC.

6.5.1 Experiment 1

In order to determine the proficiency of the proposed DTW distance-based and Euclidean distance-based classification techniques with respect to authenticating a handwritten sample, the optimal global threshold, in each case, is obtained. Recall that the optimal threshold is the point where FAR = FRR. The FAR and FRR are computed for a number of equidistant predetermined threshold values using the test set that contains *all* the positive and negative dissimilarity values of Ψ writers. These two error rates are plotted for each threshold value which results in two curves. The point where the curves intersect (FAR = FRR) constitutes the EER and is deemed to coincide with the *optimal* threshold, where the system provides the *minimum* classification error with respect to incorrectly classifying a given sample.

The performance estimates obtained from Experiment 1 are presented in Figures 6.4 and 6.5. From the results, it is clear that the DTW distance-based system yields a lower EER as compared to the Euclidean distance-based system. It can therefore be concluded that the DTW-based model is a more proficient classification technique than the Euclidean distance-based model in authenticating a handwritten sample within the context of this study. This result was expected since DTW initially optimally aligns the feature vectors before computing the dissimilarity value.

6.5.2 Experiment 2

The experimental protocol previously discussed involves $k = 3$ runs per trial and a total of $n = 100$ trials. This results in performance estimates obtained from $k \times n = 300$ evaluations.

Recall that the discrete Radon transform (DRT) is utilized to extract features from each sample text as discussed in Section 3.3. During this process, a predetermined projection profile length (feature vector dimension) d and number of projection angles (number of feature vectors) T are specified. These values constitute hyperparameters for this experiment. In addition to this, the relationship between the reference sample size K and system proficiency is investigated. The system performance for each possible combination of the said parameters is estimated. The value of d was predetermined through prior experimentation, where the optimal value was determined to be $d = 40$, which is consistent with the average dimension (in pixels) of a subimage containing

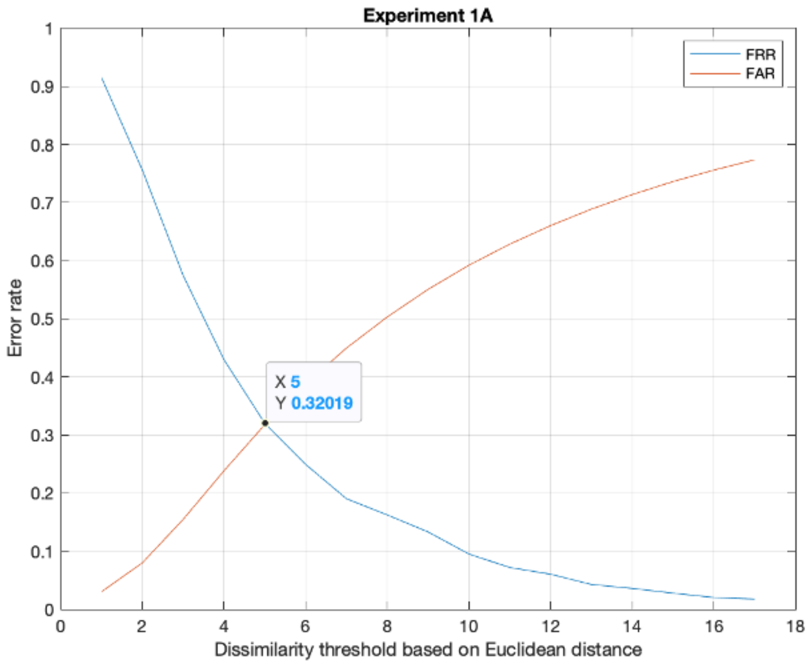


Figure 6.4: Experiment 1A. The *average* FAR (red) and FRR (blue) obtained for $\Psi = 100$ writers using the Euclidean distance-based classification technique as a dissimilarity measure. An EER of 32.01% is achieved where FAR = FRR.

the targeted word "the". This value is *constant* for all the experiments. Table 6.5 summarises the hyperparameters that will be used in the experimental procedure.

Parameter	Value(s)
d	40
T	{2, 4, 6, 16, 32, 64, 128, 256, 512, 1024}
K	{8, 12, 16}

Table 6.5: The hyperparameters considered for system evaluation in Experiment 2.

The value of T is specified to be 2^n where $n = 1, 2, \ldots, 10$ so as to ensure that each higher resolution projection angle (associated with a larger value of n) includes at least all the angles in any of the previous lower resolutions (associated with a lower value of n).

From the experimental results obtained for the proposed SVM-based and QDA-based systems as depicted in Tables 6.6 and 6.7, and Tables 6.8 and 6.9, respectively, it is clear that the SVM-based system outperforms the QDA-based system within the context of this study. The SVM-based model achieves its best result when $T = 1024$ and $K = 16$, that is an AUC of 93% and an EER of 14.93%. When similar hyperparameters are employed, the QDA-based model

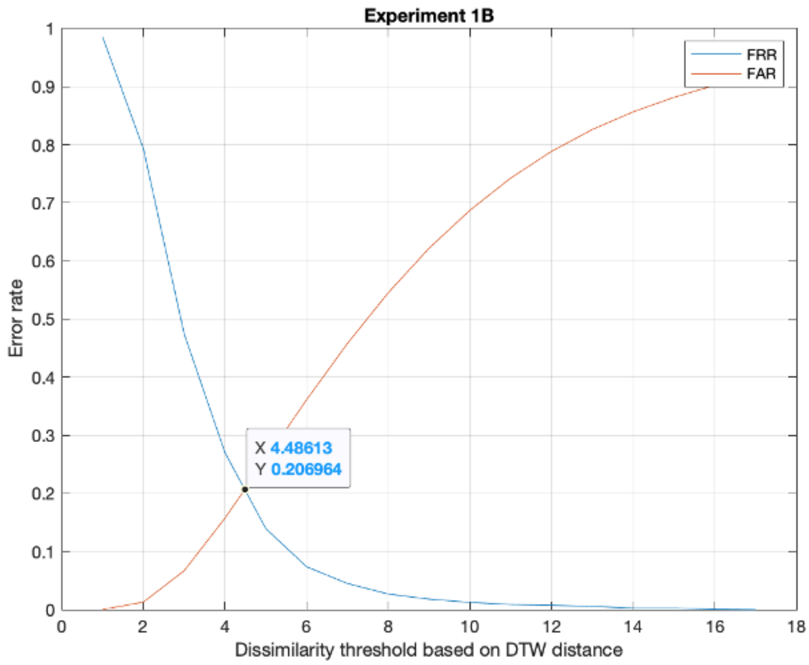


Figure 6.5: Experiment 1B. The *average* FAR (red) and FRR (blue) obtained for $\Psi = 100$ writers using DTW distance-based classification technique as a dissimilarity measure. An EER of 20.7% is achieved where FAR = FRR.

achieves an AUC of 76% and an EER of 31.24%. In the following sections the aforementioned results are discussed in more detail within the context of the proposed SVM-based system (Experiment 2A) and the proposed QDA-based system (Experiment 2B).

Experiment 2A

The mean estimates for the proposed SVM-based classifier based on the AUC and EER performance metrics are presented in Tables 6.6 and 6.7 respectively. Several observations into the effect of the K and T hyperparameters are made. For example, an increase in the reference sample size K does not yield a significant increase in system proficiency. It can therefore be concluded that, the ideal reference sample size with respect to the proposed SVM is $K = 8$ samples. With regard to the increase in the number of projection angles (number of feature vectors) T, the correlation with system proficiency is more clear. It is observed that as T increases, the system proficiency also increases. This behaviour is observed in almost all the cases irrespective of the reference sample size K. It can therefore be concluded that the ideal value for T is 1024.

AUC (%)		K			Average (T)
		8	12	16	
T	2	65	65	65	65.0
	4	67	66	67	66.7
	8	67	66	66	66.3
	16	69	70	68	69.0
	32	69	71	71	70.3
	64	71	73	73	72.3
	128	82	81	81	81.3
	256	89	88	88	88.3
	512	91	92	92	91.7
	1024	**92**	**92**	**93**	**92.3**
	Average (K)	76.2	76.4	76.4	

Table 6.6: The average AUC achieved by the proposed SVM-based system. The performance estimates associated with the optimal number of projection angles T are depicted in boldface.

EER (%)		K			Average (T)
		8	12	16	
T	2	34.42	35.19	34.57	34.73
	4	34.60	35.26	34.78	34.88
	8	34.76	34.73	34.55	34.68
	16	34.74	35.03	33.81	34.53
	32	35.43	34.77	33.89	34.70
	64	33.47	32.29	32.63	32.80
	128	24.36	26.65	25.79	25.60
	256	19.02	19.26	18.88	19.05
	512	16.28	16.54	16.16	16.33
	1024	**15.51**	**15.31**	**14.93**	**15.25**
	Average (K)	28.26	28.50	28.00	

Table 6.7: The average EER achieved by the proposed SVM-based system. The performance estimates associated with the optimal number of projection angles T are depicted in boldface.

Experiment 2B

The mean estimates for the proposed QDA-based classifier based on the AUC and EER performance metrics are presented in Tables 6.8 and 6.9 respectively. For the most part it is observed that as the reference sample size K increases, system proficiency also increases, although only marginally. With respect to the number of projection angles (number of feature vectors) T, it is observed that for the most part, the system performs better as T increases. It can therefore be concluded that within the context of the QDA-based model, the optimal values of K and T are 16 and 512 respectively.

AUC (%)		K 8	12	16	Average (T)
T	2	70	70	71	70.3
	4	69	70	70	69.7
	8	69	70	70	69.7
	16	71	69	69	69.7
	32	64	67	68	66.3
	64	65	67	68	66.7
	128	73	70	72	71.7
	256	76	73	74	74.3
	512	**75**	**76**	**76**	**75.7**
	1024	75	75	76	75.3
	Average (K)	70.7	70.7	71.4	

Table 6.8: The average AUC achieved by the proposed QDA-based system. The performance estimates associated with the optimal number of projection angles T are depicted in boldface.

EER (%)		K 8	12	16	Average (T)
T	2	39.26	39.37	38.75	39.13
	4	39.58	39.17	39.48	39.41
	8	39.21	39.66	39.02	39.30
	16	39.29	39.20	38.70	39.06
	32	39.30	39.56	38.09	38.98
	64	38.91	39.26	37.79	38.65
	128	33.34	35.97	33.53	34.28
	256	33.54	33.36	32.16	33.02
	512	32.23	31.57	31.51	31.77
	1024	**32.18**	**31.84**	**31.24**	**31.75**
	Average (K)	36.68	36.90	36.03	

Table 6.9: The average EER achieved by the proposed QDA-based system. The performance estimates associated with the optimal number of projection angles T are depicted in boldface.

6.6 Comparison with previous work

It was demonstrated in the previous section that the systems developed in this study are relatively proficient in authenticating handwritten samples. In order to put the performance of said systems into perspective, the reported results obtained are compared to those documented in the literature. As previously mentioned, no true benchmark exists for determining the proficiency of writer authentication systems. In addition to this, different data sets and experimental protocols are employed by different studies, making it challenging to directly compare the performance of different systems.

Systems that utilised the same data set as the one employed in this study, that is the CEDAR-LETTER data set, are discussed, as well as systems that employed similar classification techniques as those employed in this study (for example SVMs), albeit on different data sets. This is to give the reader some idea of how the systems developed in this study compare to existing systems. It is however important to note that a direct comparison is often not possible for the reasons mentioned above. Recall that the optimal results in this study are achieved by the proposed SVM, that is an EER of 14.93% and an AUC of 93%.

CEDAR-LETTER data set

In this section, existing systems that also employed the CEDER-LETTER data set are discussed.

Srihari *et al.* (2002) made use of an artificial neural network (ANN) for the purpose of *text-independent* handwriting authentication. The system employed a training set of 834 within-author distances and 836 between-author distances and was tested on two separate test sets. Each test set also contained 834 within-author distances and 836 between-author distances. Within the context of the respective two test sets, 83.1% (with an FRR of 14.5% and an FAR of 19.3%) and 82.7% (with an FRR of 14.4% and an FAR of 20.2%) of the data was classified correctly.

Cha and Srihari (2000) proposed a system that aims to determine the viability of the individuality of handwriting. Ten different features were extracted from the relevant documents and an ANN was trained by considering the "distance" between samples. The experiments were carried out on 1000 writers. The results (on a document level) are summarised in Table 6.10.

Number of features	5	9	10
FRR	5.3%	4.6%	3.5%
FAR	5.2%	3.5%	2.1%
Accuracy	95%	96%	97%

Table 6.10: Experimental results achieved by Cha and Srihari (2000).

SVM

Existing systems that also used an SVM-based classification protocol is now discussed.

Bertolini *et al.* (2016) proposed a writer-independent classifier for writer recognition that is based on a dissimilarity representation. The system was evaluated on the QUWI database, containing 4 068 handwritten samples from 1 017 different writers. The proposed *text-dependent* approach achieved an equal error rate (ERR) of 31.2% for English samples, while an EER of 25.7% was reported for Arabic samples.

Saranya and Vijaya (2013) proposed a *text-dependent* writer recognition model that employs an SVM. The system was tested on a data set containing 1000 data instances of 100 words written by 10 writers. The authors reported an accuracy of 94.3% when utilising a polynomial kernel, while accuracies of 77.6% and 92.3% were reported for linear and RBF kernels respectively.

6.7 Conclusion

This chapter set out by introducing the proposed experimental protocol for the experiments conducted in this study. Next, the experimental results for the individual experiments were presented. Based on the experimental findings, the following conclusions can be drawn:

SDC-based systems

1. The proposed DTW distance-based system proves to be superior to the proposed Euclidean distance-based system for the purpose of authenticating a questioned handwritten sample within the context of this study.

2. The proposed DTW distance-based system achieves an EER of 20.7%, while the proposed Euclidean distance-based system achieves an EER of 32.0%.

Machine learning-based systems

1. The proposed SVM-based system is more proficient than the proposed QDA-based system within the context of this study.

2. The most proficient SVM-based system achieves an AUC of 93% and an EER of 14.93%.

3. The most proficient QDA-based system achieves an AUC of 76% and an EER of 31.24%.

4. System proficiency increases as T, that is the number of projection angles (number of feature vectors), increases. This is true for both the SVM-based and QDA-based systems.

5. System proficiency does not notably increase as K, that is the size of the reference set, increases. This is true for both the SVM-based and QDA-based systems.

In the next chapter, recommendations for future work to improve upon this study are presented.

Chapter 7

Conclusion and future work

7.1 Conclusion

This thesis proposed a number of offline writer authentication systems. Two distinct types of systems were proposed in this study, those that make use of template matching techniques, that is those based on the Euclidean distance and those based on a dynamic time warping (DTW) distance, as well as those that make use of machine learning-based approaches, that is support vector machines (SVMs) and classifiers that employ quadratic discriminant analysis (QDA). Firstly, a raw text image of the word targeted in this study was segmented from the document in question. Following this, the image underwent various preprocessing techniques allowing for optimal feature extraction. Feature extraction was implemented by applying the discrete Radon transform (DRT) to the image in question. This resulted in a feature set, which was subsequently normalised. In the case of the proposed simple distance classifiers (SDCs), the dissimilarity of two samples was quantified by the average distance between the corresponding feature vectors, which may be based on either the Euclidean distance or a DTW distance. This dissimilarity value was subsequently normalised using writer-specific statistics.

In the case of the proposed machine learning-based systems, the normalised feature sets were converted into a dissimilarity vector-based representation by using a DTW-based dichotomy transformation for the purpose of putting into place a writer-independent framework. Fundamental techniques, which include a writer-specific normalisation strategy were subsequently employed. Following this, SVM-based and QDA-based authentication models were trained by considering the aforementioned dissimilarity vectors.

In order to quantify the proficiency of the aforementioned systems, rigorous testing was undertaken using a subset of the CEDDER-LETTER data set. The performance of the developed systems was compared to existing systems. The findings demonstrated that the performance of the proposed systems is comparable to existing systems. In addition to this, the findings showed that the proposed SVM-based system, that employs a radial basis function, proved to be the most proficient system developed in this study.

As a result, in terms of the objectives outlined in Section 1.3, this study is deemed successful

7.2 Future work

As a result of time constraints and the limited scope of this thesis, a number of additional research avenues were not pursued. A discussion is provided on areas of future research within the context of this study.

Alternative datasets

The research conducted in this study used a subset of the CEDDER-LETTER data set, that comprised of only 100 writers. As part of future research additional writers should be incorporated into the data set in order to obtain a more reliable estimate of system performance. It is also unclear how the proposed systems will perform on a different data set. As part of future research the performance of the proposed systems may therefore also be gauged by considering alternative data sets.

Text independent approach

As previously mentioned text-dependent systems may be impractical to implement for authentication purposes since the writer needs to provide very specific reference samples of the targeted word. As part of future work, the current system may be adapted in such a way that a text-independent approach is followed.

Improved image processing techniques

Several preprocessing techniques were discussed in Chapter 3. These techniques are by no means optimal for each of their respective tasks. Future research may use state-of-the-art techniques to address the various issues related to feature extraction. A key issue is that of noise reduction. Techniques such as the median filter could not be applied in this study, since the images proved too small (of too low resolution) and resulted in key features being accidentally removed. Researching other noise reduction techniques may prove useful. Another area of interest relating to image processing is to research the possibility of incorporating an automatic image grabber into the system. Said image grabber may automatically detect and "cut out" the selected word "the" from any document resulting in a usable image.

Testing on forgeries

As demonstrated in this thesis, the best system proposed in this study is sufficiently proficient. It was, however, only evaluated on samples in the handwriting of other authors. As part of future research, the incorporation of forgeries may be considered, that is deliberate attempts to imitate a known writer's handwriting.

List of References

Al-Maadeed, S., Hassaine, A., Bouridane, A. and Tahir, M.A. (2016). Novel geometric features for off-line writer identification. *Pattern Analysis and Applications*, vol. 19, no. 3, pp. 699–708.

Ballard, L., Lopresti, D. and Monrose, F. (2006). Evaluating the security of handwriting biometrics. In: *Tenth International Workshop on Frontiers in Handwriting Recognition*. Suvisoft.

Bensefia, A. and Paquet, T. (2016). Writer verification based on a single handwriting word samples. *EURASIP Journal on Image and Video Processing*, vol. 2016, no. 1, pp. 1–9.

Bertolini, D., Oliveira, L.S. and Sabourin, R. (2016). Multi-script writer identification using dissimilarity. In: *2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 3025–3030. IEEE.

Beylkin, G. (1987). Discrete radon transform. *IEEE transactions on acoustics, speech, and signal processing*, vol. 35, no. 2, pp. 162–172.

Cha, S.-H. and Srihari, S.N. (2000). Writer identification: statistical analysis and dichotomizer. In: *Joint IAPR international workshops on statistical techniques in pattern recognition (SPR) and structural and syntactic pattern recognition (SSPR)*, pp. 123–132. Springer.

Chen, L. (2019 Jan 7,). Support vector machine. CNN.
Available at: `https://towardsdatascience.com/support-vector-machine-simply-explained-fee28eba5496`

Christlein, V., Bernecker, D., Hönig, F., Maier, A. and Angelopoulou, E. (2017). Writer identification using gmm supervectors and exemplar-svms. *Pattern Recognition*, vol. 63, pp. 258–267.

Coetzer, J. (2005). *Off-line signature verification*. Ph.D. thesis, Stellenbosch: University of Stellenbosch.

Dhandra, B. and Vijayalaxmi, M. (2015). A novel approach to text dependent writer identification of kannada handwriting. *Procedia Computer Science*, vol. 49, pp. 33–41.

Fiel, S. and Sablatnig, R. (2012). Writer retrieval and writer identification using local features. In: *2012 10th IAPR International Workshop on Document Analysis Systems*, pp. 145–149. IEEE.

Fischer, A., Keller, A., Frinken, V. and Bunke, H. (2012). Lexicon-free handwritten word spotting using character hmms. *Pattern recognition letters*, vol. 33, no. 7, pp. 934–942.

Fornés, A., Lladós, J., Sánchez, G. and Bunke, H. (2008). Writer identification in old handwritten music scores. In: *2008 The Eighth IAPR International Workshop on Document Analysis Systems*, pp. 347–353. IEEE.

Ghiasi, G. and Safabakhsh, R. (2010). An efficient method for offline text independent writer identification. In: *2010 20th International Conference on Pattern Recognition*, pp. 1245–1248. IEEE.

Ghojogh, B. and Crowley, M. (2019 06). Linear and quadratic discriminant analysis: Tutorial.

Gracia, V. *et al.* (2006). State of the art in biometrics research and market survey. *HUMABIO Project (EU FP6 contract no 026990), Deliverable*, , no. 1.4.

Jain, A., Nandakumar, K. and Ross, A. (2005). Score normalization in multimodal biometric systems. *Pattern recognition*, vol. 38, no. 12, pp. 2270–2285.

Jain, A.K., Bolle, R. and Pankanti, S. (2006). *Biometrics: personal identification in networked society*, vol. 479. Springer Science & Business Media.

Kamal, P., Rahman, F. and Mustafiz, S. (2014). A robust authentication system handwritten documents using local features for writer identification. *Journal of Computing Science and Engineering*, vol. 8, no. 1, pp. 11–16.

Kumar, R. and Kaur, M. (2017). A character based handwritten identification using neural network and svm. *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*.

Liu, C. (2020 February,). A top machine learning algorithm explained: Support vector machines (svms). CNN.
Available at: `https://towardsdatascience.com/one-of-the-top-machine-learning-algorithms-for-supervised-learning-support-vector-machines-svms-fc45ac0667f4`

Manmatha, R., Han, C. and Riseman, E.M. (1996). Word spotting: A new approach to indexing handwriting. In: *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 631–637. IEEE.

Mukherjee, S. and Ghosh, I.D. (2020). Writer identification based on writing individuality and combination of features. In: *2020 IEEE Applied Signal Processing Conference (ASPCON)*, pp. 324–329.

Namboodiri, A. and Gupta, S. (2006). Text independent writer identification from online handwriting. In: *Tenth International Workshop on Frontiers in Handwriting Recognition*. Suvisoft.

Pandey, P. and Seeja, K. (2018). Forensic writer identification with projection profile representation of graphemes. In: *Proceedings of First International Conference on Smart System, Innovations and Computing*, pp. 129–136. Springer.

Pekalska, E., Paclik, P. and Duin, R.P. (2001). A generalized kernel approach to dissimilarity-based classification. *Journal of machine learning research*, vol. 2, no. Dec, pp. 175–211.

Rehman, A., Naz, S. and Razzak, M.I. (2019). Writer identification using machine learning approaches: a comprehensive review. *Multimedia Tools and Applications*, vol. 78, no. 8, pp. 10889–10931.

Rodríguez-Serrano, J.A. and Perronnin, F. (2009). Handwritten word-spotting using hidden markov models and universal vocabularies. *Pattern Recognition*, vol. 42, no. 9, pp. 2106–2116.

Saranya, K. and Vijaya, M. (2013). Text dependent writer identification using support vector machine. *International Journal of Computer Applications*, vol. 65, no. 2.

Snelick, R., Uludag, U., Mink, A., Indovina, M. and Jain, A. (2005). Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 3, pp. 450–455.

Srihari, S.N., Cha, S.-H., Arora, H. and Lee, S. (2002). Individuality of handwriting. *Journal of forensic science*, vol. 47, no. 4, pp. 1–17.

Swanepoel, J.P. (2015). *Writer-independent handwritten signature verification*. Ph.D. thesis, Stellenbosch: Stellenbosch University.

Tapiador, M., Gómez, J. and Sigüenza, J.A. (2004). Writer identification forensic system based on support vector machines with connected components. In: *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pp. 625–632. Springer.

Toft, P. (1996). The radon transform. *Theory and Implementation (Ph. D. Dissertation)(Copenhagen: Technical University of Denmark)*.

Tsiporkova, E. (2012). Dynamic time warping algorithm. Unpublished.

Vásquez, J.L., Ravelo-García, A.G., Alonso, J.B., Dutta, M.K. and Travieso, C.M. (2018). Writer identification approach by holistic graphometric features using off-line handwritten words. *Neural Computing and Applications*, pp. 1–14.

Walton, J. (1997). Handwriting changes due to aging and parkinson's syndrome. *Forensic science international*, vol. 88, no. 3, pp. 197–214.

Wang, D. (2019). Offline text-independent writer identification using different levels of features. In: *Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition*, pp. 337–341.

Yampolskiy, R.V. and Govindaraju, V. (2008). Behavioural biometrics: a survey and classification. *International Journal of Biometrics*, vol. 1, no. 1, pp. 81–113.