IDENTIFICATION OF SCRIBES FROM HISTORICAL MANUSCRIPTS

SIDRA NASIR

01-243191-012

A thesis submitted in fulfilment of the requirements for the award of degree of Masters of Science (Computer Science)

Department of Computer Science

BAHRIA UNIVERSITY ISLAMABAD

DECEMBER 2020

Approval of Examination

Scholar Name: Sidra Nasir

Registration Number: 62698

Enrollment: 01-243191-012

Program of Study: MS(CS)

Thesis Title: IDENTIFICATION OF SCRIBES FROM HISTORICAL MANUSCRIPTS

It is to certify that the above scholar's thesis has been completed to my satisfaction and, to my belief, its standard is appropriate for submission for examination. I have also conducted plagiarism test of this thesis using HEC prescribed software and found similarity index **17%**. that is within the permissible limit set by the HEC for the MS/M.Phil degree thesis. I have also found the thesis in a format recognized by the BU for the MS/M.Phil thesis.

Principal Supervisor Name: Dr. Imran Siddiqi

Principal Supervisor Signature:

Date: 5th April 2021

Author's Declaration

I, Sidra Nasir hereby state that my MS/M.Phil thesis titled is my own work and has not been submitted previously by me for taking any degree from Bahria university or anywhere else in the country/world. At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw/cancel my MS/M.Phil degree.

Name of Scholar: Sidra Nasir

Date: 10 December 2020

Plagiarism Undertaking

I, solemnly declare that research work presented in the thesis titled ________________ **"Identification Of Scribes From Historical Manuscripts"** is solely my research work with no significant contribution from any other person. Small contribution / help wherever taken has been duly acknowledged and that complete thesis has been written by me. I understand the zero tolerance policy of the HEC and Bahria University towards plagiarism. Therefore I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred / cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS/M.Phil degree, the university reserves the right to withdraw / revoke my MS/M.Phil degree and that HEC and the University has the right to publish my name on the HEC / University website on which names of scholars are placed who submitted plagiarized thesis.

Name of Scholar: Sidra Nasir

Date: 10 December 2020

Dedication

I dedicate my thesis work to my supervisor,my respected teachers, my parents, my all family members and my friends.

Acknowledgements

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main thesis supervisor, Professor Dr. Imran Siddiqi, for encouragement, guidance, critics and friendship. Without their continued support and interest, this thesis would not have been the same as presented here. I would also like to thank Dr. Isabelle Marthot-Santaniello from University of Basel, Switzerland for making the dataset available.

Librarians at Bahria University also deserve special thanks for their assistance in supplying the relevant literature. My fellow postgraduate students should also be recognised for their support. My sincere appreciation also extends to all my colleagues and others who have provided assistance at various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family members.

Abstract

Computerized analysis of historical documents has remained an interesting research area for the pattern classification community for many decades. From the perspective of computerized analysis, key challenges in the historical manuscripts include automatic transcription, dating, retrieval, classification of writing styles and identification of scribes etc. The focus of our current study lies on identification of writers from the digitized historical manuscripts. The documents are first pre-processed to segment handwriting from the background. For feature extraction and subsequent classification, we extract small patches of handwriting. These patches are extracted in two different ways, by a dense sampling of handwriting using small windows as well as by finding the key points in handwriting and using these key points as centers of small windows to extract writing fragments. Features are extracted from writing windows using a two-step fine-tuning of convolutional neural networks. First, the ConvNets are trained on contemporary handwriting samples and then fine-tuned to the limited set of historical manuscripts (Papyrus). Decisions on patches are combined using a majority vote to decide the authorship of a query document. Preliminary experiments on a set of challenging and degraded manuscripts report promising performance.

TABLE OF CONTENTS

LIST OF TABLE

LIST OF FIGURE

LIST OF SYMBOLS

T	–	Threshold
R	–	Dynamic Range
μ	–	Mean
σ	–	Standard Deviation
x_u	–	Ground Truth
e	–	Degradation
x	–	Input Image
n	–	Size of Window

CHAPTER 1

INTRODUCTION

Over the last few decades, there has been a significant increase in the trend to digitize ancient documents [2, 3]. The digitization is not only aimed at preserving the cultural heritage but also to make these documents publicly available without the need to physically access them. This contribution, in turn, led the researchers from pattern classification and the document and handwriting recognition community in particular to a new set of challenging problems [4]. Some of the prominent digitization projects include the International Dunhaung Project (IDP) [5], the Monk system [6], NAVIDOMASS (NAVIgation in Document MASSes) [4] and Madonne. Besides digitization, these projects are also supported by the development of automated tools to assist the paleographers in tasks like spotting keywords in manuscripts or retrieving documents with a particular writing style or a dropcap, etc. In the past, paleographers and historians have been hesitant in accepting computerized solutions. The key contributing factor to this resistance has been the lack of 'trust' in machine-based solutions. In recent years, thanks to the advances in various fields of image processing and machine learning, as well as the success of joint ventures between paleographers and computer scientists, experts seem to be more receptive to digital solutions in their practices [7]. The primary motivation behind such solutions is to facilitate the experts rather than replacing them. These tools can be exploited to narrow down the search space and experts can concentrate on a limited set of samples for detailed and in-depth analysis [8]. Among various challenges in computerized analysis of historical manuscripts, the identification of scribes carries significant importance. A writer of a document can be categorized by capturing the writing style which is known to be specific for each individual [9].

1.1 Motivation of Handwriting Analysis

Handwriting is an important form of communication in our culture that has developed and evolved over the years. We all learn to write according to a standard writing style at school, the 'copy book' that differs according to the temporal circumstances, geographical location and the historical and cultural background. Eventually, with the passage of time, we develop handwriting characteristics but our handwriting begins to deviate from the initially learned style. The writing differs depending on the circumstance, location, historical and cultural origin as presented in Figure 1.1. Thus, Handwriting

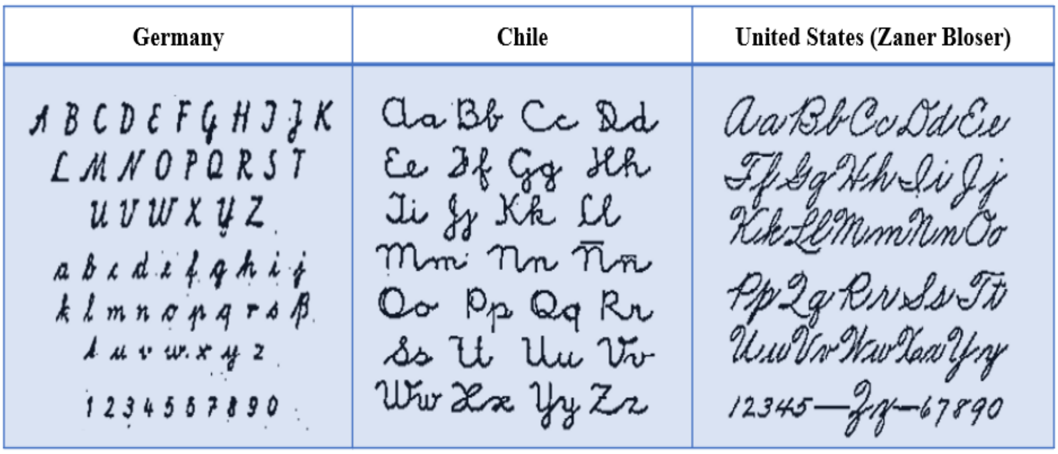


Figure 1.1: Various code book styles common in Germany, Chile and US.

is not an innate neural function but it is learned over the years, as it evolves it becomes an action that does not need much active attention [10]. These unique characteristics help in differentiating the writing of an individual from that of another, despite the fact that both texts share the same copybook style thus allowing the writer of a handwritten text to be identified [11].

From the view point of ancient manuscripts, such collections may present varied interests. They could, for instance, serve to study the form of writing and evolution of the style over the period of time which reflects the historical and cultural changes in the society. Familiarity with ligatures, abbreviations, individual letters, punctuation and how they have evolved has enabled the paleographers and historians in identifying the periods in which a manuscript is written (Figure 1.2).

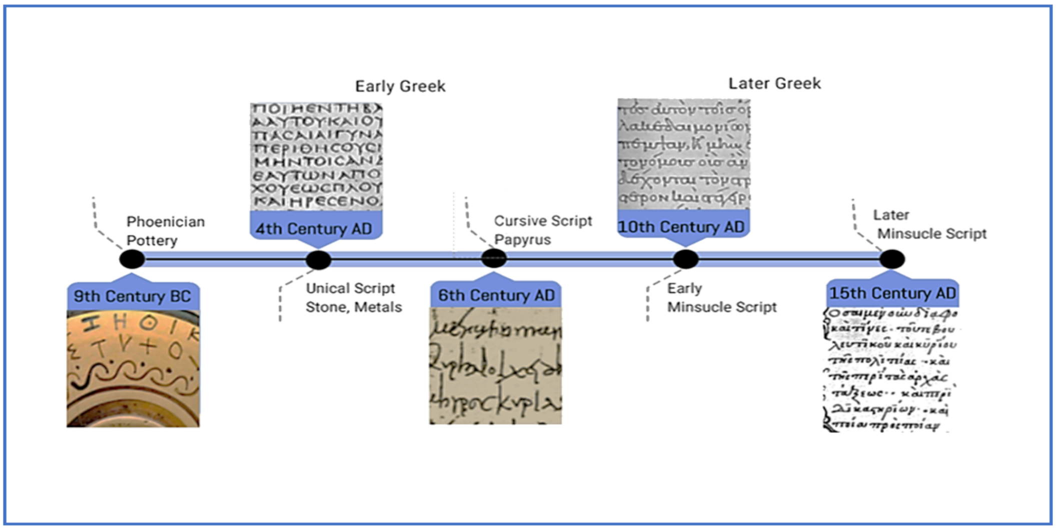


Figure 1.2: Evolution of Greek Script from Phoenicians to Unical to Cursive and finally to Minuscule Script.

Handwriting is also known to disclose demographic details such as gender, age, nationality, and handedness etc. [12]. Hence, it can be used to achieve the objective of a paleographer that is to extract information from ancient manuscripts such as keyword search, characterizing writing styles, identification of scribes, and credibility of manuscripts. Such analysis also provides potential applications for forensics. The identification of scribes carries significant importance as it can also be exploited to estimate the date and region in which the manuscript is produced by correlating with the 'active' period of the scribe [13].

A number of computerized solutions have been developed to assist the paleographers. For instance, the SPI (System for Paleographic Inspection) [14] software has been utilized by experts for their work in recent years. The tool compares and analyzes paleographic content morphologically. Such research may help a paleographer in inferring the origin of the manuscript, it is presumed that morphologically related documents could have been originated in the same cultural environment. This explicitly offers information on a manuscript that will be beneficial in identifying the scribe as well. Writer of a document can be categorized by capturing the writing style which is known to be specific for each individual. Writing style is typically exploited through a scale of observation that might be grapheme(character) level or global (page or paragraph) level. Textural features, for example, have been extensively employed to capture the writing style [15, 16, 17].

1.2 Challenges in Scribe Identification

Paleographers are particularly interested in tasks like identifying the scribe, determining the date and place of origin of a manuscript and so on. Such problems, naturally, require significant experience and domain knowledge. The writing style of every individual is exclusive. It is rarely possible for two people to have the same style of writing; even the same person cannot imitate their own handwriting. These variations in the handwriting patterns of different individuals are known as inter-class variability [18] while variation in ones own hand writing is known as intra-classs variability [19] as shown in Figure 1.3.

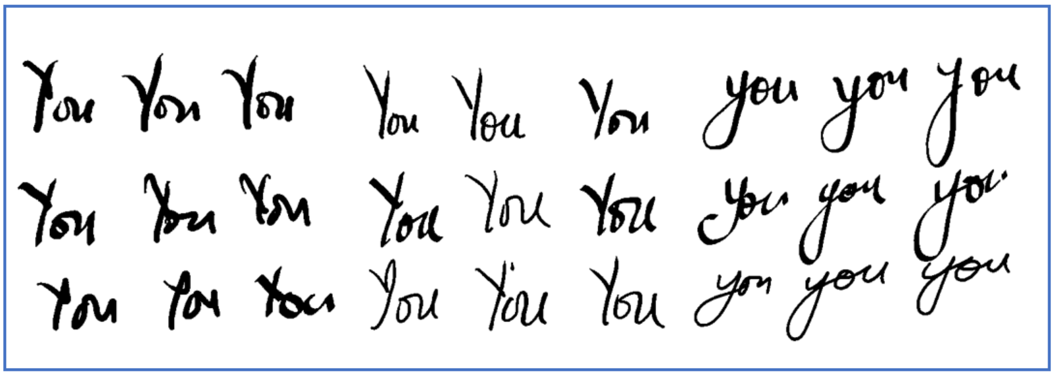


Figure 1.3: The inter and intra-class variability found in text of three separate writers with different instances of 'you' written with varying tools.

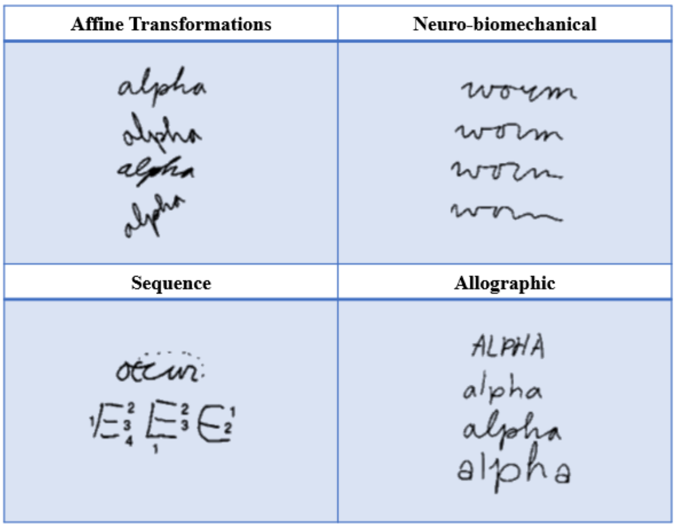


Figure 1.4: Factors responsible for variation in handwriting.

Schomaker et al. [20] have identified four factors responsible for variations in handwriting (Figure 1.4). These are affine transforms (shear, rotation, translation, scaling, etc.), allographic variations (character shapes employed by a writer), neuro-biomechanical variability(variation depending on the scribe's health, effort, and time), and sequencing variability (variable order of stroke production). Among these factors, the allographic variations provide the most useful information for automatic writer recognition [21].

Historical manuscripts, particularly from the pre-print era are much more challenging. They contain rich information and provide useful insight into the past. Drawings, shapes, embellishments, letters, signatures, and dropcaps not only provide explicit information on the content but, diverse social and cultural attributes are also manifested in the style of writing and its evolution. Historical manuscripts typically have noticeable writing visible from the back

Figure 1.5: Documents of GRK-Papyri Dataset suffering from severe damage.

of the page known as bleeding. Drawings or embellishments, such as ink blobs, note lines have a significant effect on the performance of the algorithm. These artifacts might also be affected by the aging process, holes or rips in the parchment, contrast variations that make the text indecipherable in extreme cases. Therefore, this might hinder the analysis as the foreground/background segmentation becomes difficult to discern as revealed in Figure 1.5. Furthermore, there are uneven strokes due to the different tools used for writing thus, making writer identification a more difficult task. Another important factor is the medium on which writing is produced that has evolved (stone, clay, papyrus, parchment, paper, etc.) as presented in Figure 1.6 and each medium has its unique challenges.

Recently feature learning using deep ConvNets has been widely investigated

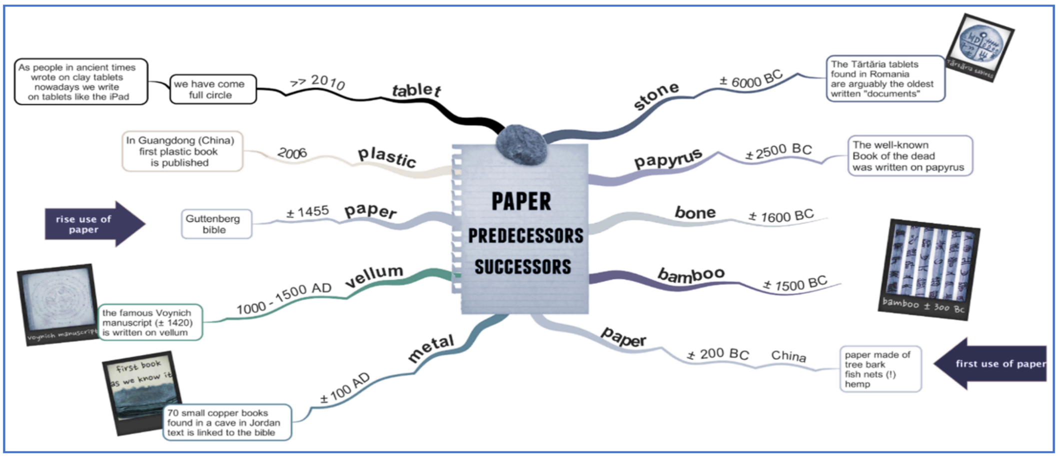


Figure 1.6: Writing material evolved from stone to paper.

to characterize the writer [22]. A major proportion of work on writer identification targets contemporary documents which do not offer the challenges encountered when dealing with ancient manuscripts. Noise removal, segmentation of text from the background, segmentation of handwriting into smaller units for feature extraction, etc. are a few of the challenges that hinder the direct application of many established writer identification methods to historical manuscripts.

1.3 Problem Statement

Scribe identification from offline images of handwriting is an interesting pattern classification problem that has been investigated for many decades. Despite significant research endeavors, the problem still remains challenging due to inter and intra-class variation found in handwriting. Handwriting is a reliable attribute in identifying the scribe. The volume of digitized data is growing and automated historical record analysis is becoming increasingly relevant. In many aspects, historical datasets differ from modern datasets where parameters are designed and evaluated, and hence techniques that are useful when handling modern datasets can not be directly used for historical data. The degradation of the writing medium also makes writer identification a more challenging task. This research aims to develop robust techniques to identify scribes from images of historical manuscripts. The goal is to exploit modern deep learning techniques and evaluate their performance on ancient manuscripts.

1.3.1 Research Questions

- Which pre-processing technique will address the degradation found in the dataset?
- How much text is eligible for writer characterization?
- What type of deep learning technique can be exploited for the problem at hand?

1.4 Key Contributions

The key contributions of the present study are listed in the following.

- Investigation of pre-processing techniques to segment handwriting from background in challenging degraded manuscripts.
- Writer characterization by studying the writing at two different scales of observation – small writing windows and patches around key points in writing
- Introduction of two-step fine-tuning of ConvNets where the models are first tuned on contemporary handwriting images and subsequently tuned on the small set of Greek handwriting on papyri.
- Validation of the proposed techniques on the benchmark collection the GRK-Papyri dataset; the reported identification rates outperforming the current state-of-the-art.

1.5 Thesis Organization

This document is organized as follows. Chapter 2 presents the state-of-the-art methods used for scribe identification. We have outlined, approaches and also common benchmarking datasets. Chapter 3 describes the method that we have adopted in order to achieve the objectives along with the key concepts behind the approaches. Chapter 4 outlines the metrics used to test our methods, describes the experiments, presents the findings we obtained and their interpretation. Chapter 5 incorporates the concluding remarks and recommendations for future work.

CHAPTER 2

LITERATURE REVIEW

Despite the emergence of digital documents, the importance of handwritten documents has prevailed. A wide range of systems utilizing pattern recognition and computer image processing techniques have been proposed to solve the problems related to automated handwriting analysis and in identifying the writer of the document. Hence, in recent years, computerized analysis of ancient handwriting has gained significant attention from the document recognition community [23, 24, 25, 26]. In our study, we will survey the approaches proposed in the last few years, thanks to the increasing interest in this domain from the document analysis community. The scale of observation at which features are computed is also critical as features can be extracted from complete pages, small patches of handwriting, text lines, words, characters, or even graphemes. These units represent the different scales of observations at which the handwriting is analyzed.

As discussed in the introductory discussion, a recent trend in writer identification is to learn features from data, typically using ConvNets. In our discussion, we will be focusing more on machine learning-based methods for writer identification as they are known to outperform the conventional handcrafted features. Readers interested in comprehensive reviews on this problem can find details in the relevant survey papers [27, 28, 12].

2.1 Writer Identification in Contemporary Documents

From the perspective of feature learning, ConvNets are either trained from scratch or pre-trained models are adapted to writer identification problem using transfer learning. Rehman et al. [23], for instance, employed the well-known AlexNet [29] architecture pre-trained on ImageNet [30] dataset as feature extractor. Handwriting images were fed to the trained model and extracted features were fed to a Support Vector Machine (SVM) for classification. In another deep learning-based solution, Xing & Qiao [24] introduced a deep multi-stream CNN termed as DeepWriter. Small patches of handwriting were fed as input to the network that was trained with softmax classification. Experiments on English and Chinese writing samples report high identification rates. Authors also demonstrated that joint training on both scripts leads to better performances.

Among other significant contributions, He et al. [31] proposed a deep neural network (FragNet) for scribe identification on single words or small text blocks of handwritten images. The proposed network was comprised of two pathways, a feature pyramid that extracts feature maps and the feature pathway to predict the writer. The method evaluated on four datasets namely IAM, CVL, Firemaker, CERUG-ER, and had reported promising results. Kumar et al. [32] proposed an optimal CNN model for the identification of writer on Indic languages. The results were computed at word and document levels. Tang & Wu [33] employed a CNN for feature extraction and the joint Bayesian technique for the identification of writers. In order to augment the size of training data, writing samples were split into words, and their random combinations were used to produce text lines. The technique was evaluated through an experimental study on the ICDAR2013 and the CVL dataset and Top-1 identification rates of more than 99% were reported in different experiments.

In another similar work, writer identification was carried out from Japanese handwritten characters using an AlexNet as the pre-trained model [34]. Subsequently, Fiel et al. [35] mapped handwriting images to feature vectors using a CNN and carried out identification using the nearest neighbor classifier. Christlein et al. [25] investigated unsupervised feature learning using SIFT descriptors and a residual network. The method was evaluated on ICDAR2017 and CLAMM16. Likewise, in [36], the authors employed a semi-supervised learning approach with ResNet. Weighted Label Smoothing Regularization (WLSR) was introduced to regulate the unlabeled data. Words in the CVL dataset were used as the original data while IAM words as the unlabeled set of data in the experimental study.

Among other studies on this problem, Keglevic et al. [26] proposed to learn the similarity between handwriting patches using a triplet network. The network was trained by minimizing the intra-class and maximizing the inter-class distances and the writing patches were represented by the learned features. He et al. [17] proposed a multitask learning method focused on a deep adaptive technique for writer identification from single word handwritten images. The re-usability of features derived for auxiliary tasks in the identification of writers was analyzed in this study. A new adaptive layer in CNN was introduced for exploiting deep features that enhanced the accuracy of the deep adaptive method as compared to simple adaptive and non-adaptive methods. Ino et al. [37] proposed an end-to-end approach, initially, a ConvNet was trained to extract local features that reflect the characteristics of individual handwriting in the whole character images and their sub-regions. Subsequently, from the training set randomly sampled image tuples were used to train the ConvNet and the local features that were extracted from the tuples were aggregated to form global features. This process was repeated for each training epoch.

Vincent et al. [38] proposed the use of CNN activation functionality as a local author recognition descriptor. A global descriptor was then obtained by Gaussian Mixture Model super-vector encoding and normalization with the KL-Kernel was also used for further improvement. In another work [39], authors presented an approach for offline text-independent writer identification focusing on the combination of the deep and traditional features. The authors also proposed deep architecture, an extended version of ResNet in which auxiliary information of handwriting thickness descriptor (HTD) was also added. In [40], authors extracted the patches from the handwritten images, and employed the hand-crafted descriptors to generate the local features. These features were assembled to form a description matrix. The vector of locally aggregated descriptors (VLAD) encoding was applied to the description matrix to extract a 1-D feature vector that represents the writer's writing patterns. A relatively recent trend was to exploit hyper-spectral imaging to capture handwriting images, mainly for forensic applications. Authors in [41], for example, demonstrated the effectiveness of employing multiple spectral responses of a single-pixel to characterize the writer. These responses were fed to a CNN to identify the writer. Experiments on the UWA Writing Inks hyper-spectral Images (WIHSI) dataset revealed that the potential of this interesting area for forensic and retrieval applications.

2.2 Writer Identification on Historical Documents

From the perspective of writer identification in historical manuscripts, as opposed to contemporary documents the literature is relatively limited [42, 43]. In some cases, standard writer identification techniques have also been adapted for historical manuscripts [44]. Recent work was reported in [45] that targets writer identification in medieval manuscripts (Avila Bible). Transfer learning was employed to detect text lines (rows) from images and the writer against each line was identified. Majority voting was subsequently applied on the row-wise decisions to assign a writer to the corresponding page and, page-level accuracy of more than 96% was reported. Sutder et al. [46] presented a comprehensive empirical study to investigate the performance of multiple pre-trained CNNs on analysis of historical manuscripts. The networks were investigated for problems like character recognition, dating, scribe identification, and handwriting style classification.

In other similar works, Cilia et al. [47] proposed a two-step transfer learning-based system to identify writers from historical manuscripts. The text rows in images were first extracted using an object detection system based on MobileNet. The CNN pre-trained on ImageNet was subsequently employed for writer identification on digitized images from a Bible of the XII century. Likewise, Mohammed et al. [48] adapted a known writer identification method (Local Naïve Bayes Nearest-Neighbour classifier [49]) for degraded documents and demonstrate high identification rates on 100 pages from the Stiftsbibliothek library of St. Gall collection [50]. The same technique was applied to the GRK-Papyri dataset [1] with FAST keypoints and reported a low identification rate of 30% (using a leave-one-out evaluation protocol).

In [51], the first-line projection was used to address the rotation problem, and images were then binarized using U-Net [52]. Later, two different feature extractors were used, SIFT [53] and pathlets. SIFT was used to extract junctions and corners while pathlet extracted the path of shapes. These features were explicitly encoded using VLAD, preceded by L2 normalization and dimensionality reduction. As a result, two global vectors were obtained which were concatenated as final style representation for writer identification. In [54] the textural information in handwriting was captured using a combination of oriented Basic Image Features (oBIFs) at different scales. The classification was carried out using the distance metrics which were combined to arrive at a final decision. An accuracy of 77.39% on ICDAR 2017 Historical WI dataset was achieved. Chammas et al. in [55] proposed to use SIFT technique to extract patches. These patches were then fed to a CNN (ResNet-20) for training. Later the results were encoded using multi-VLAD (with 5 layers) and exemplar SVM was applied to compare the results. The experimental study was carried on ICDAR2019 HDRC-IR dataset and 96.9% accuracy was achieved.

In [56] SIFT was used to extract features along with principle component analysis for dimensionality reduction, resulting in a visual vocabulary. The features were clustered using a Gaussian Mixture Model (GMM) and classified by Fisher kernel. In another work, Wang et al. [57], employed ResNet-50 for feature extraction, and introduced an optimized residual layer to obtain global descriptors. The study was carried on the ICDAR2017 (Historical-WI) dataset and reported a 72.4% identification rate.

2.3 Critical Analysis

A summarized overview of the notable writer identification techniques for contemporary and historical documents is presented in Table 2.1 and Table 2.2 respectively. A critical review of these methods reveals that features learned using ConvNets are known to outperform the conventional hand-crafted features. In recent years, researchers have followed a typical methodological flow of extracting random patches and then feeding them to convolutional neural networks. Networks are either trained from scratch or fine-tuned to the given dataset. A number of researchers have also explored the impact of the scale of observation. To achieve this, different segmentation techniques to extract words or lines have been employed while some studies have utilized feature extractors such as SIFT, oBIFs, etc. to study the discriminative property of writing patterns that might be more informative.

A major challenge in the analysis of handwritten documents is ground-truth creation. A paleographic analysis of large datasets is time-consuming hence labeled historical data remain scarce. Consequently, contributions made to computerized analysis of historical documents are limited. Furthermore, due to the challenging nature of images, as opposed to contemporary documents, the writer identification rates on an ancient manuscript offer a significant margin of improvement. Many studies are focused on fine-tuning of CNNs by employing the pre-trained ConvNets on the datasets. Moreover, CNN's are also used as feature extractors, and then these features are fed to another classifier. Historical documents are very different when compared to the natural images found in the ImageNet dataset. Therefore, training a model on ImageNet for historical document analysis may not be very useful [47, 45].

In addition, due to the complexity of historical documents, many of the methods that are known to perform well on contemporary datasets may not work well with ancient documents and significant pre-processing is required. The technique in [48] for instance, that reports high identification rates on contemporary documents, does not perform well on the papyrus dataset. The problem of writer identification in historical documents hence remains an open problem due to a wide variety of challenges such as inter and intra-class variability, changing of writing style with age, etc. All these problems increase the challenge ten folds due to the degradation found in historical datasets especially those that are written before the pre-print era for example on leaves, papyrus, etc.

Table 2.1: An overview on the Writer identification in Contemporary Documents.

Reference	Method			Dataset	Result	Limitations
	Feature	Network	Classifier			
[23], 2019	Sliding window (Patches)˜	AlexNet˜	SVM	QUWI	92.78% On English	Recent pre-trained model such as vgg16 might have performed better.
	Skew correction				92.2% On Arabic	
					88.11% On Both	
[24], 2016	Sliding window (Patches)˜	DeepWriter (CNN)		IAM	99.01%	———
				HWDB 1.1	93.85%	
[31], 2020	Sliding window (Patches)˜	FragNet (CNN)		IAM	85.10%	Cannot perform well on cursive handwriting.
				CVL	90.20%	
				CERUG-EN	77.50%	
				Firemaker	69.00%	
[33], 2016	Sliding window (Patches)˜	CNN	Joint Bayesian	ICDAR2013	99.60%	Pre-trained model can also be employed.
				CVL	99.70%	
[35], 2015	Sliding window (Patches)˜	CNN		ICDAR2013	88.50%	Pre-trained ConvNets haven't been investigated
				ICDAR2011	99.50%	
				CVL	98.90%	
[25], 2017	SIFT	Unsupervised (ResNet20)	Exemplar SVM	ICDAR2017	88.90%	———
				CLAMM16	84.10%	
[36], 2019	Line Segmentation	Semi-Supervised (ResNet50)		CVL	99.20%	———
				ICDAR2013	96.60%	
[26], 2018	SIFT	Triplet CNN		ICDAR2013	Precision: 86.1	Recent pre-trained model haven't been evaluated
	VLAD Encoding					
[39], 2020	Handwritten Thickness Descriptor(HTD)	ResNet		IAM	97.50%	The pre-trained model employed without the consideration of the dataset it was trained on.
				Firemaker	99.60%	
				CERUG-EN	89.1%	
				CVL	89.14%	

Table 2.2: An overview on the Writer identification methods for Historical Documents.

Reference	**Method**			**Dataset**	**Result**	**Limitations**
	Feature	**Network**	**Classifier**			
[45], 2020	Row Detector	VGG19		Avila Bible	96.48% on InceptionResNetV2 outperformed all	———
	MobileNetV2	ResNet50				
		InceptionV3				
	Single Shot Detector(SSD)	InceptionResNetV2				
		NasNet Large				
[46], 2019		VGG19		ICDAR2017 (Historical-WI)	DenseNet121 yield highest Precision= 34.6	Performance of model varies from dataset to dataset.
		InceptionV3				
		ResNet152				
		DenseNet121				
[48], 2018	FAST keypoints	NLNBNN		St. Gall	85.60%	———
[1], 2019	FAST keypoints	NLNBNN		GRK-Papyri	30%	The technique doesn't require contour extraction but the dataset at hand need a lot of pre-processing.
[51], 2020	Pathlet		Dimensionality Reduction	ICDAR2017 (Historical-WI)	90.10%	———
	SIFT			ICDAR2019 (HDRC-IR)	97.40%	
	VLAD Encoding					
[57], 2021		ResNet50		ICDAR2017 (Historical-WI)	72.4%	Target and the source the dataset model is trained at is not considered
[54], 2018	oBIFs		Distance Metrics	ICDAR2017 (Historical-WI)	77.39%	VLAD Encoding of the features might have increase the identification rate.
[55], 2020	SIFT	ResNet50	Exemplar SVM	ICDAR2019 (HDRC-IR)	96.60%	———

2.4 Benchmarking Datasets

In any research domain, the availability of datasets is one of the key requirements for the development and analysis of methods and the same is true for handwriting and writer identification. Over the last few years, the comparison of various methods attracted significant attention particularly from the writer identification community which led to the labeling of more data. In this section, the datasets that are used by state of the art to evaluate their approaches have been reviewed.

- **ICDAR2013 Dataset:** This dataset is collected for the writer identification competition held in conjunction with the 12th International Conference on Document Analysis and Recognition (ICDAR) [58]. It consists of images of handwritten samples by 250 writers in English and Greek. Each writer provided four samples of the text of which two are in English and two in Greek making a total of 1000 samples in the dataset. Figure 2.1 shows sample pages from the dataset.

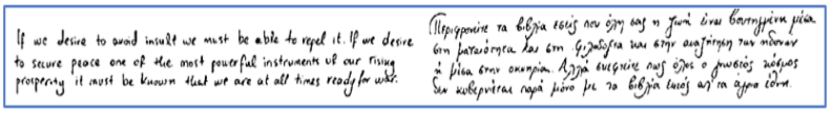


Figure 2.1: Writing sample of a scribe in English and Greek (ICDAR2013).

- **ICDAR2017 Dataset:** The Competition on Historical Document Writer Identification (Historical-WI) [59] organized by ICDAR 2017 introduced another dataset for scribe identification. There are 4782 sample images from 1114 different writers in the dataset. Documents from the digital archive of the Universitatsbibliothek Basel dating from the thirteenth to the twentieth Century are split into train and test sets. There are five images per writer out of which three images are for training and two for training. Sample images from the dataset are presented in Figure 2.2.

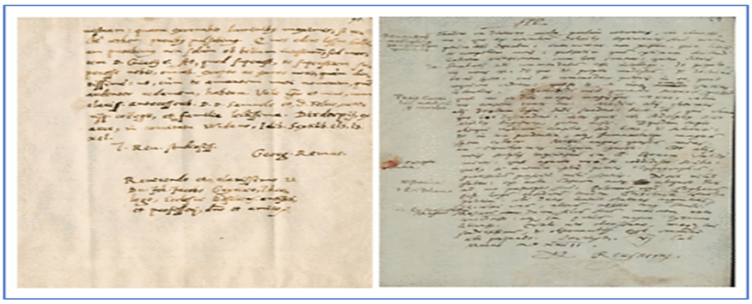


Figure 2.2: Image samples from ICDAR2017 (Historical-WI) dataset.

- **ICDAR2019 Dataset:** This dataset focuses primarily on writers of books written in the Middle Ages of Europe particularly from the ninth to the fifteenth century BC [60]. It includes 300 writers who contributed 1 page, 100 writers who contributed 3 pages, and 120 writers who contributed 5 pages, which resulted in 1,200 photos of 520 writers. The test dataset consists of 20,000 different images, where the number of samples varies from one to five samples per writer, most of the images are of the charters that are taken from Monasterium repositories.

Figure 2.3: Example of images from the Monasterium repositories (ICDAR2019 (HDRC-IR).

- **DIVA-HisDB Dataset:** For several tasks including layout-analysis, text-line segmentation, and writer recognition, DIVA-HisDB [61] a precise and annotated dataset of challenging medieval manuscripts is proposed. The database consists of 150 annotated pages with challenging formats from three different medieval manuscripts. It is a compilation of three medieval manuscripts selected, together with partners from e-codices and the Faculty of Humanities at the University of Fribourg, for the complexity of their style/layout.

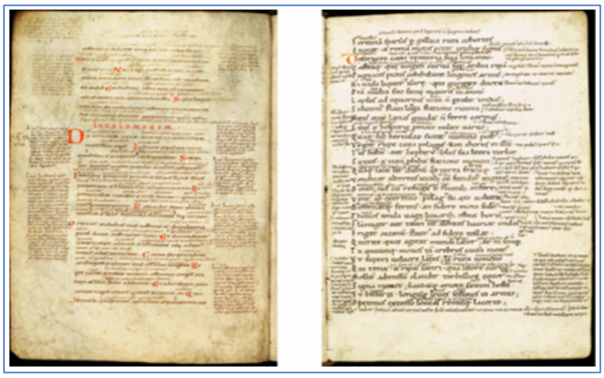


Figure 2.4: Image sample of two writers from the Medieval manuscripts.

- **St.Gall Dataset:** The archive of Saint Gall [62] is based on the Medieval Ninth Century Latin manuscript containing the hagiography of Vita Sancti Gaulli written by Walafrid Strabo. A manual copy of the work is placed in the Abbey Library of Saint Gall, Switzerland. It is written in the Carolingian script with ink on Parchment with a (probably) single experienced hand. Currently, there are 60 manuscript pages in the database.

Figure 2.5: Some of the image samples from St. Gall dataset.

- **Avila Bible** This bible [63] is written in Italy by at least nine scribes in the third decade of the twelfth century, is sent to Spain, where its decoration and text are completed by local scribes. Later, in the 15th century, the additions are made by another copyist to adapt the textual sequence to the modern liturgical sequence.

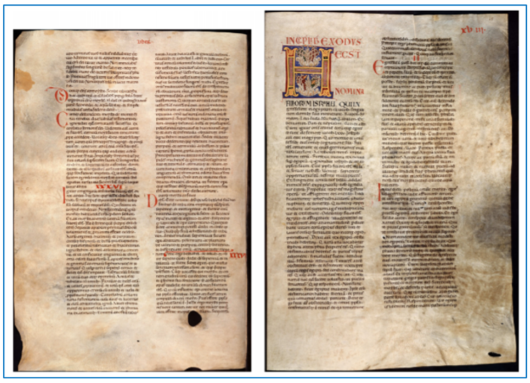


Figure 2.6: Example images from Avila Bible.

- **IAM Dataset:** IAM database [64] is initially proposed to evaluate the performance of handwritten text recognition. The initial version of the dataset comprised 1066 samples from 400 writers. But later extended to 1539 pages written by 657 writers out of which 356 writers contributed only one page while the remaining 301 wrote between two and sixty pages. The writing samples are also divided into text lines with a total of 13353 lines and an average of fourteen lines per writer.

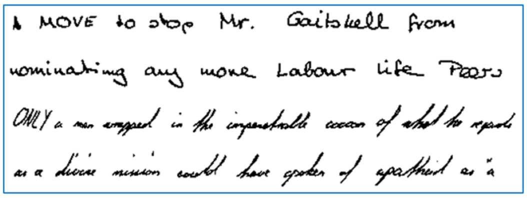


Figure 2.7: Sample of one of the writer from IAM database.

- **ClaMM16** The dataset is compiled for the ICFHR2016 competition on medieval handwriting classification [65]. There are 3000 pictures of Latin scripts from the handwritten books of 500 to 1600 CE. The dataset is divided into two thousand training and one thousand test images. The task is to automatically classify the test images into one of twelve Latin script types.

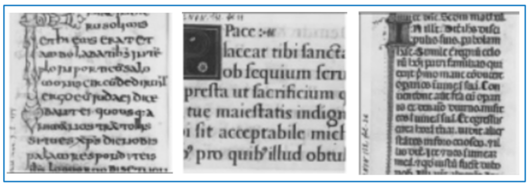


Figure 2.8: Samples of three different writers from the ClaMM dataset.

- **QUWI Dataset:** This dataset [66] includes both Arabic and English handwriting and can be used to study the performance of offline writer identification systems. It contains handwritten documents of 1,017 volunteers of varying ages, nationalities, genders, and educational backgrounds. Writers are asked to copy a specific text and generate random text that would allow the dataset to be used for both text-dependent and text-independent author identification tasks.

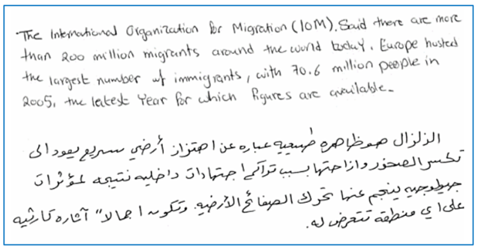


Figure 2.9: Arabic and English writing sample of the same writer in QUWI dataset.

- **CVL Dataset:** This is a public database for Writer Retrieval, Writer Identification and Word spotting [67]. A total of 310 writers participated in the dataset. 27 of which wrote 7 texts and 283 writers has to write 5 texts. The CVL database consists of pictures that have been selected from literature with German and English cursively written texts.

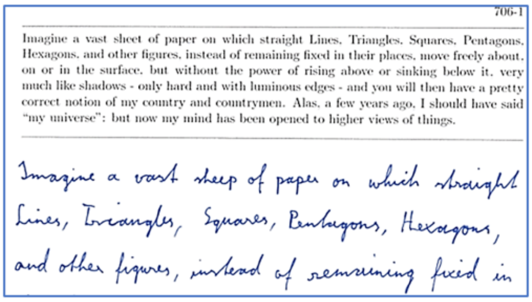


Figure 2.10: Sample page from writer 706 of CVL dataset.

- **JEITA-HP:** The corpus is initially compiled by Hewlett-Packard Japan and subsequently published by JEITA (Japan Electronics and Information Technology Association) [68]. It includes character images of 580 writers, including 480 writers in DATASET-A and 100 writers in DATASET-B writers. In principle, each writer's dataset includes 3,306 photos of 3,214 categories in which each Kanji character is written once, while each Kana/alphanumeric is written twice.

Figure 2.11: Samples from database JEITA-HP(B-99 and A-492).

- **Firemaker:** The dataset Firemaker [69] consists of a handwritten Dutch text by 251 students. Every student wrote four separate pages. A page with specified text in natural handwriting, specified text in upper case handwriting, specified text in 'forged' handwriting; and free text in natural handwriting. The handwriting pages contain a Dutch text with 612 alphanumeric characters in the upper-case script.

2.5 Summary

This chapter presented an overview of the techniques presented for the identification of scribes from handwritten images written in different languages. As a function of the type of documents, the study is organized into contemporary and historical documents. In recent years, high identification rates are reported on modern documents while from the perspective of ancient documents researchers are still exploring different techniques. Relatively higher identification rates are reported on documents that belong to the era of the 15th century or later but documents produced earlier than that suffer from severe degradations as they are written on materials such as palm leaves, papyrus, etc. thus, seems much more challenging.

CHAPTER 3

METHODOLOGY

In the preceding chapter, significant contributions in the field of writer identification both in contemporary and historical documents in the last two decades were discussed. As the previous studies indicate, ancient documents have very different image properties thus they require a substantial amount of pre-processing as they suffer from severe degradation such as bleeds rips, holes, etc. These problems have been differently addressed by different researchers. Authors in [70] for example use an image enhancement technique like CLAHE. Likewise, few studies employed Otsu binarization for ancient dead sea scrolls [71]. In our study, we start by enhancing the images and later experimented with various techniques for background and foreground segmentation. Once we achieved satisfactory results on the images, we proceeded to prepare the data for training by extracting the patches from the images. An important choice here is that how much information is sufficient to characterize the writer. It is difficult to determine how much text is optimal for writer identification [72]. Especially, ancient documents pose additional complexity and it is challenging to determine the window size for the patch extraction. Furthermore, we also investigated the effectiveness of small fragments that are extracted from the handwritten image by employing the FAST corner detector. Later the prepared data is fed to pre-trained convolutional neural networks for training. The approach primarily relies on characterizing patches of handwriting using machine-learned features in a two-step fine-tuning process. The scribe is identified on the basis of majority voting.

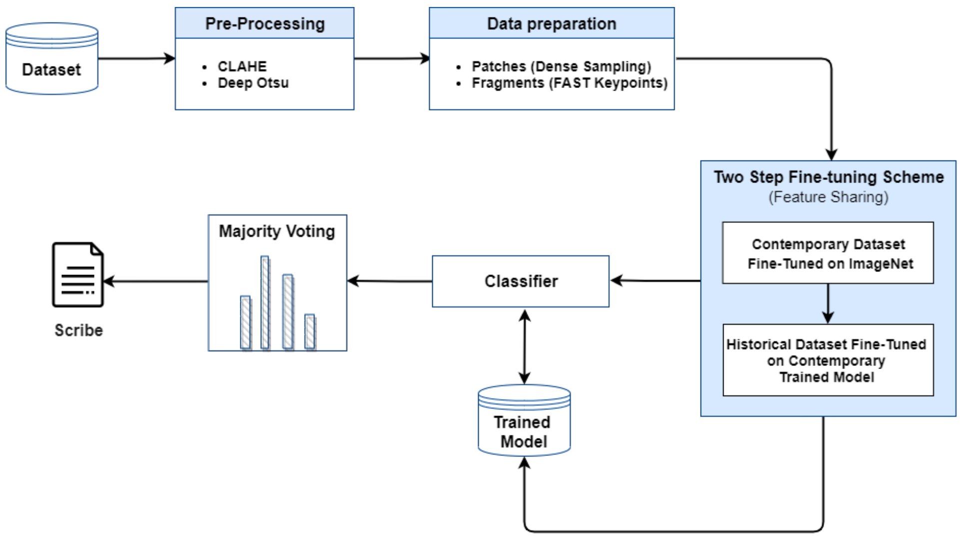


Figure 3.1: An Overview of Key processing Steps in the System.

We now present the details of the proposed method for the characterization of scribes from the challenging papyrus handwriting. We first introduced the dataset employed in our study followed by the details of pre-processing, data preparation, and writer identification through ConvNets. An overview of the key steps is presented in Figure 3.1 while each of these steps is discussed in detail in the subsequent sections.

3.1 Dataset

The experimental study of the system is carried out on the GRK-Papyri dataset presented in [1]. The dataset consists of 50 handwriting samples of 10 different scribes on papyri. All writings are in Greek and come from the 6th century A.D. The dataset has been made available for research along with the ground truth information of writers. Sample images from the dataset are shown in Figure 3.2.

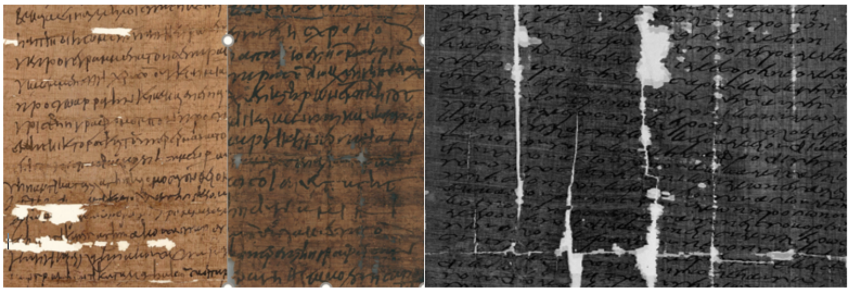


Figure 3.2: Sample Images of GRK-Papyri Dataset [1].

All images are digitized as JPEGs and height of images varies from 796 to 6818 pixels while the width values are in the range 177 to 7938 pixels. The DPI also varies from a minimum of 96 to a maximum of 2000. Few of the images are digitized as gray scale with others are three channel RGB images. The samples suffer from sever degradation including low contrast, holes and glass reflection etc. (Figure 3.2). The background contains papyrus fibers with varying sizes and frequencies adding further complexity from the perspective of automated processing. The samples are not uniformly distributed across the 10 scribes and the number varies from 4 to 7 samples per writer as presented in Table 3.1.

3.2 Pre-Processing

The primary objective of pre-processing is to improve the image quality by suppressing unwanted distortions and enhance image features that are important for further processing. Despite advanced photography and scanning equipment available, natural aging and perpetual deterioration often render many historical document images unreadable. The aging of these documents

Table 3.1: Distribution of samples per writer in the GRK-Papyri dataset.

Writer ID	Samples
Abraamios	5
Andreas	4
Dioscorus	5
Hermauos	5
Isak	5
Kyros1	4
Kyros3	4
Menas	5
Pilatos	6
Victor	7

has led to the deterioration of the writing media employed, due to factors like seepage of ink, smearing along the cracks, damage to the papyrus due to holes used for binding the manuscript leaves, and other extraneous factors such as dirt and discoloration. In order to suitably preserve these fragile materials, digital images are captured using high definition digital cameras in presence of an appropriate light source instead of scanners. Digitizing Papyrus manuscripts pose a variety of problems. They cannot be forced flat and the light source used for digital cameras is usually uneven and the very process of capturing a digital image introduces many complications. These factors lead to poor contrast between the background and the foreground text. Therefore, innovative digital image processing techniques are necessary to improve the legibility of the manuscripts. To sum up, historical document images pose several challenges to pre-processing algorithms, namely low contrast, non-uniform illumination, noise, scratches, holes, etc. Hence, Prior to feeding the images for feature extraction, we need to process the images. Since the dataset comprises both colored and grayscale images with diverse backgrounds of papyrus fiber, directly feeding raw images may lead to learning features that could be linked with the background information rather than handwriting. We therefore first convert all images to grayscale and use the image enhancement technique to improve the quality of the image to yield better performance. We first apply CLAHE on the grayscale image followed by different binarization techniques as elaborated in the following.

3.2.1 CLAHE

The modified version of the Adaptive Histogram Equalisation (AHE) is known as Contrast Limited Adaptive Histogram Equalisation (CLAHE) [73]. AHE has a disadvantage over CLAHE as it amplifies the noise. In CLAHE, this is reduced by defining a threshold, called clip limit that clips the histogram, before calculating the Cumulative Distribution Function (CDF). The comparison is presented in Figure 3.3.

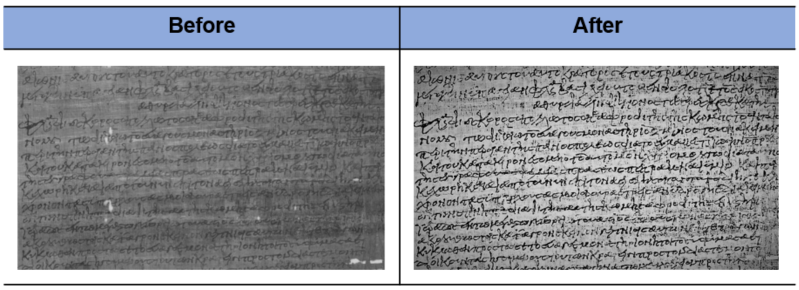


Figure 3.3: Comparison of image from dataset image before and after the implementation of CLAHE

Furthermore, we have also pre-processed images in different ways to investigate which of the representations would be optimal. These include:

- Binarization using adaptive (Sauvoloa [74]) thresholding.
- Application of Canny edge detector to preserve edges of writing strokes only.
- Edge detection on adaptively binarized images.
- Binarization of images using a recent deep learning based technique – DeepOtsu [75].

The output images resulting from these different types of processing are illustrated in Figure 3.4.

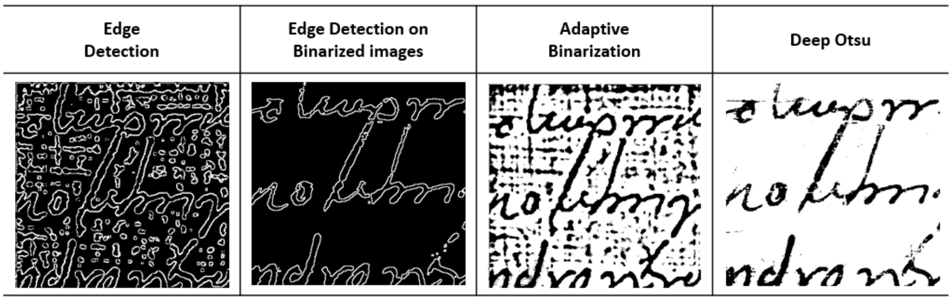


Figure 3.4: Output images resulting from different types of pre-processing.

3.2.2 Binarization using Adaptive Thresholding

Binarization is another important aspect in document image pre-processing where an input colored or a gray level image is transformed into a black and white image. This task is also performed to minimize the effect of physical document degradation. The binarization method distinguishes the content of the document from the noise by classifying each pixel as either foreground or background. Sauvola [74] presented a local thresholding technique that is useful for images where the context is not uniform, particularly for the recognition of text. Instead of computing a single global threshold for the entire image, multiple thresholds are computed for each pixel that takes into account the mean and standard deviation of the local neighborhood (defined by a window centered around the pixel). Furthermore, two new algorithms are used to determine the local threshold for each pixel. The soft decision control algorithm is used for thresholding background and local mean and variance of the gray values are applied to the textual region. Object contours and image information are better preserved by local thresholds, but it is also much more sensitive to noise. The threshold for each window is calculated as follows:

$$T\left(i,j\right) = \mu\left[1 + k\left(\frac{\sigma}{R}\right) - 1\right] \tag{3.1}$$

Where, R is the dynamic range of the standard deviation σ of the window and the parameter k takes positives values in the interval $[0.2, 0.5]$ while μ is the mean. The value of k and window size will have significant effect on the quality of image but R will have very little effect.

3.2.3 Canny Edge Detection

John Canny introduced an edge detection method [76] and since then it has been one of the most widely used edge detection techniques. The detection is based on the premise that the intensity at the edges is high in the images. In the algorithm, first, the noise is minimized by employing the Gaussian filter to smooth images.

$$g\left(m,n\right) = G_\sigma\left(m,n\right) + f\left(m,n\right) where, G_\sigma = \frac{1}{\sqrt{2 \times \pi\sigma^2}}\exp\left(-\frac{m^2+n^2}{2\sigma^2}\right) \tag{3.2}$$

Then gradient computation is achieved by computing the rate of change of intensity $(g(m,n))$ along the path of the gradient to obtain:

$$M\left(n,n\right)=\sqrt{g_n^2\left(m,n\right)+g_n^2(m,n)} \tag{3.3}$$

and

$$\theta\left(n,n\right)=\tan^{-1}\left[\frac{g_n\left(n,n\right)}{g_r\left(m,n\right)}\right]. \tag{3.4}$$

At each pixel location, pixels on the chosen threshold T are compared to pick the local maxima in a neighborhood of 3×3 along the path of gradients. This process is known as non-maxima suppression. As a result, thin, yet broken edges are formed. These broken edges are connected by using a hysteresis threshold. There are two thresholds for hysteresis thresholding: high and low thresholds. The gradient value of the pixel is higher than the high threshold then kept as an edge. But if the gradient value of pixel lies between the high and low threshold then the pixels are assessed for probable connection to the edge, if linked then kept otherwise discarded. While if the gradient value of the pixel below the low threshold then they are discarded.

3.2.4 Edge Detection on Adaptively Binarized Images

We also experimented by employing Canny edge detection on the image that is binarized by adaptive thresholding technique. The motivation is that since after binarization we still encountered noise in the images, so by employing the edge detection technique we wanted to lessen the noise as this algorithm detects edges and seems to discard pixels with lower gradient.

3.2.5 Deep Otsu

Unlike the traditional approaches in which the binary label of each pixel is predicted on the input image, in Deep Otsu [75] a neural network is trained to learn the degradations in document images and generate the uniform images, which enables the network to enhance the output iteratively. Firstly, we provide the CNN model with our degraded input image, which learns the relationships in the image and provides an output image. Secondly, that image data from CNN is added to the degraded input image to get the final output. Thirdly, we again send this output as an input to the CNN and repeat the process. This procedure of iteratively learning the relationships and providing incremental outputs as inputs helps the model to learn the 'Ground-Truth; of the image data. That ground truth is the degradation present in the image.

$$x = x_u + e \tag{3.5}$$

The x is the input image, x_u is the ground truth which is the uniform image without degradation and e represents the degradation in the image. When the model learns the relationships for the image x_u, we again give this output as an input to the model, where it again tries to learn according to the features present and thus, removes out the errors which are not representable using the new input image. This iterative procedure removes the error data from the input image giving us a candidate that can be used easily with Otsu thresholding thus, giving us the binarized image. The model used here for CNN is the U-net model (Contraction then Expansion model). Basically, we contract the features first, and then using up convolution we expand the data thus, keeping the most significant relationships in the image. This, in turn, gives us the actual features of the image and removes the degradation when used iteratively using the two iterative techniques. Recurrent refining is basically, recursively setting the output as input for some number of times to the same CNN model. This is quite useful when you have fewer features in your image and your images are not so complex. Stacked Refining uses different CNN models, thus, each model can learn different features when you have lots of important features in your image and can also refine the data much more efficiently as each model can be trained in a different way for weighing different features and degradation. Later, Otsu Global threshold employed for binarization.

3.3 Data Preparation

The division of handwriting is an important step as a 'good' division would allow exploiting most of the redundancy in writing [72]. We investigate two different methods to divide handwriting for feature extraction. These are discussed in the following.

3.3.1 Patch level

At first, we divide the handwriting image into fixed-sized windows using dense sampling. Ideally, the window size should be adjusted according to the writing details (e.g. ink thickness, character height, etc.). Moreover, when employing pre-trained ConvNets in a transfer learning framework (fine-tuning them on the target dataset), the resolution of images must match the input expected at the network. Naturally, resizing the complete page to a small square and feeding it to a network is not very meaningful as not only all writer-specific information is likely to be lost but the aspect ratio is also highly disturbed.

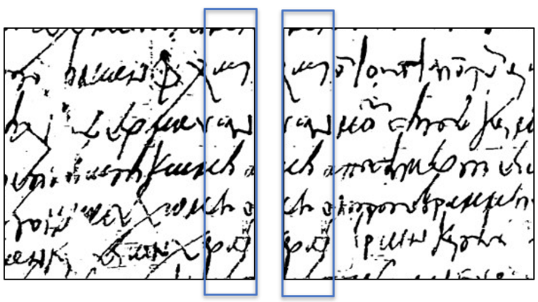


Figure 3.5: Dense Sampling Sample from the image in Dataset.

We, therefore, carry out a dense sampling of the complete image using overlapping $'n'$ size squared windows as presented in Figure 3.5. The size of the window determines the scale of observation and extracting square windows ensures that the aspect ratio is not disturbed once the extracted patches are resized to match the input layer of pre-trained ConvNet.

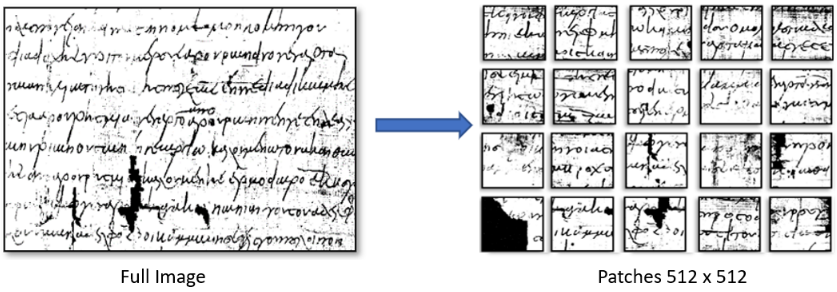


Figure 3.6: Patches extracted from a Binarized image in the Dataset.

Few patches of size 512×512 extracted from one of the images in the dataset presented in Figure 3.6. The size should be large enough to provide ample of information on the writer's style and small enough to ensure good performance in identifying the writer as represented in Figure 3.7.

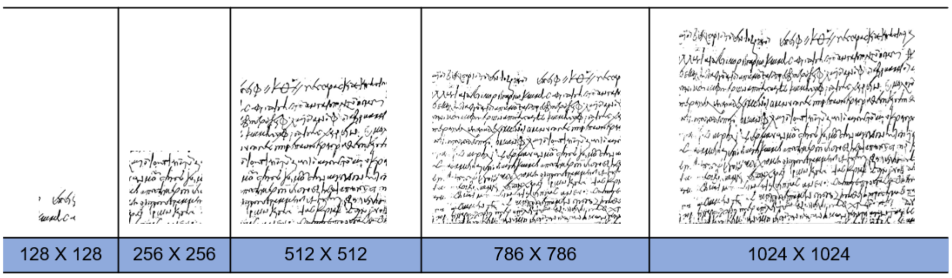
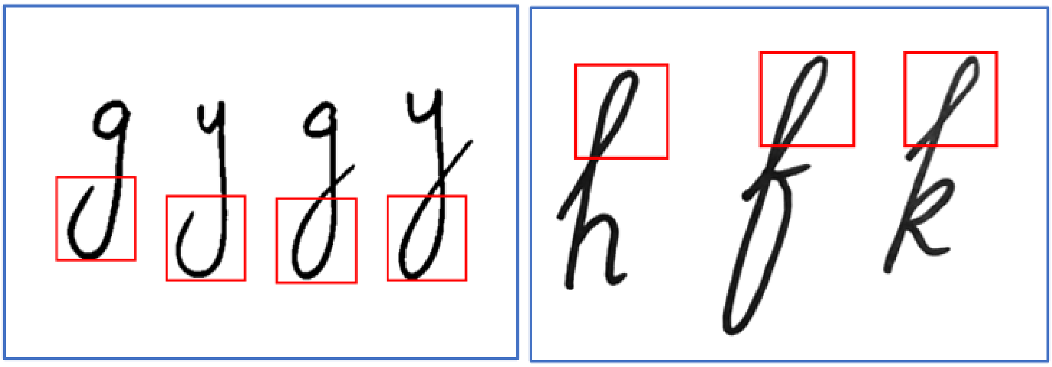


Figure 3.7: Different Size of Patches extracted from a Binarized image in the Dataset.

3.3.2 Fragment Level

It is known that all characters and their combinations are not equally discriminative to characterize the writer. Moreover, recently the research focus for writer identification has moved towards the unique patterns found in an individual's handwriting. When a person writes, some alike characters especially in cursive handwriting are produced more or less by similar hand motions as illustrated in Figure 3.8.



(a) Characters g, y

(b) Characters h, f, k

Figure 3.8: Different characters exhibiting Discriminative property.

Also, feeding compressed but informative patches could be more effective than random patches having more information. Therefore, we have also chosen to assess another technique that is to first extract keypoints in handwriting, and then using these keypoints as centres, we extract small patches around them. Extraction of key points has proved to be effective for writer identification in a number of studies [56, 55, 35]. Typically, key points represent the locations where the boundary of the object changes the direction abruptly. Likewise, the intersection between several edge segments are also potential keypoints. In this study, we have employed FAST [54] extract keypoints.

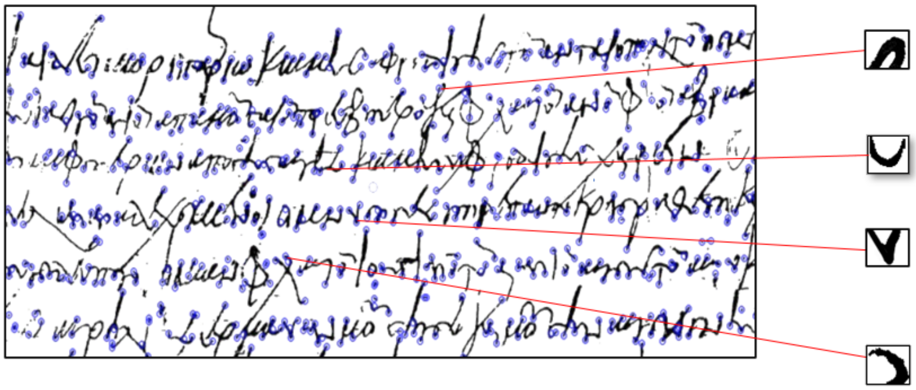


Figure 3.9: Fragments extracted from a binarized image in the dataset using FAST keypoints.

Features from accelerated segment test (FAST) is proposed in [54]. The method decides that whether a pixel in an image can be considered as a key point by evaluating the circular region around it. The technique chooses a candidate point p with an intensity I_p selecting a threshold value t. It can be computationally expensive to determine the difference of 16 pixels for every pixel in an image. Hence, to address this problem enhance version of the is detector is employed that takes the value of four points (1, 5, 9, and 13). If two or more of these four points do not satisfy the condition greater than$I_p + t$ or less than $I_p - t$ then the point is rejected. Points for which at least three out of four points satisfy the condition can be the candidate keypoint for which subsequent points are checked.

A major drawback of this technique is that it detects a large number of key-points that are very close to one another. To tackle this problem, non-maxima suppression is applied, to keep the high response points only. We then compute the sum of differences between the intensities of the 16 neighboring pixels and I_p. The scores of neighboring candidate key points are compared and the one with the lower score is eliminated. Once the keypoints are determined we extract 50×50 centered around the key points. These patches are subsequently employed as an input to the CNN as discussed in the next section.

3.4 Training and Recognition

As discussed earlier, we employ ConvNets for feature extraction and subsequent classification. The classical framework for holistic recognition techniques involves freezing the initial layers and fine-tuning later layers of pre-trained ConvNets. In the following sections, we first present an overview of CNN's followed by a discussion on how we adapt them to our problem.

3.4.1 Convolutional Neural Network

Convolutional Neural Networks has been presented in the 1990s for the first time. The research community did not pay much attention to them as they required large datasets and high-performance computing machines which were not available at that time. CNN gained attention with the availability of graphical processing units and large datasets like ImageNet [30]. Today, CNN represents state-of-the-art feature extractors and classifiers. They have outperformed conventional techniques on a number of recognition problems. A typical CNN includes convolutional and pooling layers followed by fully connected layer(s) as illustrated in Figure 3.10. For completeness, these layers are briefly discussed in the following.

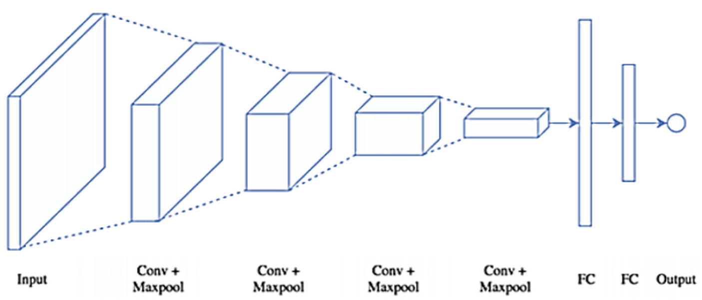


Figure 3.10: CNN block diagram.

- **Convolutional layer:** : is probably the most important layer in a deep neural network. The convolution of an image with a given filter is a feature extractor. In the initial layers, these features could be edges at different orientations or curves, etc. In the deeper layers, convolution is applied on the output map of the previous layer hence they extract high-level features that are generally domain specific. The number and size of the filter in each layer may vary. The size of output volume at each layer is a function of three hyperparameters, stride, depth, and padding. Stride refers to the jump of a filter on both horizontal and vertical direction after every convolution, the total number of filters in a layer refers to depth while the number of rows and columns added at the border of the input image to complete convolution operation is called padding.

- **Pooling Layer:** is also known as the downsampling layer. This layer serves to reduce the number of parameters as well as over-fitting. Max pooling, average pooling, and L2-norm pooling are the commonly employed techniques in this layer, max-pooling being the most popular. The idea of pooling is to consider a small neighborhood (for example 2×2) and replace the neighborhood with a single value (for example the maximum).

- **Fully Connected Layer:** Features calculated by the sequence of convolutional and pooling layers are given as input to one or more fully connected layers which act as a classifier. The final layer has neurons equal to the number of classes and each neuron is expected to fire once the network is presented with an example of the corresponding class.

3.4.2 CNN's and Transfer Learning

As discussed in the earlier sections, deep ConvNets have become the gold standard for feature extraction as well as classification. Designing a new architecture and training CNNs from scratch for every problem, however, is neither required nor feasible. In most cases, architectures and weights of ConvNets can be borrowed from those trained on millions of images and made publicly available by the research community. This concept is commonly termed transfer learning and has been successfully applied to a number of recognition tasks. Transfer learning can be implemented either by fine-tuning or using the ConVets as a feature extractor. Each of these is briefly discussed in the following.

- **Feature Extraction:**A common practice is to use the pre-trained CNN only as feature extractor use a different classifier to classify the objects under study. In such cases, the fully connected classification layers are discarded, and features extracted from the last layer of the pre-trained model are fed to the desired classifier.

- **Fine Tuning:** In fine-tuning, generally, a pre-trained network is adapted to a new dataset by continuing the backpropagation on the new dataset and replacing the last layer to match the new dataset. In some cases, the weights of initial layers are frozen as the earlier layer are meant to compute low-level features which can be common across different datasets. It is also common to use a smaller learning rate while continuing backpropagation.

3.5 Transfer Learning for Scribe identification

In our study, we employ the pre-trained ConvNets by fine-tuning them to our problem. More specifically, we employ three standard architectures namely VGG16 [29], InceptionV3 [77] and ResNet50 [78].

3.5.1 Two-step Fine Tuning

An important consideration in adapting ConvNets for feature extraction and classification is the similarity of source and target datasets. It is important to mention that most publicly available models are trained on the ImageNet [30] dataset. Since we deal with handwriting images that are different from the images in the ImageNet dataset, we employ a two-step fine-tuning. First, we fine-tune the networks using the IAM handwriting dataset [64] which contains writing samples of more than 650 writers. Although these are contemporary samples and do not offer the same challenges as those encountered in historical documents, nevertheless, since these images contain handwriting, we expect an enhanced feature learning. Once the networks are fine-tuned on IAM handwriting samples, we further tune them on the writing patches in our papyri dataset. The softmax layer of the final network is changed to match 10 scribes in our problem. The employed networks are briefly outlined in the following.

- **VGG16**: is one of the classic deep neural network architectures which is proposed by Karen Simonyan and Andrew Zisserman from the University of Oxford. In 2014, this architecture is evaluated on ImageNet dataset which is compiled for ImageNet Challenge and reported 92.7% Top-5 test accuracy. The major improvements found in this architecture, as compared to AlexNet, included the replacement of large filters of sizes 11 and 5 in the first and second convolutional layers, respectively with multiple filters of 3×3 one after another. The architecture of the model is presented in Figure 3.11.

- **InceptionV3**: is a well-known convolutional neural network and is the third version of GoogleNet [79] which is widely used in object detection

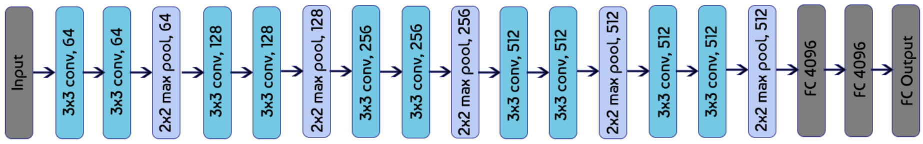


Figure 3.11: Architecture of VGG16.

and image analysis. The network has 48 layers in total. The model contains symmetric and asymmetric building blocks that include factorized 7 × 7 convolutions, label smoothing, and an auxiliary classifier that is used as a regularizer along with average pooling. Batch Normalization is also used extensively throughout the model and also applied to activation inputs. These changes in the architecture have minimized the error rate. The model of the architecture in Figure 3.12.

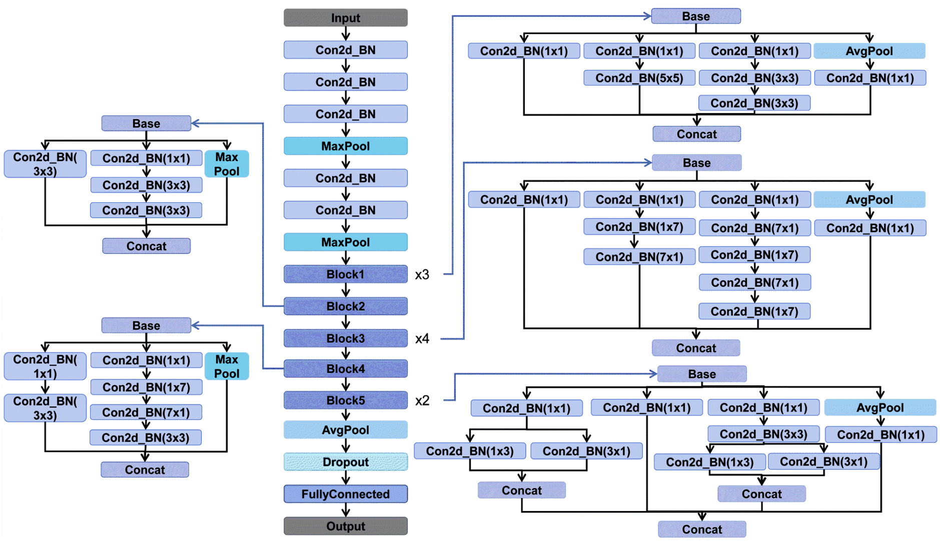
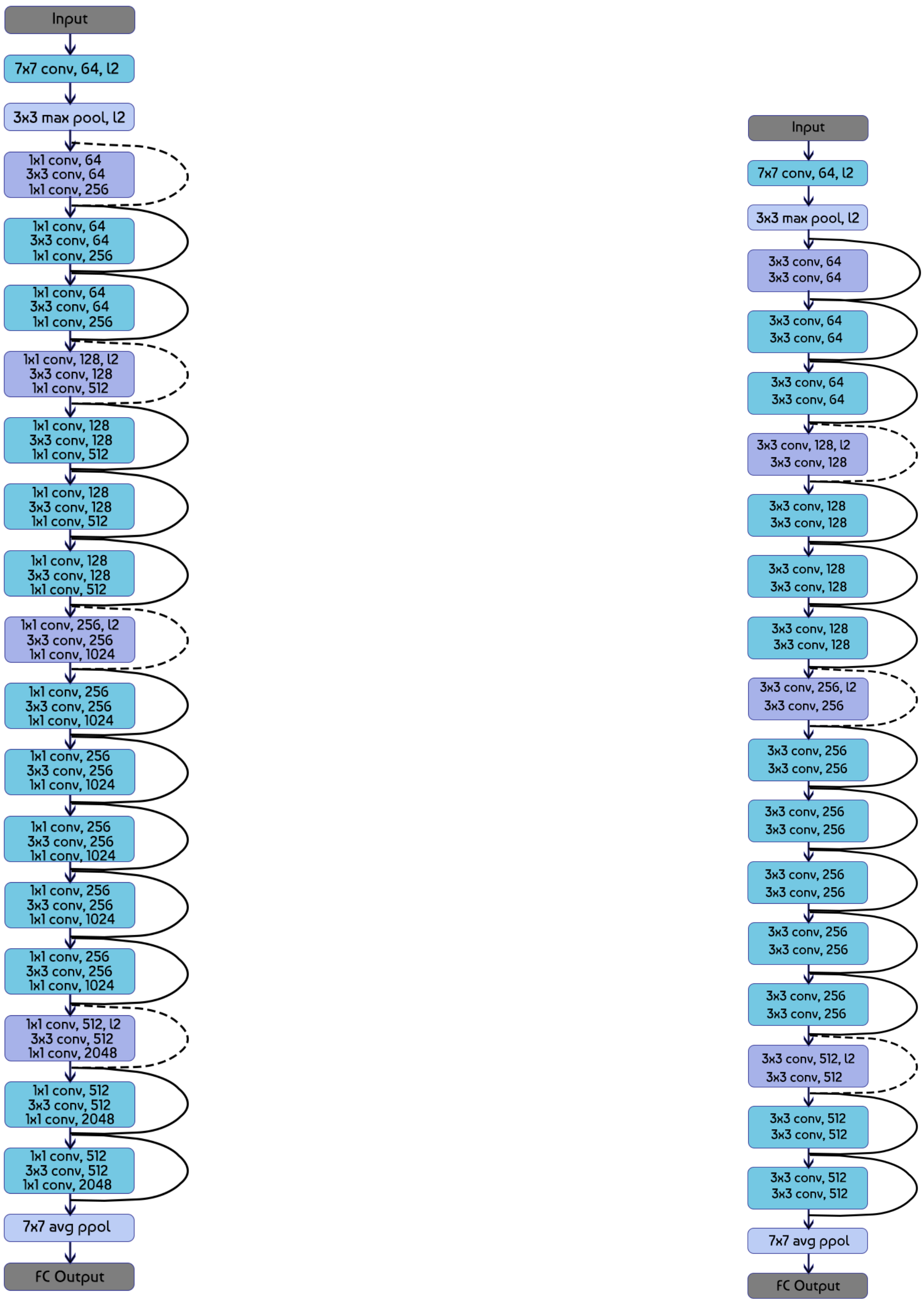


Figure 3.12: Architecture of InceptionV3.

- **ResNet**: Deeper networks are capable of learning more complex input representations and functions. However, some of the research found that adding more layers has an adverse effect on the performance of the model known as degradation problem [78]. This problem is resolved by adding residual blocks in which intermediate layers learn a residual function (unique and new feature map) with reference to the input of the block. It can be regarded as a refinement step, if it is no longer needed then intermediate layers gradually learn to adjust the weight towards zero such, residual block represents an identity function. There are different versions of ResNet, including ResNet-18, ResNet-34, ResNet-50, and so on.

In our study, we have employed ResNet-34 and ResNet-50 to analyze how the model evolves if the model is shallower or deep. ResNet34 network uses a 34-layer plain network architecture inspired by VGG-19 in which then the shortcut connections are added. These shortcut connections then convert the architecture into the residual network as shown in the Figure 3.13b. As opposed to the version of ResNet34, in ResNet50 the shortcut connections have skipped three layers and also 1×1 convolution layers added as illustrated in Figure 3.13a.



(a) Architecture of ResNet50

(b) Architecture of ResNet-34

Figure 3.13: Difference Between Two Variants of ResNet Architecture.

Table 3.2: Summary of Pre-trained ConvNets.

Network	Model	Contribution	Parameters
Series	VGG16	Homogeneous Topology Small Size Kernels	138M
DagNet	InceptionV3	Handle the Problem of representational Bottleneck Replaced Large kernels with Asymmetrical small kernels	23.6M
	ResNet	Residual Learning Identity mapping Based skip connections	25.6M

A summary of different architectures employed in our study is presented in Table 3.2.

For each of the models, we first fine-tuning it on the IAM database. We employed a batch size of 32, the activation function is ReLU, the optimizer is Adam and the loss function is categorical cross-entropy. With VGG16, we freeze the first 4 layers of the model. Similarly, in InceptionV3 10 layers while for ResNet-50 and ResNet-34, the first 12 and 8 layers are frozen respectively. Furthermore, the classifier is discarded to be replaced as per the dataset (650-writers) with softmax as an activation function and then fine-tuned on the IAM dataset.

In order to avoid over-fitting, we opted to employ a dropout layer (with a dropout rate of 0.2) and kernel regularizer L1 as it shrinks the less important feature's coefficient close to zero thus, removing some features altogether. When we fine-tuned the model that is trained on the IAM dataset on the Papyrus dataset the fully connected layers are replaced by 10 writers as per the dataset with softmax as an activation function. Once the models are tuned, we feed patches/fragments of writing in a test image and identify the writer for each patch/fragment. Subsequently, the decisions are combined for a document using a majority voting. Details of these experiments are presented in the next chapter.

3.6 Summary

This chapter introduced the proposed scribe identification technique along with the pre-processing techniques investigated. Features are extracted using patches of handwriting as well as by first finding key points in the handwriting and then extracting features. A number of pre-trained ConvNet models are employed in a two-step fine-tuning framework where networks are first tuned on contemporary and then on the Greek handwriting. The next chapter presents the details of the experimental study carried out to validate the proposed methods.

CHAPTER 4

ANALYSIS & RESULTS

This chapter presents the details of the different experiments with results and discussion. We first introduce the experimental protocol followed by the scribe identification performance of different models using patches and fragments of handwriting.

4.1 Experimental Protocol

The GRK-Papyri dataset is provided to carry out writer identification tasks in two experimental settings.

- Leave-one-out Approach.

- A training set of 20 and a test set of 30 images.

Since we employ a machine learning-based technique, experiments under a leave-one-out approach would mean training the system 50 times for each evaluation. We, therefore, chose to employ the training and test set distribution provided in the database i.e. 20 images in train and 30 in the test set. At this distribution there might be a contradiction that data is imbalanced but here we would like to clarify that the length of text in images of test is fairly small. Thus, number of images per class is more in test as compare to train.

4.1.1 Document Level Vs Patch Level

During the evaluation phase, the query document is divided into patches where each patch's probabilistic score is maintained which indicates the probability of patch being produced by each of the writer in reference base is $S_i = \{P_i, P_2, \ldots, \ldots, P_n\}$ where, S_i is the score of patch i (Patch Level) and N indicates the number of writers. The identity of the writer is decided by combining the scores of patches using the majority Voting to come to conclusion at Document Level (meaningful from viewpoint of application). The pipeline of how document level Scores are computed is illustrated in Figure 4.1.

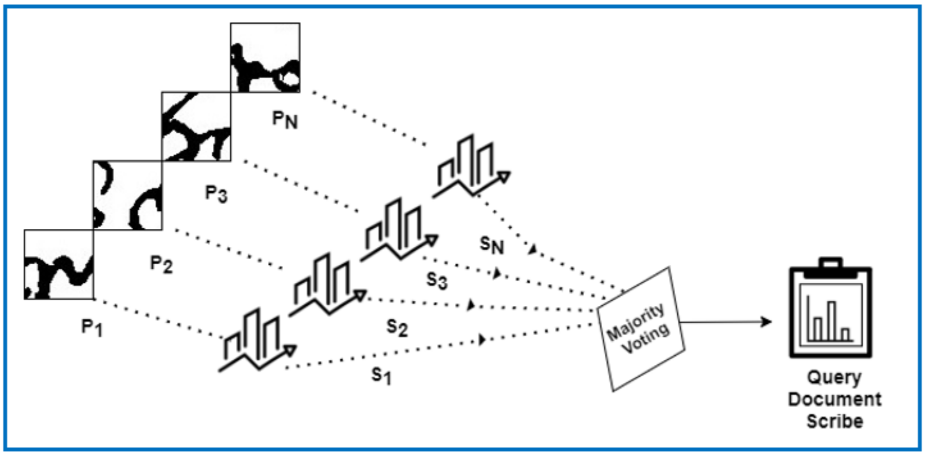


Figure 4.1: Pipeline of how Document Level Scores are computed.

4.2 System Performance

The efficiency of the system is exploited on various paradigms. We first present the identification rates as a function of different pre-processing techniques. These classification rates are computed by fine-tuning InceptionV3 first on the IAM dataset and subsequently on the training images in the GRK-Papyri dataset. The results are evaluated at the patch level as well as document level by applying a majority vote on the patch level decisions.

Table 4.1: Writer identification rates for different pre-processing techniques (Two-step fine-tuning of InceptionV3).

	Patch Level	**Document Level**
Adaptive Binarization [74]	0.11	0.32
Canny Edge Detection [76]	0.10	0.27
Edge Detection+Binarization	0.38	0.16
Deep Otsu [75]	**0.27**	**0.48**

It can be seen from Table 4.1 that among the different pre-processing techniques investigated, DeepOtsu reports the highest identification rates of 27% at the patch level and 48% at the document level. The subsequent experiments are therefore carried out using DeepOtsu as the pre-processing technique.

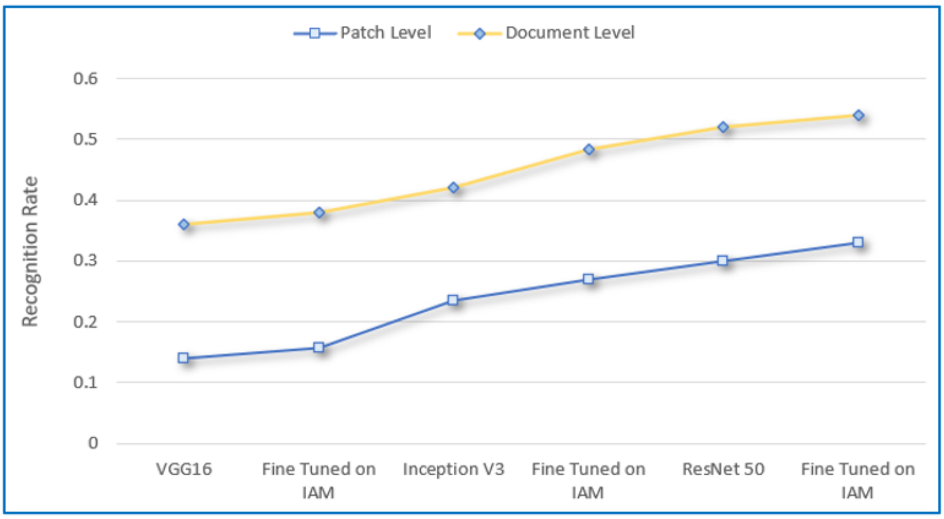


Figure 4.2: Writer identification rates as a function of patch size.

In the next experiment, we present the identification rates by directly fine-tuning the models from ImageNet to our dataset (single-step tuning) as well as by first tuning them on the IAM dataset and subsequently on the papyri dataset (two-step tuning). It can be seen in Table 4.2 that in all cases two-step fine-tuning serves to enhance the identification rates (by 2% to 6%). The highest document level identification rate is reported by fine-tuning ResNet-50 and reads 54%. Considering the complexity of the problem and the small set of training samples, the reported identification rate is indeed very promising.

Table 4.2: Performance of single and two step fine tuning on different pre-trained ConvNets (Patches).

	Fine-Tuning Scheme	**Patch Level**	**Document Level**
VGG16 [80]	ImageNet→Papyri	0.14	0.36
	ImageNet→IAM→Papyri	0.16	0.38
InceptionV3 [77]	ImageNet→Papyri	0.24	0.42
	ImageNet→IAM→Papyri	0.27	0.48
ResNet-50 [78]	ImageNet→Papyri	0.30	0.51
	ImageNet→IAM→Papyri	**0.33**	**0.54**

We also study the impact of patch size (scale of observation) on the identification rates. The Document level identification rates with two-step fine-tuning of InceptionV3 and ResNet-50 as a function of patch size are summarized in Figure 4.3. It is interesting to observe that both the models exhibit more or less similar trends and the highest identification rates are reported at a patch size of 512×512, i.e. 48% and 54% for Inception and ResNet respectively. Too small or too large patches naturally report relatively lower identification rates indicating that the scale of observation is a critical parameter that must be carefully chosen.

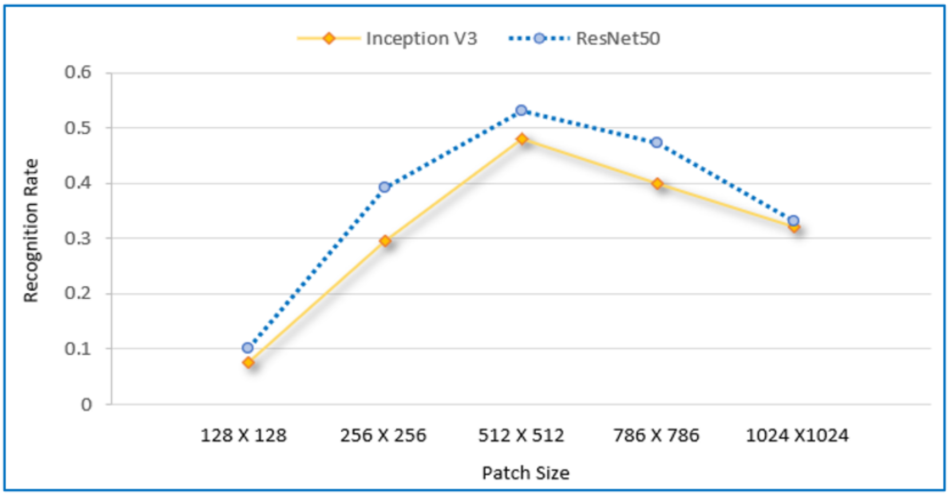


Figure 4.3: Writer identification rates on Patches of different sizes.

In the next series of experiments, rather the using patches of handwriting, we employ small fragments extracted using FAST key points.

Table 4.3: Performance of single and two step fine tuning on different pre-trained ConvNets(Fragments).

	Fine-Tuning Scheme	**Fragment Level**	**Document Level**
VGG16 [80]	ImageNet→Papyri	0.16	0.32
	ImageNet→IAM→Papyri	0.19	0.41
InceptionV3 [77]	ImageNet→Papyri	0.25	0.45
	ImageNet→IAM→Papyri	0.32	0.57
ResNet-50 [78]	ImageNet→Papyri	0.35	0.58
	ImageNet→IAM→Papyri	**0.38**	**0.61**
ResNet-34 [78]	ImageNet→Papyri	0.33	0.60
	ImageNet→IAM→Papyri	**0.40**	**0.645**

Keeping these key points as centers, small patches around them are extracted as small fragments of writing. The fragment and document level

identification rates are summarized in Table 4.3 and illustrated in Figure 4.4. Similar to experiments with random patches, two-step fine-tuning outperforms single-step fine-tuning. The highest reported identification rate reads 64% with ResNet-34. It is also important to note that ResNet34 which is a relatively shallower network as compared to ResNet-50 reports higher identification rates. This indicates that a deeper model is more prone to over-fitting in our case as the dataset is limited.
We also interpreted that indeed fragments facilitated in characterizing the discriminative patterns in handwriting, which also resulted with the increase in the identification rate.

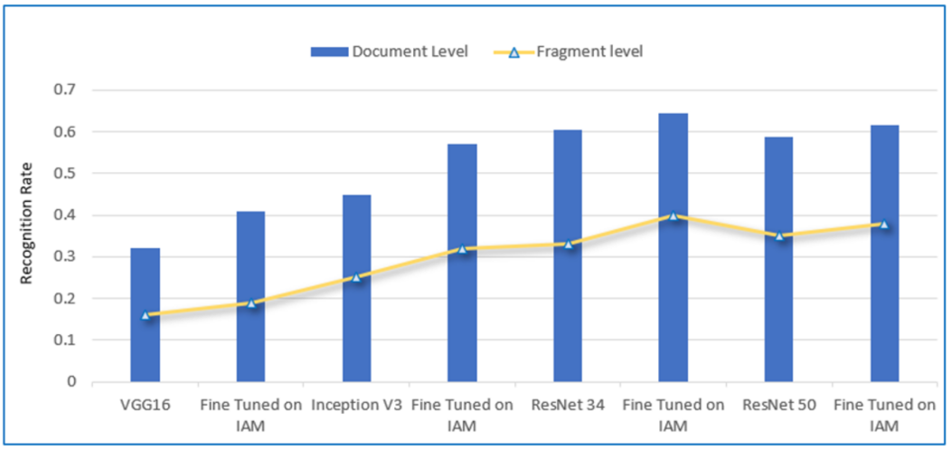


Figure 4.4: Writer identification rates on fragments extracted from FAST key-points.

4.3 Performance Comparison

From the perspective of comparing the approaches, we employed with the identification rates reported on this dataset using Normalized Local Naïve Bayes Nearest-Neighbor with FAST key points in [1].

Authors report an identification rate of 30.0% with leave-one-out protocol

Table 4.4: Comparison with the state-of-the-art.

	Method	***Identification Rate***
[1]	NLNBNN with FAST	30%
Proposed (Dense Sampling)	Two-Step Fine Tuning	54%
Proposed (FAST)	Two-Step Fine Tuning	64.5%

and, 26.6% identification rate with the distribution of data into training and test set. Using the same distribution of 20 images in the training and 30 in the test set, we report an identification rate of 54% using a two-step fine-tuning approach which seems to be quite encouraging. Furthermore, using fragments of writing extracted using FAST key points, the identification rate improves to 64.5% which indeed is very promising.

4.4 Summary

This chapter presented the details of experiments we carried out and the results we achieved with a discussion on our findings. We experimented with a number of well-known, state-of-the-art pre-trained models using two-step fine-tuning. A number of pre-processing techniques were also explored prior to feature extraction and classification. Overall, considering the challenging set of images at hand, the reported identification rates are indeed very encouraging.

CHAPTER 5

CONCLUSION & FUTURE WORK

This study aimed at the identification of scribes from historical manuscripts. More specifically, we investigated the problem of Greek handwriting on papyrus. A major challenge in the analysis of historical documents is to handle the degradations over time. A number of pre-processing techniques were investigated to effectively preserve the writing stroke and eliminate the background noise. Features are extracted from handwriting patches by fine-tuning state-of-the-art ConvNets. A two-step fine-tuning is carried out by first tuning the models to contemporary handwritings and subsequently to the papyri dataset. Patch level identification decisions are combined to document level using a majority voting and, identification rates up to 54% are reported. Patches are extracted using dense sampling of writing with windows as well as by first identifying the handwriting keypoints and then extracting windows around these keypoints. Considering the challenging set of writing samples, the realized identification rates are indeed very promising with a document level identification rate of around 64%.

In our further study on this subject, we intend to extend the analysis to other relevant problems like the classification of writing styles and dating. Furthermore, the current study revealed that pre-processing is a critical step in analyzing such documents and further investigating different pre-processing techniques could indeed be an interesting study. In addition to standard pre-trained models, relatively shallower networks can also be trained from scratch to study the performance evolution. It is expected that the findings of this study would be helpful for the pattern classification community in general and the handwriting analysis community in particular.

REFERENCES

[1] H. Mohammed, I. Marthot-Santaniello, and V. Märgner, "Grk-papyri: A dataset of greek handwriting on papyri for the task of writer identification," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pp. 726–731, 2019.

[2] H. S. Baird, V. Govindaraju, and D. P. Lopresti, "Document analysis systems for digital libraries: Challenges and opportunities," in *International Workshop on Document Analysis Systems*, pp. 1–16, Springer, 2004.

[3] F. Le Bourgeois, E. Trinh, B. Allier, V. Eglin, and H. Emptoz, "Document images analysis solutions for digital libraries," in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, pp. 2–24, IEEE, 2004.

[4] K. P. Sankar, V. Ambati, L. Pratha, and C. Jawahar, "Digitizing a million books: Challenges for document analysis," in *International Workshop on Document Analysis Systems*, pp. 425–436, Springer, 2006.

[5] A. Klemme, "International dunhuang project: The silk road online," *Reference Reviews*, 2014.

[6] T. Van der Zant, L. Schomaker, and K. Haak, "Handwritten-word spotting using biologically inspired features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 1945–1957, 2008.

[7] A. Hamid, M. Bibi, I. Siddiqi, and M. Moetesum, "Historical manuscript dating using textural measures," in *2018 International Conference on Frontiers of Information Technology (FIT)*, pp. 235–240, IEEE, 2018.

[8] A. Hamid, M. Bibi, M. Moetesum, and I. Siddiqi, "Deep learning based approach for historical manuscript dating," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pp. 967–972, 2019.

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

[10] L. Schomaker, "From handwriting analysis to pen-computer applications," *Electronics Communication Engineering Journal*, vol. 10, pp. 93–102, June 1998.

[11] I. Siddiqi, "Classification of handwritten documents: writer recognition," *Université Paris Descartes*, 2009.

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

[13] S. He, P. Samara, J. Burgers, and L. Schomaker, "Image-based historical manuscript dating using contour and stroke fragments," *Pattern Recognition*, vol. 58, pp. 159–171, 2016.

[14] F. Aiolli and A. Ciula, "A case study on the system for paleographic inspections (spi): challenges and new developments.," *Computational Intelligence and Bioengineering*, vol. 196, pp. 53–66, 2009.

[15] H. E. Said, T. N. Tan, and K. D. Baker, "Personal identification based on handwriting," *Pattern Recognition*, vol. 33, no. 1, pp. 149–160, 2000.

[16] Z. He, X. You, and Y. Y. Tang, "Writer identification using global wavelet-based features," *Neurocomputing*, vol. 71, no. 10-12, pp. 1832–1841, 2008.

[17] S. He and L. Schomaker, "Deep adaptive learning for writer identification based on single handwritten word images," *Pattern Recognition*, vol. 88, pp. 64–74, 2019.

[18] F. Aiolli and M. Giollo, "'a study on the writer identification task for paleographic document analysis," in *Proc. 11th IASTED Int'l. Conf. on Artificial Intelligence and Applications—AIA*, 2011.

[19] C. Adak, B. B. Chaudhuri, and M. Blumenstein, "An empirical study on writer identification and verification from intra-variable individual handwriting," *IEEE Access*, vol. 7, pp. 24738–24758, 2019.

[20] L. Schomaker, "From handwriting analysis to pen-computer applications," *Electronics & Communication Engineering Journal*, vol. 10, no. 3, pp. 93–102, 1998.

[21] L. Schomaker and M. Bulacu, "Automatic writer identification using connected-component contours and edge-based features of uppercase western script," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 787–798, 2004.

[22] L. Xing and Y. Qiao, "Deepwriter: A multi-stream deep cnn for text-independent writer identification," in *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 584–589, IEEE, 2016.

[23] A. Rehman, S. Naz, M. I. Razzak, and I. A. Hameed, "Automatic visual features for writer identification: A deep learning approach," *IEEE access*, vol. 7, pp. 17149–17157, 2019.

[24] L. Xing and Y. Qiao, "Deepwriter: A multi-stream deep cnn for text-independent writer identification," in *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 584–589, 2016.

[25] V. Christlein, M. Gropp, S. Fiel, and A. Maier, "Unsupervised feature learning for writer identification and writer retrieval," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 01, pp. 991–997, 2017.

[26] M. Keglevic, S. Fiel, and R. Sablatnig, "Learning features for writer retrieval and identification using triplet cnns," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 211–216, 2018.

[27] S. M. Awaida and S. A. Mahmoud, "State of the art in off-line writer identification of handwritten text and survey of writer identification of arabic text," *Educational Research and Reviews*, vol. 7, no. 20, pp. 445–463, 2012.

[28] G. J. Tan, G. Sulong, and M. S. M. Rahim, "Writer identification: a comparative study across three world major languages," *Forensic science international*, vol. 279, pp. 41–52, 2017.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[30] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.

[31] S. He and L. Schomaker, "Fragnet: Writer identification using deep fragment networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3013–3022, 2020.

[32] B. Kumar, P. Kumar, and A. Sharma, "Rwil: Robust writer identification for indic language," in *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 695–700, 2018.

[33] Y. Tang and X. Wu, "Text-independent writer identification via cnn features and joint bayesian," in *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 566–571, Oct 2016.

[34] R. Nasuno and S. Arai, "Writer identification for offline japanese handwritten character using convolutional neural network," in *Proceedings of the 5th IIAE (Institute of Industrial Applications Engineers) International Conference on Intelligent Systems and Image Processing*, pp. 94–97, 2017.

[35] S. Fiel and R. Sablatnig, "Writer identification and retrieval using a convolutional neural network," in *International Conference on Computer Analysis of Images and Patterns*, pp. 26–37, Springer, 2015.

[36] S. Chen, Y. Wang, C.-T. Lin, W. Ding, and Z. Cao, "Semi-supervised feature learning for improving writer identification," *Information Sciences*, vol. 482, pp. 156–170, 2019.

[37] H. T. Nguyen, C. T. Nguyen, T. Ino, B. Indurkhya, and M. Nakagawa, "Text-independent writer identification using convolutional neural network," *Pattern Recognition Letters*, vol. 121, pp. 104–112, 2019.

[38] V. Christlein, D. Bernecker, A. Maier, and E. Angelopoulou, "Offline writer identification using convolutional neural network activation features," in *German Conference on Pattern Recognition*, pp. 540–552, Springer, 2015.

[39] M. Javidi and M. Jampour, "A deep learning framework for text-independent writer identification," *Engineering Applications of Artificial Intelligence*, vol. 95, p. 103912, 2020.

[40] A. Sulaiman, K. Omar, M. F. Nasrudin, and A. Arram, "Length independent writer identification based on the fusion of deep and hand-crafted descriptors," *IEEE Access*, vol. 7, pp. 91772–91784, 2019.

[41] A. U. Islam, M. J. Khan, K. Khurshid, and F. Shafait, "Hyperspectral image analysis for writer identification using deep learning," in *2019 Digital Image Computing: Techniques and Applications (DICTA)*, pp. 1–7, 2019.

[42] I. Bar-Yosef, I. Beckman, K. Kedem, and I. Dinstein, "Binarization, character extraction, and writer identification of historical hebrew calligraphy documents," *International Journal of Document Analysis and Recognition (IJDAR)*, vol. 9, no. 2-4, pp. 89–99, 2007.

[43] D. Fecker, A. Asit, V. Märgner, J. El-Sana, and T. Fingscheidt, "Writer identification for historical arabic documents," in *2014 22nd International Conference on Pattern Recognition*, pp. 3050–3055, IEEE, 2014.

[44] L. Schomaker, K. Franke, and M. Bulacu, "Using codebooks of fragmented connected-component contours in forensic and historic writer identification," *Pattern Recognition Letters*, vol. 28, no. 6, pp. 719–727, 2007.

[45] N. Cilia, C. De Stefano, F. Fontanella, C. Marrocco, M. Molinara, and A. S. Di Freca, "An end-to-end deep learning system for medieval writer identification," *Pattern Recognition Letters*, vol. 129, pp. 137–143, 2020.

[46] L. Studer, M. Alberti, V. Pondenkandath, P. Goktepe, T. Kolonko, A. Fischer, M. Liwicki, and R. Ingold, "A comprehensive study of ima-

genet pre-training for historical document image analysis," *arXiv preprint arXiv:1905.09113*, 2019.

[47] N. D. Cilia, C. De Stefano, F. Fontanella, C. Marrocco, M. Molinara, and A. S. Di Freca, "A two-step system based on deep transfer learning for writer identification in medieval books," in *International Conference on Computer Analysis of Images and Patterns*, pp. 305–316, Springer, 2019.

[48] H. Mohammed, V. Märgner, and H. S. Stiehl, "Writer identification for historical manuscripts: Analysis and optimisation of a classifier as an easy-to-use tool for scholars from the humanities," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 534–539, 2018.

[49] S. McCann and D. G. Lowe, "Local naive bayes nearest neighbor for image classification," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3650–3656, IEEE, 2012.

[50] P. E. Pagels, "e-codices-virtual manuscript library of switzerland," 2016.

[51] S. Lai, Y. Zhu, and L. Jin, "Encoding pathlet and sift features with bagged vlad for historical writer identification," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3553–3566, 2020.

[52] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[53] T. Lindeberg, "Scale invariant feature transform," 2012.

[54] G. Abdeljalil, C. Djeddi, I. Siddiqi, and S. Al-Maadeed, "Writer identification on historical documents using oriented basic image features," in

2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 369–373, 2018.

[55] M. Chammas, A. Makhoul, and J. Demerjian, "Writer identification for historical handwritten documents using a single feature extraction method," in *19th International Conference on Machine Learning and Applications (ICMLA 2020)*, 2020.

[56] S. Fiel, F. Hollaus, M. Gau, and R. Sablatnig, "Writer identification on historical glagolitic documents," in *Document Recognition and Retrieval XXI*, vol. 9021, p. 902102, International Society for Optics and Photonics, 2014.

[57] Z. Wang, A. Maier, and V. Christlein, "Towards end-to-end deep learning-based writer identification," *INFORMATIK 2020*, 2021.

[58] G. Louloudis, B. Gatos, N. Stamatopoulos, and A. Papandreou, "Icdar 2013 competition on writer identification," in *2013 12th International Conference on Document Analysis and Recognition*, pp. 1397–1401, IEEE, 2013.

[59] S. Fiel, F. Kleber, M. Diem, V. Christlein, G. Louloudis, S. Nikos, and B. Gatos, "Icdar2017 competition on historical document writer identification (historical-wi)," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 1, pp. 1377–1382, IEEE, 2017.

[60] V. Christlein, A. Nicolaou, M. Seuret, D. Stutzmann, and A. Maier, "Icdar 2019 competition on image retrieval for historical handwritten documents," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1505–1509, IEEE, 2019.

[61] F. Simistira, M. Seuret, N. Eichenberger, A. Garz, M. Liwicki, and R. Ingold, "Diva-hisdb: A precisely annotated large dataset of challenging medieval manuscripts," in *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 471–476, IEEE, 2016.

[62] A. Fischer, E. Indermühle, H. Bunke, G. Viehhauser, and M. Stolz, "Ground truth creation for handwriting recognition in historical documents," in *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*, pp. 3–10, 2010.

[63] C. De Stefano, F. Fontanella, M. Maniaci, and A. S. di Freca, "A method for scribe distinction in medieval manuscripts using page layout features," in *International Conference on Image Analysis and Processing*, pp. 393–402, Springer, 2011.

[64] U.-V. Marti and H. Bunke, "The iam-database: an english sentence database for offline handwriting recognition," *International Journal on Document Analysis and Recognition*, vol. 5, no. 1, pp. 39–46, 2002.

[65] F. Cloppet, V. Eglin, D. Stutzmann, N. Vincent, *et al.*, "Icfhr2016 competition on the classification of medieval handwritings in latin script," in *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 590–595, IEEE, 2016.

[66] S. Al Maadeed, W. Ayouby, A. Hassaïne, and J. M. Aljaam, "Quwi: An arabic and english handwriting dataset for offline writer identification," in *2012 International Conference on Frontiers in Handwriting Recognition*, pp. 746–751, IEEE, 2012.

[67] F. Kleber, S. Fiel, M. Diem, and R. Sablatnig, "Cvl-database: An off-line database for writer retrieval, writer identification and word spotting," in *2013 12th International Conference on Document Analysis and Recognition*, pp. 560–564, 2013.

[68] O. Velek and M. Nakagawa, "The impact of large training sets on the recognition rate of off-line japanese kanji character classifiers," in *International Workshop on Document Analysis Systems*, pp. 106–110, Springer, 2002.

[69] L. Schomaker and L. Vuurpijl, *Forensic writer identification: A benchmark data set and a comparison of two systems*. NICI (NIjmegen Institute of Cognitive Information), Katholieke Universiteit ..., 2000.

[70] J. Chiang, C. Hsia, H. Tu, H. T. H. Giang, and T. Lin, "Adaptive image enhancement method for document," in *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pp. 417–420, 2017.

[71] M. A. Dhali, S. He, M. Popović, E. Tigchelaar, and L. Schomaker, "A digital palaeographic approach towards writer identification in the dead sea scrolls," in *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods-Volume 1: ICPRAM*, vol. 2017, pp. 693–702, Scitepress; Setúbal, 2017.

[72] A. Brink, M. Bulacu, and L. Schomaker, "How much handwritten text is needed for text-independent writer verification and identification," in *2008 19th International Conference on Pattern Recognition*, pp. 1–4, 2008.

[73] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

[74] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern recognition*, vol. 33, no. 2, pp. 225–236, 2000.

[75] S. He and L. Schomaker, "Deepotsu: Document enhancement and binarization using iterative deep learning," *Pattern Recognition*, vol. 91, p. 379–390, Jul 2019.

[76] W. Rong, Z. Li, W. Zhang, and L. Sun, "An improved canny edge detection algorithm," in *2014 IEEE International Conference on Mechatronics and Automation*, pp. 577–582, IEEE, 2014.

[77] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[78] S. Targ, D. Almeida, and K. Lyman, "Resnet in resnet: Generalizing residual architectures," *arXiv preprint arXiv:1603.08029*, 2016.

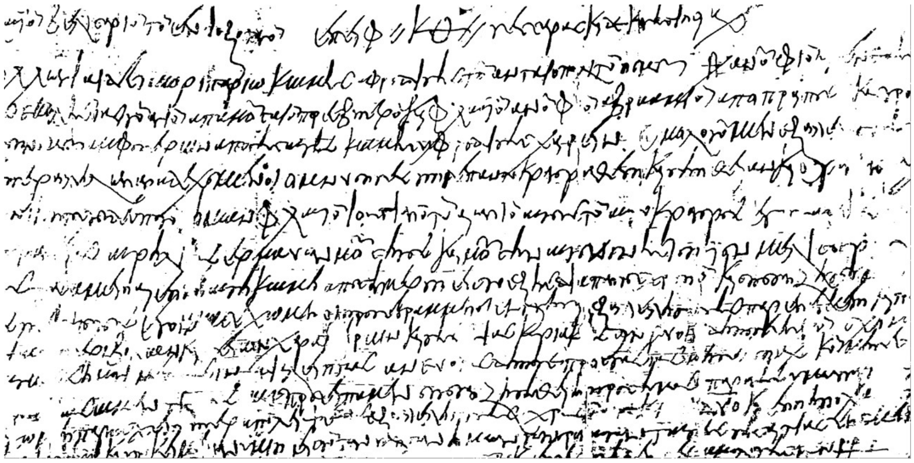
[79] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[80] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

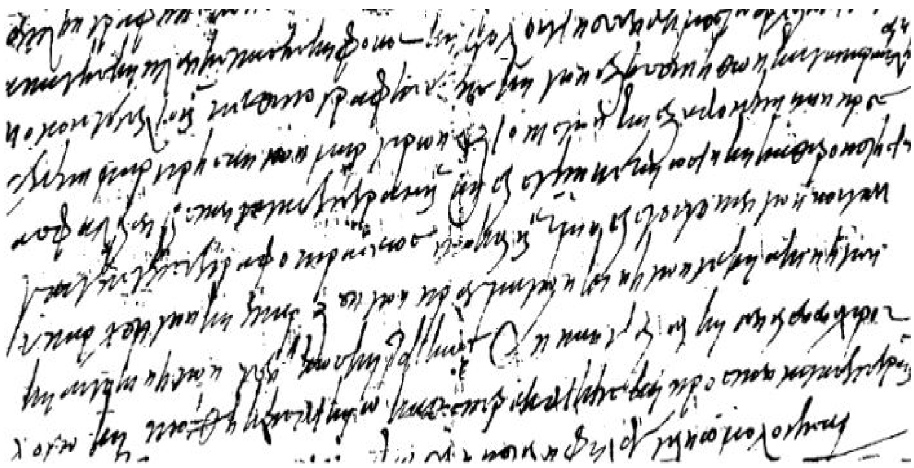
APPENDIX A

Binarized Images

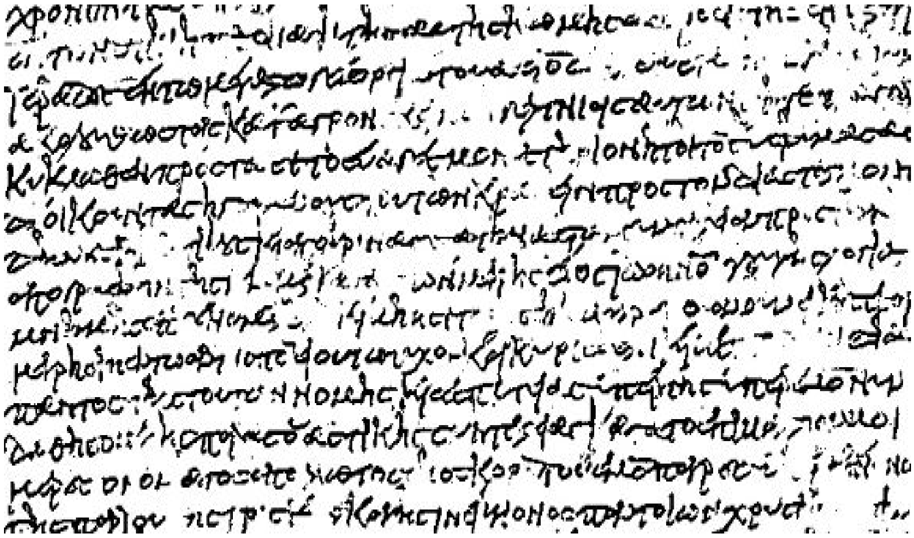
- Abraamios



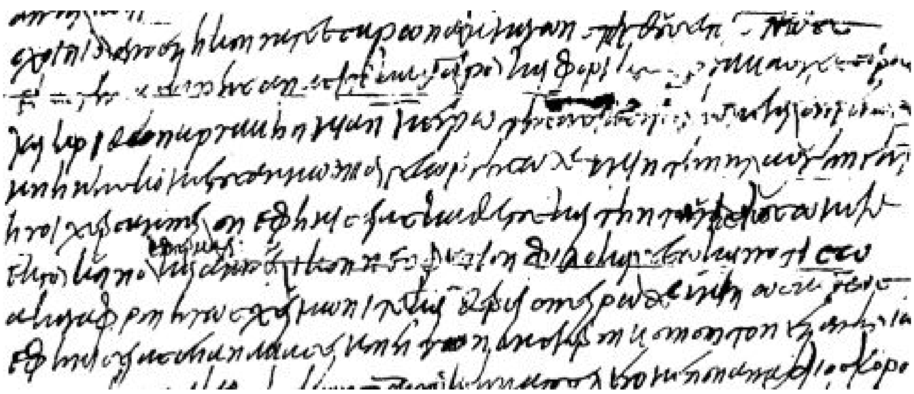
- Andreas



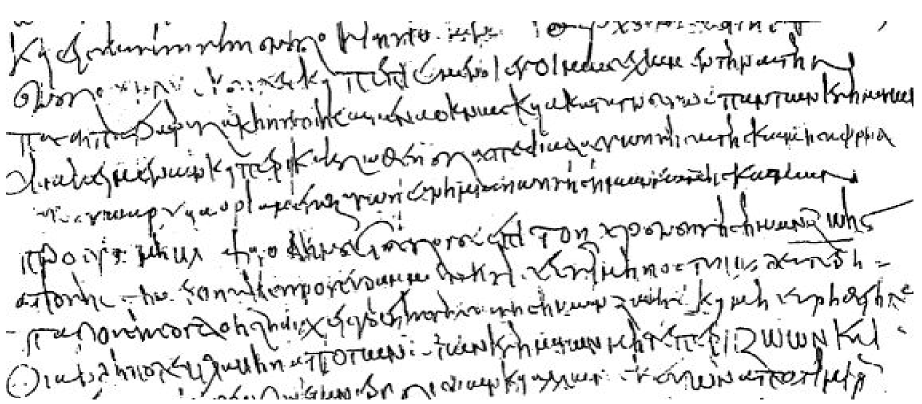
- Dioscorus



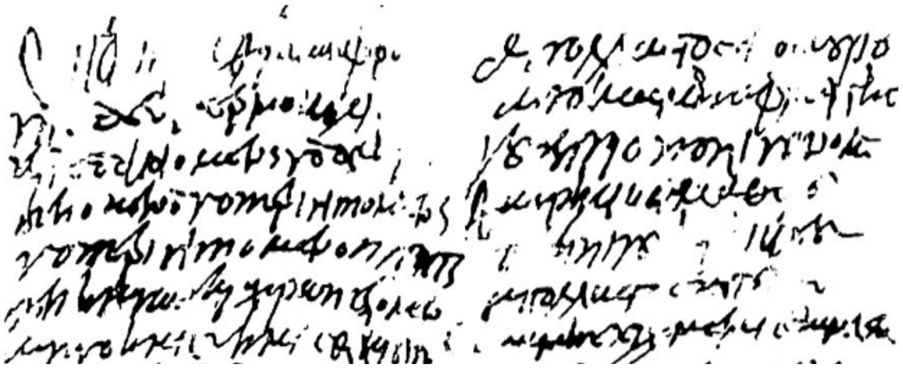
- Hermauos



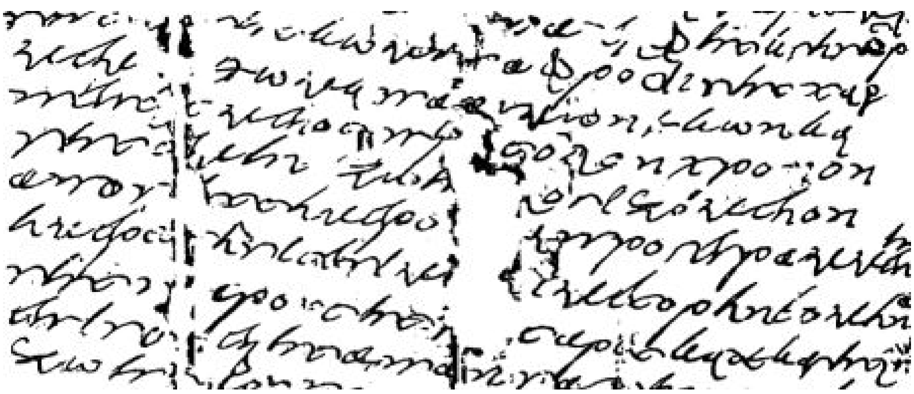
- Isak



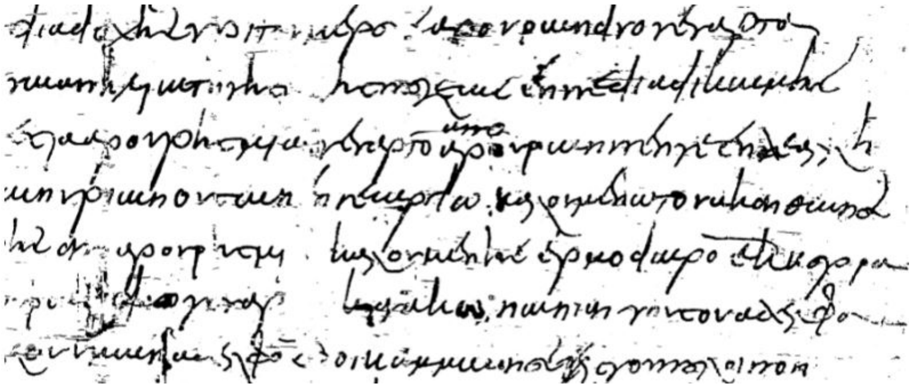
- Kyros1



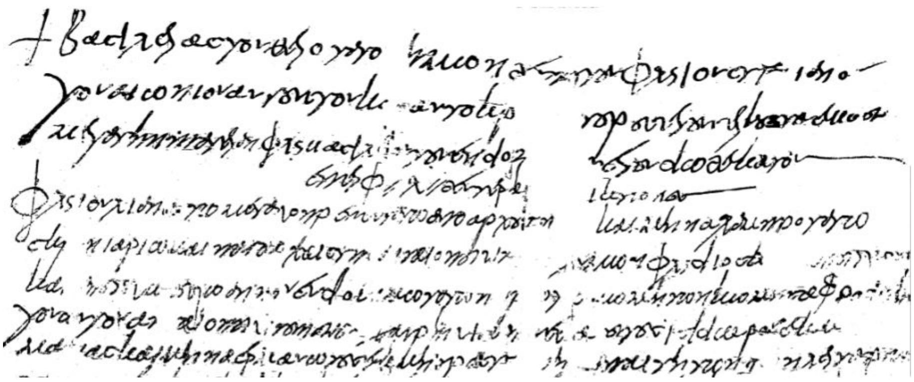
- Kyros3



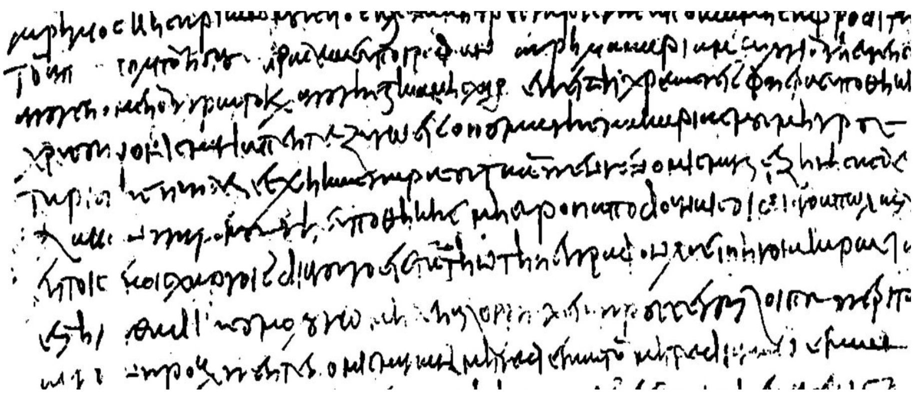
- Menas



- Pilatos



- Victor



turnitin **Turnitin Originality Report**

Scribe Identification by Sidra Nasir

From Papers (Article Similarity Check)

Processed on 01-Mar-2021 13:07 PKT
ID: 1521130968
Word Count: 13213

Similarity Index	Similarity by Source	
17%	Internet Sources:	8%
	Publications:	11%
	Student Papers:	3%

sources:

1 1% match (publications)
Anmol Hamid, Maryam Bibi, Momina Moetesum, Imran Siddiqi. "Deep Learning Based Approach for Historical Manuscript Dating", 2019 International Conference on Document Analysis and Recognition (ICDAR), 2019

2 1% match (publications)
Hussein Mohammed, Isabelle Marthot-Santaniello, Volker Margner. "GRK-Papyri: A Dataset of Greek Handwriting on Papyri for the Task of Writer Identification", 2019 International Conference on Document Analysis and Recognition (ICDAR), 2019

3 < 1% match (student papers from 29-Oct-2020)
Submitted to Higher Education Commission Pakistan on 2020-10-29

4 < 1% match (Internet from 09-Dec-2019)
https://scikit-image.org/docs/dev/auto_examples/segmentation/plot_niblack_sauvola.html

5 < 1% match (Internet from 29-Dec-2018)
https://epdf.tips/image-analysis-and-processing-iciap-2011-proceedings-part-i-lecture-notes-in-com.html

6 < 1% match (publications)
Ammad Ul Islam, Muhammad Jaleed Khan, Khurram Khurshid, Faisal Shafait. "Hyperspectral Image Analysis for Writer Identification using Deep Learning", 2019 Digital Image Computing: Techniques and Applications (DICTA), 2019

7 < 1% match (publications)
Hung Tuan Nguyen, Cuong Tuan Nguyen, Takeya Ino, Bipin Indurkhya, Masaki Nakagawa. "Text-Independent Writer Identification using Convolutional Neural Network", Pattern Recognition Letters, 2018

8 < 1% match (Internet from 01-Nov-2019)
http://www.ijsred.com/volume2/issue4/IJSRED-V2I4P62.pdf

9 < 1% match (publications)
Vincent Christlein, Martin Gropp, Stefan Fiel, Andreas Maier. "Unsupervised Feature Learning for Writer Identification and Writer Retrieval", 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017

10 < 1% match (Internet from 25-Sep-2019)
https://qspace.qu.edu.qa/handle/10576/11934?locale-attribute=ar&show=full

11 < 1% match (publications)
Rapeeporn Chamchong. "Comparing Binarisation Techniques for the Processing of Ancient Manuscripts", IFIP Advances in Information and Communication Technology, 2010

12 < 1% match (Internet from 25-Mar-2020)
https://tel.archives-ouvertes.fr/tel-02514007/document

13 < 1% match (publications)
Gattal Abdeljalil, Chawki Djeddi, Imran Siddiqi, Somaya Al-Maadeed. "Writer Identification on Historical Documents using Oriented Basic Image Features", 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2018

14 < 1% match (Internet from 01-Apr-2017)
http://lvdmaaten.github.io/publications/papers/TR%20New_Dataset_2009.pdf