Nicole Dalia Cilia — Tiziana D'Alessandro — Claudio De Stefano — Francesco Fontanella

Writer Identification from Handwriting on Greek Papyri

1 Introduction

Over time, man has developed a series of well-defined signs capable of fixing an articulated language to transmit it to succeeding generations, thus giving rise to writing. Thanks to it, it is possible to reconstruct an awareness of past events; thus, it represents our most important historical source. The importance of writing in historical reconstruction has spawned a plethora of disciplines devoted to learning, reconstructing, and interpreting these sources: one of these is papyrology.

Studies in the field of papyrology do not focus only on the transcription of every papyrus fragment, but also on other challenging tasks such as digitization,[1] the document enhancing,[2] the writer identification,[3] the recognition of materials used, the definition of the place and the period in which the transcription was executed. This project aims to build an end-to-end system to support researchers during the writer's identification of ancient Greek papyri.[4] In past years, this challenge was faced by different researchers with many approaches.[5] The following paragraphs show the dataset of papyri used and the three different approaches evaluated.

2 Dataset

The reference dataset considered in this work comprises images representing Greek papyri dating back to about the 6th century AD, which have been selected and cataloged by experts in the field of papyrology. All the documents used are part of the richest archive of the Byzantine period, belonging to Dioscorus of Aphrodito, which collects more than

1 See Jayanthi – Indu – Hasija – Tripathi 2022.
2 See Gupta – Kumar – Gupta – Chaudhury – Joshi 2007; He – Schomaker 2019.
3 See Rehman – Naz – Razzak 2019.
4 The project is conducted at the Department of Computer Engineering, University of Enna "Kore" (N. D. Cilia) and at the Department of Electrical and Information Engineering (DIEI), University of Cassino and Southern Lazio (T. D'Alessandro, C. De Stefano, F. Fontanella).
5 See Nasir – Siddiqi 2021; Nasir – Siddiqi – Moetesum 2021; Christlein – Marthot-Santaniello – Mayr *et al.* 2022; Peer – Sablatnig 2023; Cilia – D'Alessandro – De Stefano *et al.* 2024.

https://doi.org/10.1515/9783111070162-020

700 texts written in cursive Greek.[6] The base of this dataset is the GRK-Papyri,[7] which is used to identify the writers and is composed of 50 images distributed unequally among the 10 available writers. Subsequently, other papyri have been added to this starting set, which partly increased each existing writer's number and introduced new authors.[8] Finally, a number of images equal to 122 was reached for a total number of 23 writers.

2.1 Problems

The papyrological transcription and identification tasks are extremely difficult due to the deteriorated condition of most papyrus fragments. Every papyrus is different as each is affected by a certain level of degradation, showing missing pieces, holes, and bending marks. These problems are due to the natural deterioration of the papyrus material or the exposure to harsh natural conditions. The writing may have become illegible, or the ink may have faded with time. Despite being uniquely different in content and composition, they mostly differ because of the deterioration process.

The papyri available in the reference dataset are located and exposed in different parts of the world, representing important documents and part of the artistic and cultural heritage. This means that some images of these papyri may have been acquired with different tools and resolutions and under different light conditions; also, for some of them, it is possible to notice the reflection of the glass that covers them for preservation purposes. Moreover, images have different sizes and are saved with different numbers of color channels; some of them are greyscale, while others are RGB.

Table 1: Writers and the number of papyri that belongs to them.

Writer	no. of Papyri
Abraamios	21
Amais	1
Andreas	4
Anouphis	1
Apa Rhasios	4
Dios	15
Dioscorus	5
Daueit	1

6 See Fournet 2008.

7 See Mohammed – Marthot-Santaniello – Märgner 2019; https://d-scribes.philhist.unibas.ch/en/grk-papyri. All hyperlinks last accessed on 21.4.2024.

8 See Cilia – De Stefano – Fontanella *et al.* 2021.

Writer	no. of Papyri
Hermauos	5
Ieremias	2
Isak	8
Kollouthos	2
Konstantinos	2
Kyros (1)	9
Kyros (2)	1
Kyros (3)	5
Menas	5
Philotheos	3
Pilatos	10
Psates	2
Theodosios	5
Victor (1)	10
Victor (2)	1

3 Experimental procedure

This project aims to build a support system for the writer's identification of ancient Greek papyri. The system relies on Deep Learning (DL) techniques to automatize the process of writer recognition according to handwriting. Because of the problems mentioned in the previous paragraph, a preprocessing step on the images was necessary, trying to uniform them and highlight the handwriting. Following, three different approaches are described. Each one is composed of image preprocessing and a classification step.

3.1 First Approach

Image Pre-processing The image preprocessing step of the first approach tried to solve more than one problem and made the dataset as uniform as possible. The intention was to apply DL techniques on those images, so the dataset needed homogeneity. Also, every papyrus' background had to be as uniform as possible. First, the background uniformity was obtained by filling the holes with the same color as the papyrus background and turning the image greyscale.

Then, images were resized and rotated, if necessary, to extract rows of different lengths. After this step, two datasets were obtained from the initial one:

- a dataset of extracted rows of 1232 pixels width;
- a dataset of extracted rows of 500 pixels width.

Classification The classification step required a well-balanced dataset among the various classes (writers); this led us to arrange the data in several ways according to the number of writers to recognize and the availability of images for each writer. That is why the classification phase was performed on five arranged datasets from the two sets of images discussed in the previous section. We adopted a K-Fold Cross Validation strategy to improve the system's stability. Four Convolutional Neural Networks (CNNs), pre-trained on the public dataset ImageNet,[9] were tested considering the DL techniques of Transfer Learning and Fine Tuning. The chosen CNNs are VGG19,[10] ResNet50,[11] InceptionV3[12] and InceptionResNetV2.[13]

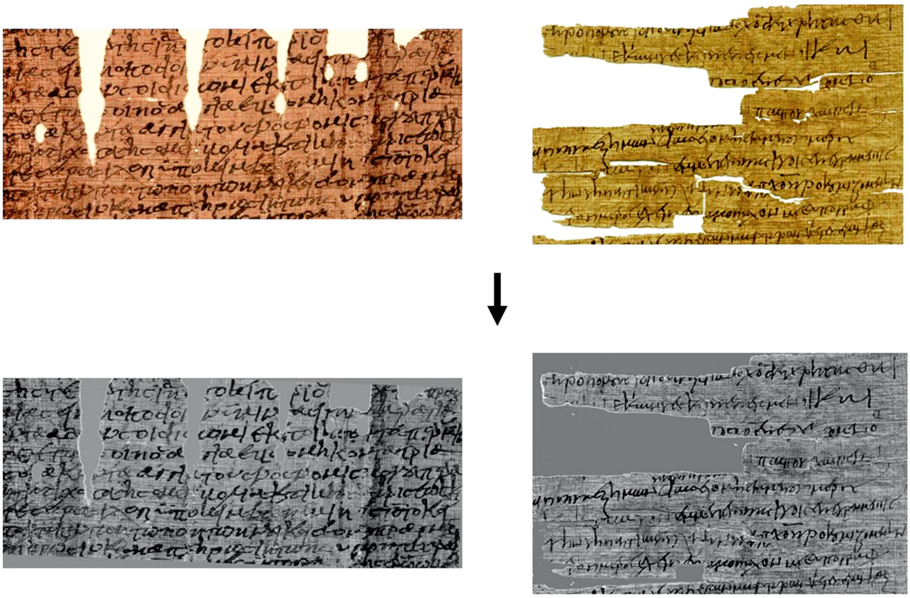

Fig. 1: Processing on original images turned into greyscale with background uniformity.

9 See Deng – Dong – Socher *et al.* 2009.
10 See Simonyan – Zisserman 2015.
11 See He – Zhang – Ren – Sun 2016.
12 See Szegedy – Vanhoucke – Ioffe *et al.* 2016.
13 See Szegedy – Ioffe – Vanhoucke 2016.

Fig. 2: Extraction of rows of different lengths.

Classification on rows. The classification step was performed on the following arranged datasets:

- Dataset 1: 11 classes (Abraamios, Andreas, Dios, Dioscorus, Hermauos, Isak, Kyros 1, Kyros 3, Menas, Pilatos and Victor), 1232 pixels of row width;
- Dataset 2: 2 classes (Abraamios and Dios), 1232 pixels of row width;
- Dataset 3: 2 classes (Abraamios and Dioscorus), 1232 pixels of row width;
- Dataset 4: 2 classes (Abraamios and Dioscorus), 500 pixels of row width;
- Dataset 5: 4 classes (Abraamios, Dioscorus, Pilatos and Victor), 500 pixels of row width.

Dataset 1 obtained the worst performance because the classification task was more challenging, with an increasing number of classes to recognize. The best performance was obtained for dataset 2, particularly with the CNN InceptionResNetV2. This dataset comprised only two classes with a lot of samples for each. Neural Networks are known to work better with a huge and well-balanced dataset. Another interesting trend is that dataset 4 outperformed dataset 3 most of the time. Although they contain the same classes, dataset 4 contains 500 pixels-wide rows instead of 1232 pixels, so the number of images is higher for dataset 4. Dataset 5, with four writers, did not achieve a good performance.

Table 2: Accuracy achieved by classifying on rows.

Model	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
VGG19	35.98	86.65	63.91	80.03	42.75
ResNet50	38.17	97.68	67.68	79.85	44.19
InceptionV3	34.03	97.45	65.86	78.54	46.40
Inc.ResNetV2	42.07	98.35	66.81	48.07	42.55

Classification on Papyri Once the classification result was obtained for every row, it combining the predictions obtained from rows belonging to the same papyrus image was possible. The chosen combination rule was the Majority Vote, so if most of rows from a papyrus were classified as belonging to a certain class, then the entire papyrus was classified as belonging to that class. The application of the combining rule involved only two of the datasets previously considered:

- Dataset 3: 2 classes (Abraamios and Dioscorus), 1232 pixels of row width;
- Dataset 4: 2 classes (Abraamios and Dioscorus), 500 pixels of row width.

Table 3: Accuracy achieved by classifying on rows.

Model	Dataset 3	Dataset 4
VGG19	67.86	46.43
ResNet50	76.79	83.93
InceptionV3	73.21	80.36
Inc.ResNetV2	65.48	88.10

The majority vote rule was applied on images belonging to the same papyrus to obtain a prediction on the entire papyrus instead of single fragments. This rule was applied to the third and fourth datasets, improving the classification result.

3.2 Second Approach

Image Pre-processing The image preprocessing step of the second approach starts from the greyscale images of the entire papyri, obtained during the first approach. We manually detected groups of two and then four consecutive characters from these. This procedure was repeated for every papyrus with the software LabelImg, which returned as output a .xml file containing the information regarding every ROI (Region Of Interest) detected on the image. A Python script received the images and the .xml files as input, returning all the patches containing two and four characters detected with LabelImg.

Classification on groups of two or four characters. As in the previous case, different datasets were generated to obtain balanced sets of two or four characters among the classes. A K-Fold Cross Validation strategy was adopted, and four CNN models, pre-trained on the public dataset ImageNet, were tested considering the DL techniques of Transfer Learning and Fine Tuning. The considered datasets are listed below:

- Dataset 1: 4 characters images, Abraamios and Dios;
- Dataset 2: 2 characters images, Dios vs. All;
- Dataset 3: 2 characters images, Dioscorus and Hermauos;
- Dataset 4: 2 characters images, Abraamios, Dios, Kyros and Victor.

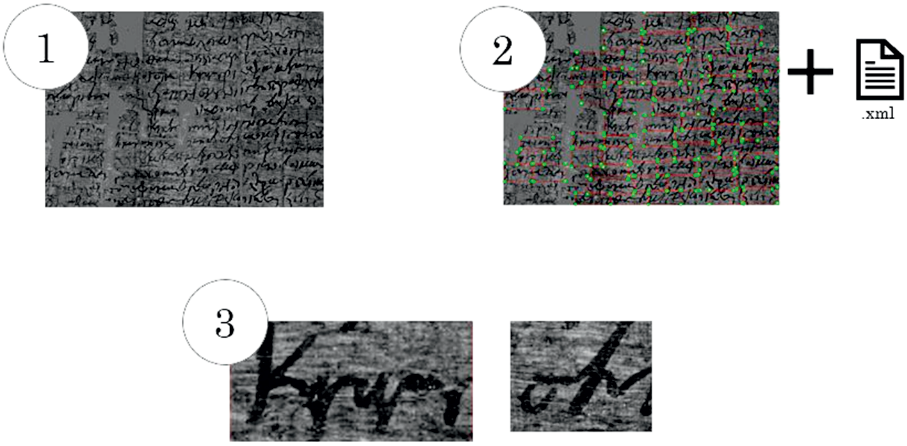

Fig. 3: Extraction of patches of two/four character from greyscale images, through LabelImg.

Table 4: Accuracy achieved by classifying on patches datasets.

Model	Dataset 1	Dataset 2	Dataset 3	Dataset 4
VGG19	59.12	60.36	51.72	52.3
ResNet50	94.58	95.08	73.71	58.8
InceptionV3	97.97	98.27	82.3	69.54
Inc.ResNetV2	98.59	97.69	85.48	63.75

The first dataset obtained the best performance, considering images containing four characters. The second dataset also achieved good results; we noticed that this was a common trend when one of the classes referred to Dios. The dataset that obtained the worst performance was the fourth.

Classification on Papyri. Once the classification result was obtained on every patch, combining the predictions obtained from patches belonging to the same papyrus image was possible. The chosen combination rule was the Majority Vote, applied to the following datasets:
- Dataset 1: majority vote on 4 characters images, Abraamios and Dios;
- Dataset 2: majority vote on 2 characters images, Dios vs. All;
- Dataset 3: majority vote on 2 characters images, Dioscorus and Hermauos;
- Dataset 4: majority vote on 2 characters images, Abraamios, Dios, Kyros and Victor.

Table 5: Accuracy achieved by classifying on patches datasets.

Model	Dataset 1	Dataset 2	Dataset 3	Dataset 4
VGG19	72.22	100	50	40.98
ResNet50	97.22	100	87.5	49.18
InceptionV3	94.44	100	75	59.01
Inc.ResNetV2	97.22	100	75	59.01

In the case of patches, the performance increased for the second dataset when applying a combination rule, but not for the others.

3.3 Third Approach

Image Preprocessing. The image preprocessing step of this approach was divided into two procedures: enhancement and binarization. Document enhancement involves improving the perceptual quality of document images and removing degradations and artifacts from the images. Document binarization separates each pixel belonging to the text from those belonging to the background. The enhancement is also a preprocessing step for binarizing degraded document images to remove unnecessary noise. The enhancing step was done according to the DeepOtsu method proposed by Sheng He and Lambert Schomaker, from the University of Groningen.[14] They built a network to improve input images by removing noise and correcting the degradation. Thus, the neural network's output was the improved version of the input with supervised learning. Many iterations could be performed according to the desired enhancing outcome. After several iterations, the output was the improved input version and could be binarized through classic binarization algorithms such as Otsu. The third approach required extracting patches of two characters from the original images, not the greyscale ones. The extraction process was the same as described for the second approach.

Classification This was a new approach, so the classification involved a dataset of only two writers: Hermauos and Isak. A K-Fold Cross Validation strategy was adopted, and four CNNs, pre-trained on the public dataset ImageNet, were tested considering the DL techniques Transfer Learning and Fine Tuning.

14 He – Schomaker 2019.

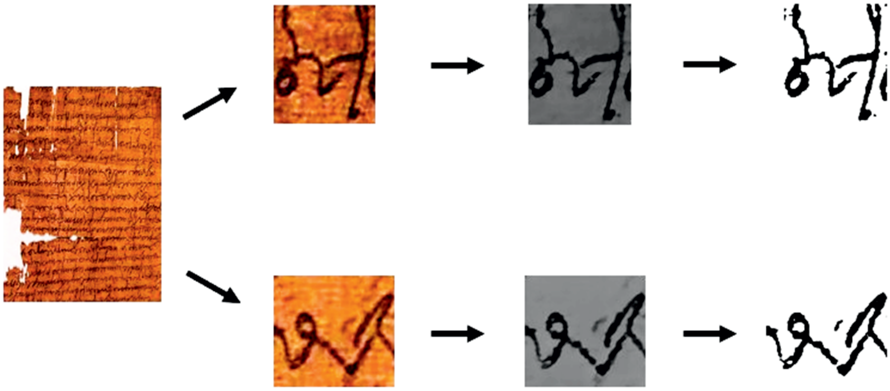

Fig. 4: Extraction of patches from original images, followed by enhancing and binarization.

Table 6: Accuracy achieved by classifying on binarized patches.

Model	Dataset
VGG19	74.42
ResNet50	74.53
InceptionV3	73.72
Inc.ResNetV2	72.56

The best classification result was obtained with the CNN ResNet50, though the results were almost identical for the other networks.

4 Conclusions

The study's primary objective was to develop a classification system based on deep learning to identify the authors of Greek papyri. This undertaking poses significant challenges due to various factors. One challenge is the limited quantity and quality of available data, which is essential for training AI models. Ancient papyri, the focus of this study, often suffer from constraints in data availability, as they do not cover a wide range of writers and stylistic variations. Moreover, preservation issues associated with these historical artifacts make it difficult to extract clear and accurate handwriting samples. Factors such as faded ink, smudges, tears, and other damage to the material can obscure the original script, complicating AI analysis.

However, despite these challenges, AI can be a valuable tool in authorship identification for papyri. Researchers can utilize methodologies involving image processing, pattern recognition, and machine and deep learning algorithms to analyze existing data and make informed determinations. The initial findings from our study are promising and consistent with results from similar investigations. Future efforts will focus on several enhancements. Firstly, there will be a focus on integrating a balanced dataset, even if it requires removing numerous samples. A fine-tuning process will be implemented using a repository of handwritten texts. More binarization strategies will also be explored to eliminate background information and allow the network to concentrate solely on the personality of the writer's handwriting.

Bibliography

Christlein, V. – Marthot-Santaniello, I. – Mayr, M. – Nicolaou, A. – Seuret, M. (2022), *Writer Retrieval and Writer Identification in Greek Papyri*, in *Intertwining Graphonomics with Human Movements*, ed. by C. Carmona-Duarte – M. Diaz – M. A. Ferrer – A. Morales, Cham, 76–89, https://doi.org/10.1007/978-3-031-19745-1_6.

Cilia, N. D. – De Stefano, C. – Fontanella, F. – Marthot-Santaniello, I. – Scotto di Freca, A. (2021), *Papyrow: A Dataset of Row Images from Ancient Greek Papyri for Writers Identification*, in *Pattern Recognition. ICPR International Workshops and Challenges*, ed. by A. Del Bimbo – R. Cucchiara – S. Sclaroff – G. M. Farinella – T. Mei – M. Bertini – H. J. Escalante – R. Vezzani, Cham, 223–34, https://doi.org/10.1007/978-3-030-68787-8_16.

Cilia, N. D. – D'Alessandro, T. – De Stefano, C. – Fontanella, F. – Marthot-Santaniello, I. – Molinara, M. – Scotto Di Freca, A. (2024), *A Novel Writer Identification Approach for Greek Papyri Images*, in *Image Analysis and Processing – ICIAP 2023 Workshops (Udine, September 11–15, 2023) Proceedings*, ed. by A. Del Bimbo – R. Cucchiara – S. Sclaroff – G. M. Farinella – T. Mei – M. Bertini – H. J. Escalante – R. Vezzani, Heidelberg, II, 422–36, https://doi.org/10.1007/978-3-031-51026-7_36.

Deng, J. – Dong, W. – Socher, R. – Li, L.-J. – Li, K. – Fei-Fei, L. (2009), *Imagenet: A Largescale Hierarchical Image Database*, in *2009 IEEE Conference on Computer Vision and Pattern Recognition (Miami, 20–25 June 2009)*, Piscataway, 248–55, https://doi.org/10.1109/CVPR.2009.5206848.

Fontanella, F. – Colace, F. – Molinara, M. – Scotto Di Freca, A. – Stanco, F. (2020), *Pattern Recognition and Artificial Intelligence Techniques for Cultural Heritage*, Pattern Recognition Letters 138, 23–9, https://doi.org/ 10.1016/j.patrec.2020.06.018.

Fournet, J.-L. (2008), ed., *Les archives de Dioscore d'Aphrodité cent ans après leur découverte. Histoire et culture dans l'Egypte byzantine*, Paris.

Gupta, A. – Kumar, S. – Gupta, R. – Chaudhury, S. – Joshi, S. (2007), *Enhancement of Old Manuscript Images*, in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, Piscataway, II, 744–8, https://doi.org/10.1109/ICDAR.2007.4377014.

He, S. – Schomaker, L. (2019), *DeepOtsu: Document Enhancement and Binarization using Iterative Deep Learning*, Pattern Recognition 91, 379–90, https://doi.org/10.1016/j.patcog.2019.01.025.

He, K. – Zhang, X. – Ren, S. – Sun, J. (2016), *Deep Residual Learning for Image Recognition*, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (Las Vegas, 27–30 June 2016)*, Piscataway, 770–8, https://doi.org/10.1109/CVPR.2016.90.

Jayanthi, N. – Indu, S. – Hasija, S. – Tripathi, P.: *Digitization of Ancient Manuscripts and Inscriptions: A Review*, in *Advances in Computing and Data Sciences: First International Conference, ICACDS 2016, Ghaziabad, India, November 11–12, 2016, Revised Selected Papers*, Cham, I, 605–12, https://doi.org/10.1007/978-981-10-5427-3_62.

Mohammed, H. – Marthot-Santaniello, I. – Märgner, V. (2019), *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), Sydney, Australia*, Piscataway, 726–31, https://doi.org/10.1109/ICDAR.2019.00121.

Nasir, S. – Siddiqi, I. (2021), *Learning Features for Writer Identification from Handwriting on Papyri*, in *Pattern Recognition and Artificial Intelligence*, ed. by C. Djeddi – Y. Kessentini – I. Siddiqi – M. Jmaiel, Cham, 229–41, https://doi.org/10.1007/978-3-030-71804-6_17.

Nasir, S. – Siddiqi, I. – Moetesum, M. (2021), *Writer Characterization from Handwriting on Papyri Using Multi-Step Feature Learning*, in *Document Analysis and Recognition – ICDAR 2021 Workshops: Lausanne, Switzerland, September 5–10, 2021, Proceedings*, Berlin – Heidelberg, I, 451–65, https://doi.org/10.1007/978-3-030-86198-8_32.

Peer, M. – Sablatnig, R. (2023), *Feature Mixing for Writer Retrieval and Identification on Papyri Fragments*, in *7th International Workshop on Historical Document Imaging and Processing, San Jose, CA, USA, August 2023*, New York, https://doi.org/10.1145/3604951.3605515.

Rehman, A. – Naz, S. – Razzak, M. I. (2019),: *Writer Identification Using Machine Learning Approaches: A Comprehensive Review*, Multimedia Tools and Applications 78, 10889–931, https://doi.org/10.1007/s11042-018-6577-1.

Simonyan, K. – Zisserman, A. (2015), *Very Deep Convolutional Networks for Large-Scale Image Recognition*, https://arxiv.org/abs/1409.1556.

Szegedy, C. – Ioffe, S. – Vanhoucke, V. (2017), *Inception-v4, Inception-Resnet and the Impact of Residual Connections on Learning*, in *AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (San Francisco, February 4–9, 2017)*, ed. by S. Singh – S. Markovitch, https://dl.acm.org/doi/10.5555/3298023.3298188.

Szegedy, C. – Vanhoucke, V. – Ioffe, S. – Shlens, J. – Wojna, Z. (2016), *Rethinking the Inception Architecture for Computer Vision*, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (Las Vegas, 27–30 June 2016)*, Piscataway, 2818–26, https://arxiv.org/abs/1512.00567.