International Symposium on the Conservation of Monuments in the Mediterranean Basin

(2024)

Proceedings of the 11th MONUBASIN (2024)

Creating a dataset of digital images of Byzantine frescoes for analysis and classification using machine learning techniques

Antigoni Vlisidi, Antonia Moropoulou

doi: 10.12681/monubasin.8183

Creating a dataset of digital images of Byzantine frescoes for analysis and classification using machine learning techniques

Antigoni Vlisidi, NTUA, Athens, Greece
antigonivlisidi@gmail.com
Antonia Moropoulou, NTUA, Athens, Greece

Abstract. The methodology for creating a database of fresco images, mainly of the Macedonian School (13th-14th century), is presented here. This base will be used later for digital image processing for classification using machine learning. For this purpose, modern digital data analysis tools in Python were used. The collection of appropriate material, the digitization of images, the input of information and the categories examined are presented in detail, along with the tools used. One of the main problems of the research was that there were no open digital libraries to cover with images of the specific period examined, so the material gathered comes from corporate sources such as books, studies and collectors. After analyzing the art historical information from Byzantine frescoes not only for their measurability and reliability, but also for availability, a python application that manages the data as an object-oriented relational graph is created. This is a first approach to such material that combines easy-to-use tools with historical knowledge and provides a ready-to-use and development base.

Keywords: Byzantine art, Fresco, Macedonian School, Python, Dataset, Digital image processing

1. Introduction

In recent decades, there has been a leap forward in machine learning techniques in many fields such as medicine, science and art. The development and simplification of systems that manage data, graphics cards and ever-growing computer vision have contributed significantly to this. This study presents in detail the creation of a database consisting of color images of frescoes, mainly from the late Byzantine period, historical observations entered as categorical data and the organization of disparate material structured so as to be useful for further analysis. For this purpose, a system was created with open-source tools in Python that manages the data.

All stages are developed from finding the appropriate material, gathering information, and creating systems that manage files, up to and including the completion of the database that has the possibility of enrichment and development. This database will later be used by us for machine learning for classification. The purpose of this research is to make an approach to the classification of Byzantine wall paintings which we do not know for sure as to which artist created them. It essentially combines art history with computer vision and can be used as an additional tool for scholars.

Firstly, the research began with art historians collecting as much detailed information as possible about Byzantine art and more specifically about the last period of Byzantine fresco painting.

The data we collected is focused on:

- Consistency in the availability of historical data
- Ease of finding information
- Reliability/Objectivity of historical observations
- Possibility of using them as categorical data

Byzantine art from its beginnings (4th century) is a purely religious art, which introduced symbolism and spirituality into its subject matter. Thus, within the bosom of the church, art developed, evolved, matured and reached its peak in the last centuries of the empire (13th to 14th centuries) presenting works of exceptional dynamics and aesthetics. It is no coincidence that the specific period was called the "Palaeologan Renaissance" [1], as not only intellectual people gathered in the two great urban centers of the empire, Constantinople and Thessaloniki, and made them nuclei of philosophical and theological thought, but also, art flourished with now famous artists [2].

There were two main Schools that dominated at that time, the School of Constantinople as the capital of the empire, and the Macedonian School based in Thessaloniki which flourished artistically after the

Proceedings of the 11th International Symposium on the Conservation of Monuments in the Mediterranean Basin, 17-19 of June 2024, Athens, Greece
https://doi.org/10.12681/monubasin.8183

first fall of Constantinople (1204) by the Crusaders [3]. The general characteristics of art during the late Byzantine period can be summarized as follows [4]:

- Mainly mural art as the empire was in decline murals were more economical to make than mosaics made of precious and semi-precious stones
- Multiple compositions: the compositions are enriched with forms that frame the main characters of the works and tell a story
- Figures are not static but possess movement which is distinguished in the grace with which the clothes are drawn or in the limbs of the saints
- The background plays an important role in the performance, in contrast to previous centuries when the saints stood in front of a monochromatic, mostly blue or gold background. Architectural elements are presented which suggest a place, a place, interior or exterior. Central perspective, which was never a requirement of Byzantine artists, is inverted and architecture or objects are presented from multiple viewing angles
- Enrichment of the pictorial program with new themes such as the miracles of Christ, the lives of the saints and the life of the Virgin Mary
- Distortion of forms: lengthening of the faces and limbs -disproportion of the body which conflicts with their striving for standards from antiquity
- Expressiveness in expression in terms of facial expressions, limb movement and body posture
- The known, so far, names of the painters who worked at that time are Michael Astrapas and Eutychios [5], of George Kalliergis [6], of Manuel Panselinos [7], and of Michael Proelefsis, as either signatures and monographs have been found on monuments or historical references and notes in monastery diaries.

But there are a lot of works that we don't know who made them because there is no strong evidence. First, art historians through macroscopic observations have recorded and analyzed the technique, style and painterliness in rendering the works of the above artists and have speculated about the creators of the unsigned.

Then, after a thorough study of Byzantine art as a whole, the selection of the most important monuments of the Paleologan era was made based on the frescoes they contain. Thus, we ended up studying mural ensembles from Thessaloniki, Serbia, Mount Athos and Macedonia.

Analysis methods that take advantage of machine learning algorithms work best with large datasets. So, the goal at this stage is to collect as many photos of the frescoes as possible. Ideally, a digital archive of these works would meet the following criteria:

- there should be no distortion of the murals due to the shooting angle. This could be achieved by orthophotography or photogrammetry techniques [8]
- color fidelity to the original works through
- to be able to calculate the scale of the projects

Here comes the first and very important problem we had to face and that is that there is no open library with photos of Byzantine frescoes as there are correspondingly other periods of art from international organizations, museums, etc.. The private image collectors and research theses studied [9] met some of the basic criteria as they provided some information on image capture, equipment used, lighting, etc., but were incomplete in terms of specific criteria. The Byzantine art books, although rich in images, did not meet the specifications to be used as an exclusive document of the work they depict, as they did not provide scale information or the ability to correct the image's color. The material from the internet was photos from sites mostly of Byzantine churches and religious and historical content, which met specific requirements as they were intended for information. Such material was also not ideal because any attempt, either to control the shooting angle or to correct the color for uniformity

within the data set, could not be carried out. Additionally, scale is completely ignored as the images are presented in various sizes. For this research, material from the internet, private scholars of Byzantine art, research and studies that have been carried out in institutions such as the NTUA and Byzantine art books were used. As none of the sources could provide digitization to use as the primary source for a work, a system that can hold more than one representation of a work came up without considering anyone as the definitive source of truth for the work. With this focus on the quantity of samples per project and modern data augmentation techniques, the present material's deficiency in quality could be compensated for.

1.1 Implementation

In the present study, the following procedure was followed for the pre-processing stage of the data:

- Any processing done has been done exclusively with the derivatives of the original images, which are generated dynamically and not preserved with the original material.

- To achieve maximum flexibility in the selection of tools at each stage of pre-preparation, the application supports pre-preparation chains and temporary storage of the calculations made.

- The 'View' class creates a set of _Photos_ with a chain of edits,
- e.g. 'fp-patch_512'. Hyphens separate routines and underscore parameters. Each process corresponds to a class defined in 'process.py'.
- All intermediate steps are kept in '/derivatives' after their initial creation and are not recalculated unless deleted by the user. After it is created, the 'View' class maintains an 'items' parameter, a dictionary '{row_id: [image list]}', with derivatives for whatever photos we have supplied it for further processing. The class maintains its own Loader and access the data

From the Python libraries where used:

- NumPy manipulation of data in the form of arrays [10]
- Pandas, which allow importing data from various file formats and images are translated into large data tables [11]
- Matplotlib, which allows visual access to massive amounts of data [12]

After gathering 123 color images from Byzantine frescoes, in order to compile and utilize the information material that had been gathered, the entities are organized in a chart as follows:

- Picture: (Photograph), of a project. Where image data and Exif metadata are included, Project reference, ownership and usage rights
- Work: This includes the name, location and possibly an author reference
- Artist: (Creator), where the artist's name (if know) and historical information about his life, are included

In Fig.1. [13] shows an example of entering information divided into categories. This approach was applied to each of the images studied. We examined which fields had values that could be used without processing, such as the title of the Project, the monument it belongs to, the date of its implementation, and the name of the Creator. Some of the fields needed more development, such as the life dates of the painters, which are not known, as there are no records to indicate such facts. For this reason, the specific dates were indicated with the indication "xx" in the last two elements as we knew the century of their activity from the date of creation of their works. In addition, Unknown Artist is the largest number of works we study. So, the specific field was left blank, indicating the lack of information.

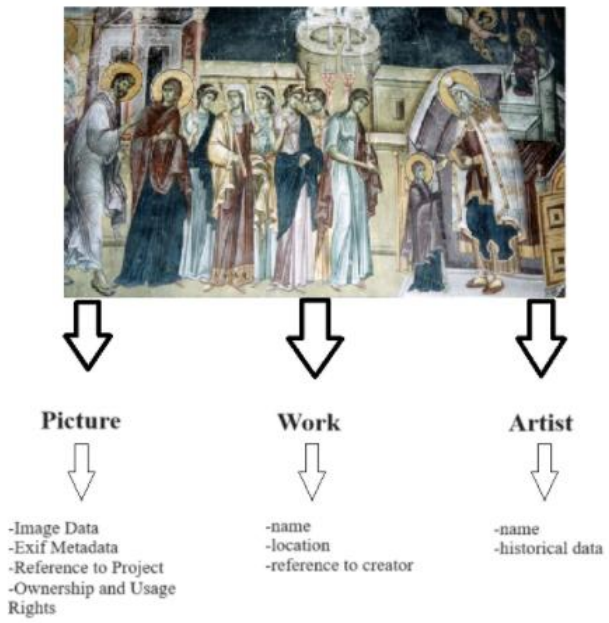

Fig. 1. *: Presentation of Virgin Mary from Zitcha monastery*

Files are stored in a horizontal structure with names derived from the sha256 hexdigest and their format, as detected by Pillow [14]. This process acts as (cryptographically secure) identification and avoids the problem of double importing files. The application manages entities in an object-oriented manner, allowing each metadata entity to be changed and stored independently. The application for data entry was done from a console, where the system asked recursively per image. For each photo, all possible fields are automatically recalled, and all previous values of the relevant field are presented as options.

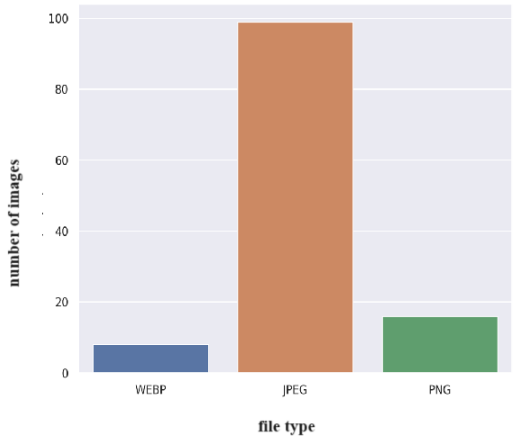

Fig. 2.*: Types of files*

The collected files are shown in detail in the diagrams (Figure 2 and Figure 3) where the unevenness of the information can be seen. The diagram [Figure 2] shows the type of files that were initially collected from the available sources in JPEG, PNG and WEBP formats and, in the diagram in Figure 3 the number of pixels per image used were recorded.

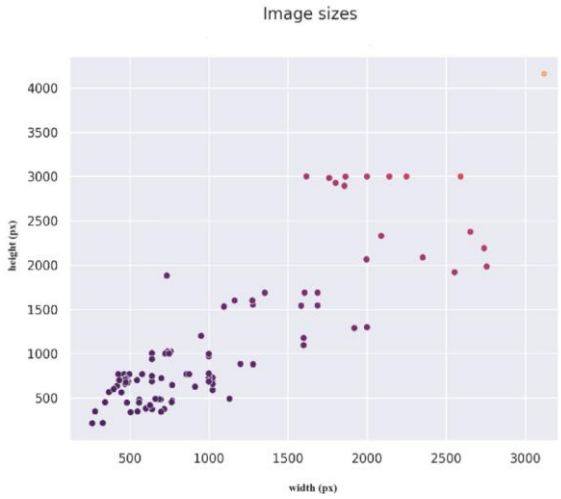

***Fig. 3.**: pixels per image*

.

1.2. Development prospects

This base is the first approach with images that are not in open libraries and their quality varies. So the challenge also lies in whether this kind of material is usable for further analysis. The presented database will be used for machine learning in order to make an approach to image classification and artist recognition. In addition, the dataset can be constantly enriched with new data and developed. As many images as possible can be entered in this way that we presented in the database, the more reliable results we will get. Also, more categories of data such as the geographic location of the project with GIS or CIELab color analysis methods [15], can be useful in extracting results.

Acknowledgments

«This research is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-5000432), implemented by the State Scholarships Foundation (IKY)» The authors would especially like to thank the personnel of the Research Centre for their support and technical cooperation.

References

1. Runciman, St., The last Byzantine Renaissance, Athens Domos, (1991)
2. Panselinou, N., Byzantine painting -Byzantine society and its icons, Kastaniotis publications
3. Radivoj, R., The Fall of Constantinople, 1204, Encyclopedia of Greater Hellenism, Constantinople, (2008)
4. Vlisidi, A., Vesic, N., Doulamis, A. and Moropoulou, A.,(2016) “Determination of criteria for the digital image processing of byzantine fresco paintings”, ISSC2016 Conference The Logics of Image: Visualization, Iconicity, Imagination and Human Creativity, Santorini, Greece, 25-30 July
5. Vesic, N. of Predrag, 2012, "Materials and conservational interventions on the byzantine monasteries of the Balkan painted by Michael and Eutychios Astrapas", PhD thesis supervised by Antonia Moropoulou, School of Chemical Engineering, Department of Materials Science and Technology NTUA
6. Pelecanidis, St., Culture of all Thetalia excellent painter, En Athens Archaeological Society, Athens, (1994)
7. Tsigaridas, E. N., Manuel Panselinos, From the Holy Temple of Protatos, Agioreitiki Estia, (2008)
8. Tokmakidis K., Mapping monuments and archaeological sites. Imprinting by photogrammetry methods., Version: Aristotle University of Thessaloniki, (2014)
9. Kyriakou Ch., Moropoulou A., Characterization of Serbian frescoes of the 14th century and their correlation with those of the Greek area, Workshop Development cooperation between Greece and Serbia for the protection of common cultural heritage -Cultural Paths in Serbian monasteries», Belgrade (2011)
10. Oliphant, T., Guide to NumPy, Open Source, (2007)
11. McKinney W., Pandas: a Foundational Python Library for Data, ResearchGate, (2011)

12. Matplotlib, matplotlib.org, Matplotlib: Visualization with Python: matplotlib.org, (2020)
13. Photo archive N. Vesic
14. https://python-pillow.org
15. Biscontin G., Bakolas A., Longega G., Moropoulou A., Zendri E., Proposta di una metodologia per la valutazione della pullitura di superfici lapidee, in Proc. 3 rd International Symposium on the Conservation of Monuments in the Mediterranean Basin, ed. V. Fassina, H.Ott & F. Zezza, Publ. Sopritendenza ai Beni Artistici e Storici di Venezia, Venice (1994)