A 3D digitisation workflow for architecture-specific annotation of built heritage

Marissia Deligiorgi[a,*], Maria I. Maslioukova[b,*], Melinos Averkiou[b,c], Andreas C. Andreou[b], Pratheba Selvaraju[d], Evangelos Kalogerakis[d], Gustavo Patow[e], Yiorgos Chrysanthou[b,c], George Artopoulos[a]

[a] *The Cyprus Institute, 20 Konstantinou Kavafi Str., 2121, Aglantzia, Nicosia, Cyprus*
[b] *University of Cyprus, 1 Panepistimiou Av., 2109, Aglantzia, Nicosia, Cyprus*
[c] *RISE, 1 Constantinou Paleologou, Tryfon Building, 1011, Nicosia,Cyprus*
[d] *University of Massachusetts Amherst, 140 Governors Dr, Amherst, MA 01002, USA*
[e] *ViRVIG, Universitat de Girona, 17003 Girona*

Abstract

Contemporary discourse points to the central role that heritage plays in the process of enabling groups of various cultural or ethnic background to strengthen their feeling of belonging and sharing in the society. Safeguarding heritage is also valued highly in the priorities of the European Commission. As a result, there have been several long-term initiatives involving digitising, annotating and cataloguing artefacts of tangible cultural heritage in museums and collections. In the case of built heritage, a specific challenge is that historical monuments such as buildings, temples, churches and city fortification infrastructures are hard to document due to historic transformations, destruction, reuse of material, urban development that covers traces and changes the spatial configuration of the site. The ability to reason about a monument's form is crucial for efficient documentation and cataloguing. This paper presents a 3D digitisation workflow through the involvement of reality capture technologies for the annotation and structure analysis of built heritage with the use of 3D Convolutional Neural Networks (3D CNNs) for classification purposes. The present workflow can contribute to the identification of a building's architectural components (e.g., arch, dome) and to the understanding their stylistic influences (e.g., Gothic, Byzantine) can assist in tracking its history, identifying its construction period and comparing it to other buildings of the same period. This process can contribute to educational and research activities, as well as facilitate the automated classification of datasets in digital repositories for scholarly research in digital humanities.

Keywords: convolutional neural network, monument, architectural structure, 3D data, deep learning

*Corresponding authors

Email addresses: m.deligiorgi@cyi.ac.cy (Marissia Deligiorgi), migari01@cs.ucy.ac.cy (Maria I. Maslioukova), m.averkiou@rise.org.cy (Melinos Averkiou), aandre28@cs.ucy.ac.cy (Andreas C. Andreou), pselvaraju@cs.umass.edu (Pratheba Selvaraju), kalo@cs.umass.edu (Evangelos

1. Introduction

One of the most widely used applications of computational methods in human sciences is the automatic annotation and classification of large datasets (of text, images, etc.) in digital libraries that otherwise would require a highly laborious annotation process by trained and skilled users (Canul-Ku et al., 2018; Engel et al., 2019; Dhali et al., 2020). In the case of image-based datasets, computer vision methods have been used to analyse and annotate photos e.g., the geo-reference of series of aerial photos (Cantoro, 2014), the semantic analysis of a digital library of museum artefacts, or a collection of old photographs for the semantic categorisation of contents (Eramian et al., 2017).

Datasets generated by European efforts in digitising cultural content for preservation and online access, in response to European Commission's 2011 recommendations (https://ec.europa.eu/digital-single-market/en/digitisation-digital-preservation), have been growing and have started to include not only 2D information about cultural artefacts of museums and collections, but also 3D models and assets, as in EUROPEANA (https://pro.europeana.eu/project/3d-content-in-europeana), Horizon 2020 projects with relevant APIs, standards, metadata schemata in 3D linked datasets (https://share3d.eu/, https://www.inception-project.eu/en), as well as centralised online repositories for semantically enriched 3D representations of cultural assets (https://sketchfab.com/tags/europeana).

In this paper, we present a 3D digitising workflow specifically designed to assist scholars and researchers in the humanities with the annotation and analysis of cultural heritage monuments. The present work is part of ANNFASS (An Artificial Neural Network Framework for understanding historical monuments Architectural Structure and Style, http://annfass.cs.ucy.ac.cy), a project funded by Research & Innovation Foundation (RIF, https://www.research.org.cy/en/). The RIF Call through which ANNFASS was funded encourages local research and technological innovation, focusing on several areas one of whom is the promotion of the cultural heritage of Cyprus. ANNFASS embraces this goal, aiming to develop an online platform framework for digital humanities, that will integrate 3D CNNs for the classification of elements of Cypriot historical monuments. The Cypriot architecture monuments are rich in complex combinations of structural and stylistic elements, and serve as a challenging test scenario of our platform's abilities and possible restrictions.

2. Related Work

Recently researchers and scholars in the humanities have started to share and exchange 3D datasets of cultural heritage assets beyond the scale of museum and collection artefacts. For example, there have been efforts to create datasets for structures closer to the size of an architectural construction (e.g., monuments), and often whole excavated archaeological sites (Prasomphan and Jung, 2017), with the support of the appropriate computational tools as well as resources for online visualisation. Beyond the management, accessibility, and the "FAIRification" (Harrower et al., 2020) of these

Kalogerakis), gustavo.patow@udg.edu (Gustavo Patow), y.chrysanthou@rise.org.cy (Yiorgos Chrysanthou), g.artopoulos@cyi.ac.cy (George Artopoulos)

unstructured 3D datasets, a major challenge emerges for researchers with regards to their analysis and interpretation in the context of humanities-driven enquiries. The sheer number of digitised cultural heritage assets, their complexity and the sophisticated computational interfaces that are required to be used by humanities scholars for their study, require new tools and techniques that would aid and accelerate research, as well as cross-disciplinary enquiries (Charalambous and Artopoulos, 2018). In addition, these big data can only benefit researchers if they are in the form of linked data (`https://inspire.ec.europa.eu/training/introduction-linked-data`) that can be mined for the identification of patterns, trends and macroscopic mapping of cultural production through time (Fiorucci et al., 2020).

These efforts in digitising the content of study in humanities led to the development of the cross-disciplinary field of Digital Humanities that involves the automatic semantic analysis of visual content. Arguably this process is more complex than solely the annotation of a cultural object's, e.g., a vase or a column provenance and general description (Dallas, 2003). One of the fundamental classification processes in humanities and archaeological research is the periodisation, i.e., to classify artefacts based on historical period (Jiménez-Badillo et al., 2010). This is typically done by experts, who classify artefacts chronologically based on spatial and social context, technique of production, provenance, style and geometric or material features (Baratin et al., 2012). Up until recently it has been very difficult to apply this process in the case of large-scale cultural products, such as buildings or monumental structures, due to their inherent geometry and architectural complexity which resulted in challenges in data management. However, the exponential growth of Machine Learning (ML) algorithms, combined with the large size of 3D datasets being generated in digital humanities and cultural heritage, recently enabled researchers to apply ML techniques in their practice for the interpretation of 3D spatial data of buildings (Grilli and Remondino, 2019) or urban scales (Dirk et al., 2018).

Until recently, most of research in architecture style analysis has been done with 2D data, e.g., images, architectural drawings, or floor plan configurations (Hillier et al., 1987), but advances in neural networks have allowed researchers to handle more complex and bigger datasets, drawing from developments in other fields, cf. 2D image-based retrieval of information by means of Convolutional Neural Networks (CNN) (Llamas et al., 2017). CNNs were first introduced in 1980s LeCun et al. (1989) but were popularised more recently (Krizhevsky et al., 2012) because of their successful application in image and video recognition, recommender systems, image classification, and medical image analysis. CNNs drive inspiration from biological processes in that the connectivity pattern between neurons resembles the organisation of animals' visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. Each neuron or node in a CNN depending on the layer it is located has a different functionality. Input layer nodes, are responsible for receiving data/patterns from the environment (analogous to human senses), and pass them to the next layer to be processed. Finally, the output layer assesses the extracted features to classify the network's input, e.g., whether the input represents a door or a window (da Silva et al., 2017).

In parallel, technological developments — both in spatial 3D documentation equipment as well as in computer vision, e.g., photogrammetry — enabled the acquisition of high resolution and precision 3D data, e.g., point clouds and meshes (Georgopoulos and

Ioannidis, 2004), which naturally capture more information than 2D images of building façades and floor plans, as previously used in architecture history for didactic purposes. These developments and new research opportunities, together with the evolution of deep learning methods in processing 3D data, such as 3D point cloud classification for object identification and semantic annotation (Qi et al., 2017; Wang et al., 2017; Kalogerakis et al., 2017; Su et al., 2018; Bassier et al., 2019; Poux et al., 2017; Messaoudi et al., 2018; Wang et al., 2018; Malinverni et al., 2019; Feng et al., 2020; Pierdicca et al., 2020; Morbidoni et al., 2020), have been the source of inspiration for the authors in applying 3D CNNs, in 3D architectural elements and whole buildings' historical understanding and interpretation, involving annotation and classification processes. To do so the ANNFASS project aims to develop an online platform and software tool that will classify architectural styles and the relevant historical period of built heritage based on 3D analysis instead of 2D image-based analysis (Mathias et al., 2012).

3. Dataset Generation Pipeline

The professional methodologies of measuring and documenting heritage buildings, requires highly developed skills and precision, and is time consuming. However, many of the activities involved in these methodologies, which are human-led, often result in data loss during the transfer of information from the building site into 2D measurements in order to create 3D representations. It is not rare that important details for the conversion would be missed or neglected, as physical observation by an expert is required, especially because of the difficulty of capturing in and representing physical complexity of the building site through 2D drawings (Dallas, 2003).

Typical classification workflows of building elements and architectural style of heritage comprise of the following steps:

- Production of 2D plans and diagrams.
- Literature review for the monument based on cross-referencing literature and onsite findings and,
- Interpretation based on facts and assumptions, as stated by the architect.

A classification workflow involves the analysis of individual architectonic elements of the building in great resolution and detail. Through this analytical process an architect is able to map and interpret singular human interventions on the building structure, or the stylistic influences the owner or its mason drew inspiration from, in order to arrive at the classification of distinct elements.

Nowadays it is common for a built heritage to be documented in 3D through terrestrial laser scanning or photogrammetric techniques for conservation and geometric analysis purposes. This is enabled by advanced software that becomes increasingly accessible to the public as well as by the advancement in computing power available to researchers and professionals in the field. ANNFASS exploits these developments and adapts to current best practices of architects and archaeologists. Thus, it starts the first steps of its methodology by means of 3D documented models of architectural buildings.

ANNFASS' state of the art pipeline involves the use of reality capturing technology, such as 3D scanning technology and photogrammetric techniques, as an exercise to match archaeological and built heritage documentation practices to eventually provide, when completed, a tool that could be used by practitioners, scholars and educators who work with 3D reconstructions of buildings. Apart from the need to create this pipeline based on the current state-of-the-art techniques in 3D reconstruction of complex geometry and materiality of heritage buildings, other reasons for the authors' choice to use 3D reality capture models of historic monuments were to (a) avoid simplified manually modeled representations, (b) and assess the limitations and challenges of developing a tool that in the future would allow the public to upload 3D models of structures documented in the wild by non-experts (recently mobile devices started integrating lidar sensors for 3D capture).

The architectural study of the surveyed monuments was complemented with the theoretical documentation and classification of each building through the analysis of its various characteristics, such as building façade, form, shape, structure, material, colour, openings, ornamentation, roof type, as well as localised factors, including environmental conditions, site topography or cultural aspects.

Stathopoulou and Remondino (2019) trained CNNs on 3D datasets of historic building façades in Italian cities created with a photogrammetry-based documentation pipeline, for the generation of labelled 3D datasets of heritage buildings. Contributing to this inquiry, the ongoing research of ANNFASS explores the use of 3D CNNs trained on 3D point cloud datasets generated by means of photogrammetry, and aims to develop an integrated interface for researchers and scholars in digital humanities for the guided annotation and classification (in terms of architecture style, typology and period) of heritage buildings, in three dimensions. Using historic buildings and monuments of Cypriot architecture, which often are located outside a continuous urban fabric, and therefore have more than two elevations in addition to a courtyard, to train the 3D CNNs, is helping the authors to develop a flexible integrated tool for the labelling and classification of all the 3D mass of a building structure, in addition to its façade.

The authors expect that, when the project is completed and the 3D CNNs will be trained with a large enough 3D building dataset, the ANNFASS tool will enable for a more accurate classification of the building typology, based not only on architectonic features but also on whole building components (e.g., its roof). The authors expect that the capacity of the tool for analysing the building as a whole, combined with the annotation of the style and period of the architecture features and decoration of the façade, will contribute more information to the architect user of the tool than other methods. The 3D CNNs of ANNFASS were trained using architecture 3D datasets (labelled dataset of 2000 buildings, of different types, an ongoing effort at the University of Massachusetts) and is currently further used in labelling the many architectural features of Cypriot heritage, listed below.

Below partial results of the conducted research in the context of the ANNFASS project will be presented. Particularly, the theoretical considerations and practical challenges related to the development of a 3D documentation workflow customised specifically to the needs of reconstructing objects of architectural scale, involving 3D mesh adaptation (3.1), as well as the parameters of the annotation processes (3.2). These two steps in the process are essential in generating the necessary input that can be used to train the 3D CNNs of the project.

The buildings were chosen because of their historical and cultural significance, with the aim to cover as many historical architectural periods as possible, and are located in Nicosia, as the authors wanted to limit the study of exemplar cases to a historically complex location where significant hybridisation / exchanges of stylistic features were identified. It is worth mentioning that the historic city of Nicosia is ideal for collecting data from various time periods, as it has a history of thousands of years and has been at the crossroads of many empires and civilisations (Michaelides, 2012). A summary of the selected buildings per period/architectural style can be seen in Table 1. In order to test the developed workflow with the particularities of more contemporary architecture comprised of different geometric characteristics, the authors included in the dataset examples of built heritage of the recent past and, in particular, select cases of Modernist architecture. The following section explains in detail the challenges of annotating the selected examples in the list below, and presents the results of the aforementioned workflow.

3.1. Building 3D documentation

The interactive visualisation of these heritage buildings and monuments on the ANNFASS interface extends typical vector-based and orthophotographic visualisation practices. This interface allows the user of the tool to have a photo-realistic 3D representation of the structure that enables both expert and non-expert audiences to study and compare architectonic elements, stylistic variations, and their integration in the form of the building. The steps of the hybrid 3D documentation process, that involves terrestrial photogrammetry methods and manual 3D modeling, are described below. An integrated photogrammetry method is defined as a hardware/software configuration that produces photogrammetric products from digital imagery using manual and automatic techniques. Close range photogrammetry is a measurement technology that can be used for the extraction of 3D points from the images, and by extension, these points are useful for the accurate 3D modeling and visualisation. Digital photogrammetry derives all the appropriate measurements from the images themselves rather than taking measurements directly from the objects (Lachambre et al., 2017).

The architectural structure 3D documentation process involved the following steps:

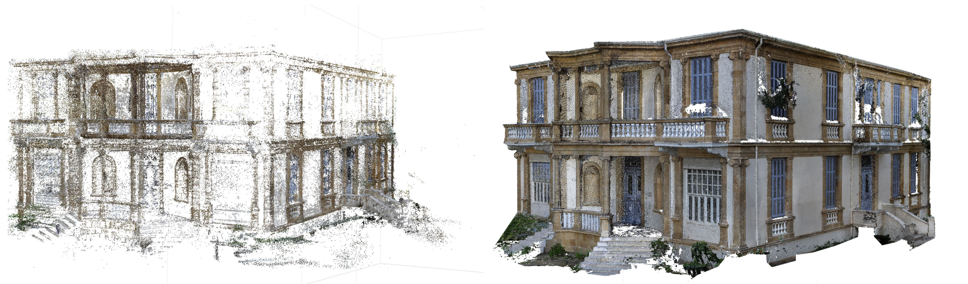
On-site collection of data

The result of this procedure was the complete dataset of 3D models of every building in the list, in terms of geometry and texture. In this first step, the aim was to gather multiple photos of the monument from various angles, along with the basic external measurements. In some cases, it was necessary to have a more detailed geometry and texture, in combination with the main geometry, at different elements of a monument such as doorways, ornaments, pilasters. For example, an ornament's texture which covers a large area of a model might be enhanced by adding a higher resolution detail texture at a much smaller scale which shows small details and imperfections in the ornament. The documentation process of the geometry and texture of architectural details was the same followed for regular, plain building surfaces, except that shots were taken more closely to the building, focusing on a small area. However the part of the building that was documented in order to produce the texture had to be as generic as possible, as detail textures are typically tiled many times across an object. In practice, these shots were taken at the closest focus limit of the camera and lens combination used in the documentation.

Lusignan - Gothic Architecture	
	Armenian Church
	Cathedral of Our Lady Hodegetria
	Augustinian Hermits
	St. Catherine's Church
Ottoman Architecture	
	Buyuk Han
	Bayraktar Mosque
	Ayios Michael Tripiotis
	Axiothea House
	Hadjigeorgakis Kornesios Mansion
Venetian Architecture	
	Famagusta gate
	Kyrenia Gate
	Stavros tou Misericou
British - Colonial architecture	
	Pafos Gate
	English school
	The club of the British Cavalry
Neo-classicism - Greek-revival architectural style	
	Faneromeni Church
	Parthenagogio of Faneromeni
	Pancyprian Gymnasium in Nicosia
	The Severios Library
	Archaeological Research Unit (ARU) Building
	Cyprus Archaeological Museum portico
Vernacular construction methods	
	Townhouses in Ayioi Omologites
	Townhouses in Strovolos
Modernist architecture principles	
	Stavrou Economou Building
	Lefkaritis Building
	Nicolaou Building

Table 1: List of selected monuments, categorised by period/architectural style.

Lastly, due to the lack of access to private properties, in some cases there were parts of the building that were covered from plants or civic equipment, and also due to location of the buildings in the densely populated historic core of Nicosia the aerial recording was at the time not an option for documentation. These restrictions resulted in the photogrammetric process producing partial 3D models of the heritage buildings in the list rather than a complete documentation of them in terms of geometry and texture. (fig. 1).



(a) ARU Building, Alignment of the collected data (b) ARU Building, Dense 3D Point cloud

Figure 1: Examples of processing the collected data.

In technical details, the survey equipment used consisted of a Canon EOS6D, self-calibration was used, and a tripod for image stabilisation (see Table 2).

Property	**Value**
Dimensions	5472×3648
ISO speed	ISO-100
F-stop	f/7.1
Exposure time	1/250 sec.
Flash mode	No Flash
Focal Length	35 mm
Output	.CR2(raw) & .JPEG

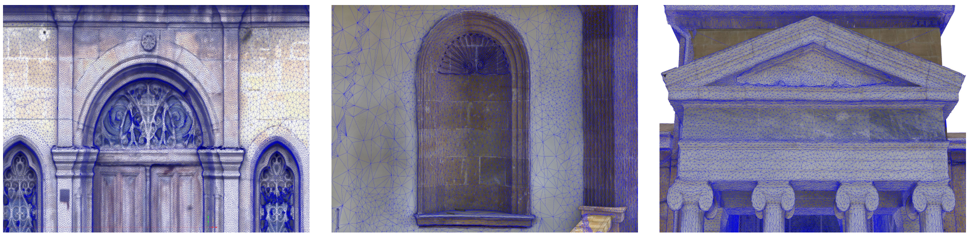
Table 2: Rectified photography properties

Generation of dense 3D point-clouds

The next step of the documentation process was the generation of dense 3D point clouds. The photographs acquired were imported into photogrammetric modeling software, *Agisoft Metashape*. Reference points were selected on the monument, e.g., the corners of the building which are easily identifiable and separable. This process is known as orientation. Furthermore, the reconstruction application compares the shapes in the photos (Alignment) to generate a high resolution 3D point could and by extension the mesh model. The color contained in the pictures is then transferred to either the point cloud and mesh vertex colors (Colourise) or to textures used on the surface of the mesh.

The photogrammetric software generated an extremely high-resolution point cloud and mesh model. This was not suitable for use in 3D modeling software and the online annotation tool because rendering, processing and interaction with its geometry were prohibitively slow. We decimated the original models with photogrammetric software (see Table 3). We observed that the important architectural details for the classification of the building from the high resolution 3D mesh were preserved. The texture transfer from the original mesh model to the decimated one was done with a texture baking process.

In detail, texture baking generally refers to the process of recording as an image, some aspect of the texture or mesh characteristics of a model. The baking tool starts with a low-resolution model and casts rays inwards towards to the high-resolution mesh model. When a ray intersects the high-resolution mesh model, it records the surface detail and saves that into a texture map, using the first model's Texture Coordinates. In other words, in texture baking, what is originally a procedural texture can be recorded as an image. Sometimes various "channels" of a material can be consolidated into a single image, simplifying the number of texture images used. In normal baking, the mesh normals can be recorded – this results in very specialised images with RGB values based on normal vectors. Usually baking requires having the model UV-unwrapped and -mapped, so the resulting image is properly fit to the model (fig. 2).



(a) Townhouses in Ayioi Omologites, Detailed mesh model of the fanlight.

(b) ARU, Detailed mesh model of entrance ornamentation.

(c) Cyprus Archaeological Museum portico, Detailed mesh model of the entrance.

Figure 2: Examples of monuments' detailed geometry and texture.

Due to the fact that the mesh quality and resolution may vary significantly depending on the algorithms involved during the decimation, this process required several attempts, in order to decimate the generated 3D meshes appropriately, so as to maintain the geometric features and architectural details of the building (and the original 3D model), but at the same time to produce a 3D model that would meet the constraints and specifications of the annotation digital platform (fig. 3). The main constraints were computer memory and processing time, which resulted in a trade off between the model resolution/detail and loading time (proportional to model size). This led to the proposed methodology, creating hybrid models with enough detail to resemble closely the original and reducing memory consumption (especially for mobile devices) and loading/rendering time to the minimum.

Table 3 lists the details of the original photogrammetric models of some of the buildings in the list, the points of which range from 422,123,606 points, in the case of the Archaeological Research Unit Building, to 77,773,370 points for the townhouses in Ayioi Omologites, while the poly-counts that consisted the mesh models range from 158,554,674 (Townhouses in Ayioi Omologites) to 20,057,150 (Severios Library). After the decima-


(a) Townhouses in Ayioi Omologites


(b) Townhouses in Ayioi Omologites, High poly-count 3D mesh model, 15.000.000 triangles


(c) Townhouses in Ayioi Omologites, Medium poly-count 3D mesh model, 6.000.000 triangles


(d) Townhouses in Ayioi Omologites, Low poly-count 3D mesh model, 800.000 triangles

Figure 3: Mesh model decimation

tion of the original 3D mesh models, the final low poly-count model of the monuments was limited to a range between 441,804 (for the Townhouses in Ayioi Omologites) and 1,555,196 (English School). In detail, the process of the mesh model decimation was done in the photogrammetric modeling software *Agisoft Metashape*, the re-meshing and the mesh smoothing steps were done with the use of the Autodesk's *Mesh Mixer* open source software, while the final processing and synthesis of all 3D models was done with the use of Autodesk's *Maya* design software.

Monuments 3D documentation				
	Cameras	**Point Cloud**	**Mesh model**	**Final Model**
Colonial / Hybrid architectural style				
English school	596 / 650	440,971,995	88,209,900	1,555,196
Neo-classicism / Greek-revival architectural style				
Severios Library	333 / 357	100,285,759	20,057,150	1,293,236
ARU Building	915/ 930	422,123,606	101,435,980	1,427,309
Cyprus Archaeological Museum	486 / 487	108,395,083	21,679,015	1,089,696
Vernacular construction methods				
Townhouses in Ayioi Omologites	242 / 255	77,773,370	158,554,674	441,804
Townhouses in Strovolos	234 / 234	256,645,209	51,329,037	1,248,625

Table 3: Statistics of Modeling Procedure

Post-processing of the produced 3D mesh and manual 3D modellings

All the above restrictions in the resolution of the 3D models forced the authors to create a hybrid method (partly automated and semi-supervised) for the production of the monument 3D models. This process involved the creation of a 3D model that combined high mesh quality of the period relevant and style-characteristic architectural features (e.g., windows, doors, pilasters), with a reduced resolution of the mesh of the not so distinct elements (e.g., walls, roof, etc.), some of which were even modelled manually (fig. 4).

Figure 4: Hybrid 3D model of townhouses in Ayioi Omologites.

Segmentation of the 3D model into elements

The last step of the modelling process proved more challenging than the original estimations, due to the required segmentation of each building's 3D model into its elements, in order for the user of the ANNFASS annotation tool to be able to easily annotate the architectural parts. In order to respond to this requirement, the process had to be developed as follows: as soon as the final 3D model of a building was produced, a series of segmentations on the generated final surfaces needed to be created, be named, and grouped together with all the relevant architectural elements (e.g., other windows), that is, semantically structuring the elements of the building. The goal of this stage was to speed up the selection of elements from the user, who can select on the (ANNFASS) tool the command `<EXPAND>` to be allowed to add into the selection new elements of the same architectural characteristics as the already selected elements. The following paragraphs present details related to the factors that were integrated in the tool for the function of monument classification (fig. 5).

(a) Severios Library

(b) Archaeological Research Unit Building

(c) Cyprus Archaeological Museum portico

(d) Townhouses in Ayioi Omologites

Figure 5: 3D mesh models of selected historical monuments

3.2. Building classification and stylistic definition factors

In ANNFASS, the period and architecture style classification is based on a building's main architectural parts (e.g., tower, roof type, courtyard, openings), as well as the architectonic features on its elevations. "Architectonic" is used here to refer to architecture, formal and structural aspects of a building feature. The building elevations are divided into two sets of features, i.e., structural parts and decoration. The selected heritage buildings were analysed and compared according to the indicators of their elevations, such as:

- Form and shape: in terms of main form, geometric or irregular and dimension;
- Building structure: in terms of type of structure (skeletal frame, load bearing), type of column, material and shape of brackets and material of balustrade;
- Building Material: in terms of usage, color, function;
- Openings: in terms of size, shape, position on the building façade;
- Door, Entrance, Window: in terms of height, width, position;

- Balcony / Bay window / Semi open space: in terms of location, function, size;
- Ornamentation: in terms of period, style, date; and,
- Proportions and plan space configuration.

The methodology of building classification included the factor of visual analysis, such as observation and 3D documentation, as well as the study of literature of historical sources, in order to understand all the elements and the details that compose the selected monuments. The main aim was to analyse the design of the buildings in terms of morphological characteristics. In this context, the buildings in the list were selected as exemplar cases of historic architecture in Cyprus to train the algorithms (as briefly discussed in section 4.2 below) and give the opportunity to users of the ANNFASS platform to make a comparative assessment between different buildings of various historical periods of Cypriot architecture.

The list of architectural components and stylistic factors, along with the corresponding labels used for the building annotation purposes can be found in Table 4.

Architectural Components		
Arch bay	Balcony	Bay window
Beam	Bell tower	Canopy
Column	Chimney	Door
Doorway	Fanlight	Keystone
Minaret	Ornamentation	Pilaster
Railing	Roof	

Table 4: List of labels used to name the architectural components.

The selected heritage buildings was a set of pilot cases to assess the performance of algorithms through the use of Cypriot monuments.They presented probably some of the most challenging examples of architecture to classify. This is due to the irregular and variable combination of decoration and features borrowed by their builders from multiple historical periods. This difficulty was tackled by dividing them in separate objects per architectural part and engaging groups of experts to guide the classification process, as presented below. Thus the classification would regard individual building parts and not the whole monument. The authors' motivation was to use these difficult examples in order to showcase the value of the ANNFASS tool in educational environments.

4. Features of the ANNFASS Framework

In the previous section (section 3 Dataset), the procedure of creating the hybrid 3D models of monuments was described (fig. 6), here the functionalities and purpose of the annotation tool and platform will be discussed as the next step in the workflow presented by the article.

It is worth mentioning that prior to the development of the annotation tool and ANNFASS platform, the authors sought for digital tools developed for similar scope and

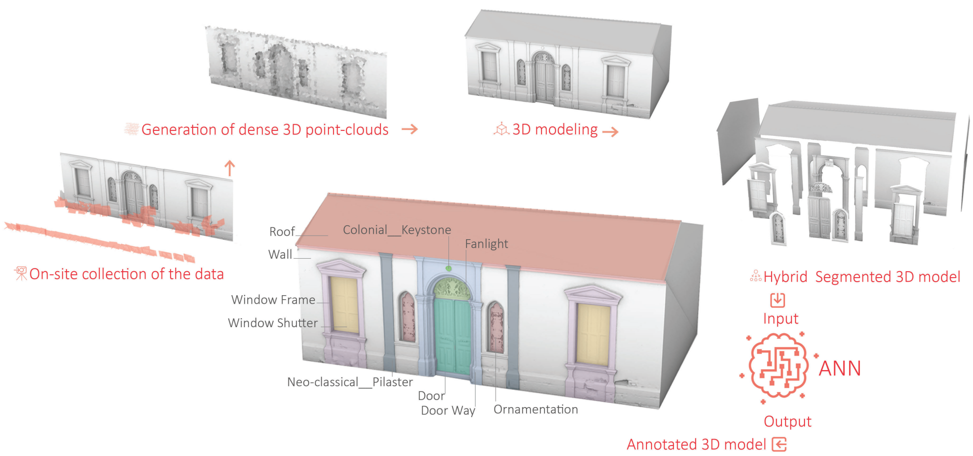


Figure 6: 3D digitisation workflow for architecture-specific annotation of built heritage.

research purpose aimed to be used in digital humanities by scholars and cultural heritage experts in comparative, analytical studies. As it was difficult to identify tools that offer similar functionalities to ANNFASS', the authors had to assess the needs of experts through primary research, in this case participatory methods such as focus groups and workshops. In doing so, several meetings with a group of history of architecture experts were held, introducing the group of stakeholders to the objectives of the ANNFASS, and discussing the needs, challenges, and particularities of their research. Those meetings were necessary and helpful for the development of a platform tailored to their requirements.

4.1. Annotation Tool

In these meetings, the architecture experts had the chance to use an annotation tool inspired by the annotation tool of the BuildingNet platform Selvaraju et al. (2020), and further developed based on the objectives of ANNFASS. The meeting consisted of a short demonstration of the tool's fundamental functionalities, followed by an annotation session where the experts were asked to annotate structural components (wall, door, etc.) of monuments. This process not only offered to the target user group a hands-on experience of the tool, but also allowed the authors to collect valid labels for their dataset of monuments and buildings. Feedback from the experts was collected by means of a previously validated questionnaire that captured input specifically on the user interface of the online ANNFASS tool, accessibility, friendliness and usefulness/purpose of the tool for specialists.

In detail, the user experience involves the following steps: once a user accesses the annotation tool, a monument is loaded in two visualisation forms, the 3D model (stripped from any textures) and its textured twin (fig. 7). The model to be annotated appears on the left side of the screen, but since sometimes architectural features can not easily be distinguished in the plain model, the textured version of the 3D model is also presented next to it in order to enhance the user's perception and allow for detailed observation. Through this interface a building 3D model can be observed by rotating it 360° in any direction and zooming in/out to provide to the user a better look at the different elements of the building.

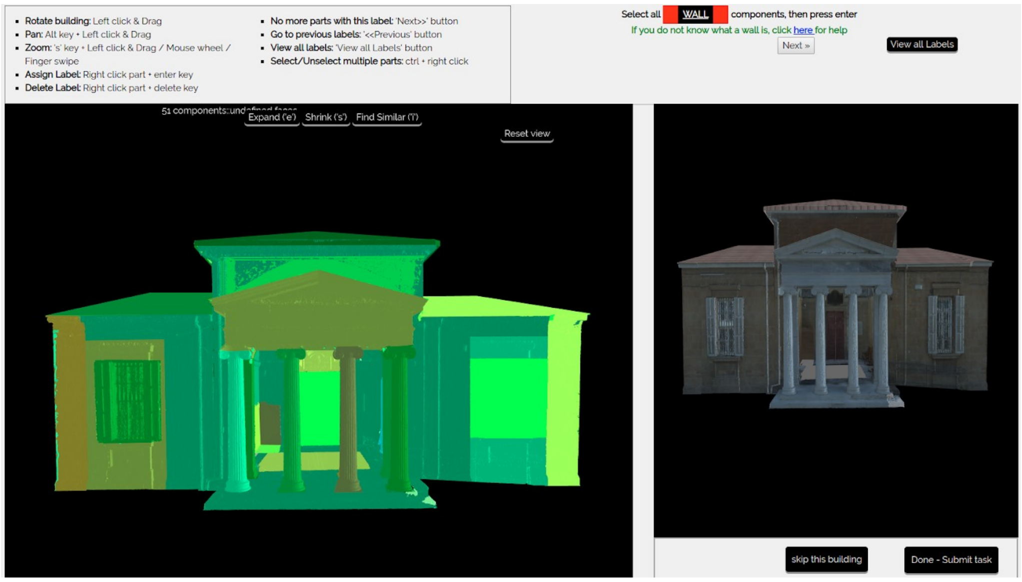


Figure 7: 3D model in annotation tool. Geometry of a the model (left) and textured monument (right).

For the annotation process to start, the user must click on the component of interest and assign it a label. A label can be chosen in one of the following ways: by navigating to the desired label on the upper right corner of the screen or pressing the `<View all labels>` button, for the entire list of labels to appear between the two building model views, and selecting the desired one. The authors are well aware that an architectural element might appear in a variety of configurations and designations, thus to eliminate confusions and mislabellings, a set of example images (fig. 8) for each label is provided by clicking on the link below the label in question. The extent and specificity of the labels is limited to the most frequently found architectural elements, in order to cover the majority of buildings on the platform while avoiding to overwhelm the user. In the event of lacking a corresponding label for a component, it can remain unlabelled or the user can select the label "cannot_label".

Furthermore, to speed up the annotation process, similar elements were grouped together semantically based on function/operation during the subdivision step in the modeling process, so that when the `<EXPAND>` button is selected, all similar elements will be selected and labelled at once. The selected item(s) turn white on the plain model to be distinguished from the rest of the components, while a floating, highlighted bounding box appears surrounding the item on the textured model, to help the user of the tool identify it faster. Once, a component is labelled it will be given the corresponding colour of its label, so that it is differentiated from those elements that remain unlabelled. To avoid relabelling when an annotated element is picked, a message will appear to the user parapet

Figure 8: Explanatory images for label “parapet”.

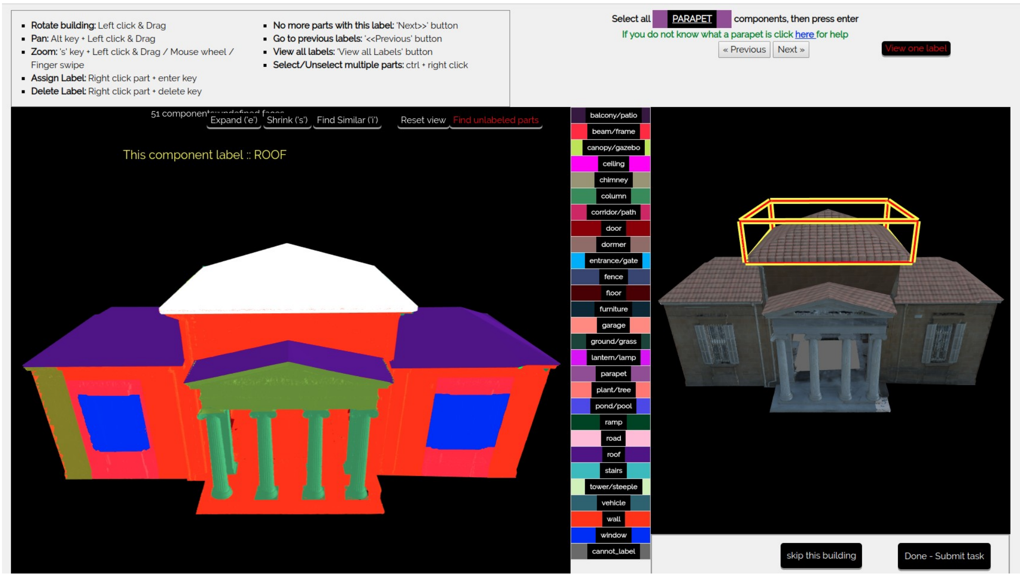


Figure 9: Partially annotated monument.

stating the current label of the element (fig. 9). Naturally, an architectural element can be relabelled if the user deems it necessary.

Finally, when the user is satisfied with the annotation completion and accuracy, she/he can select the `<Done - Submit task>` button in order to continue to the annotation of another building or exit the tool. In case a user is given a monument that is not of their interest, she/he can click on `<skip this building>` to load the next monument in the list. On the submission of a building, the annotation choices are saved and used in the next stage of the workflow which includes the development and testing of the platform's automated functionality briefly mentioned in the following section, and currently under development.

4.2. Platform

Following the demonstration of the ANNFASS platform for heritage annotation in the expert group meetings, feedback was received from the participants, regarding their experience with the tool, including suggestions about improving the user interface and the need for other desirable functionalities of the platform. These served as design and development guidelines for the final platform that will be delivered to the architecture history experts, at the end of the ANNFASS project.

The primary outcome of the questionnaires collected, was that this tool is much needed to the education community in particular, since it modernises, automates and speeds up the currently used methods and procedures in scholarly analysis in architecture. Responding to this feedback, the final ANNFASS platform, currently under development, will consist of both essential features for the visualisation of the 3D model of the building, as well as automated modules for comparative analysis. Essential tool features and functionalities include: viewing a monument's 3D model, or displaying its description and historical information.

Specifically, every model in the platform is accompanied with the following information: monument name, brief description / chronicle, architectural style(s) / stylistic influences, location (country and geographic position), the 3D model and lastly, a cover photo for preview purposes. During the expert group meetings, it was indicated that this is the minimal information needed to perceive the main context. More information can be added, if available, to enrich this perception, such as architect(s), construction start and end period, and a photo gallery of the monument. The ANNFASS tool addresses didactic needs in educational environments and the main purpose is not to automatically determine the construction period of the monument under study but rather to assist in identifying the historical references of decorative and architectural features of the building. In addition to the visualisation features, the platform will also be equipped with the following automated modules:

- Architectural Component Recognition
- Construction Period Recognition

Architectural Component Recognition

Analysing and identifying the various architectural elements and components (e.g., roof, door) is a necessary step in the process of studying a building's structure and style, that is also a time consuming and tedious procedure. The current advances in the field of machine learning allow the automation of this procedure, with the use of 3D CNNs, and specifically the employment of deep learning methods. The success of neural networks in image-based tasks, induced by the introduction of convolutional layers, inspired researchers to develop convolutional layers that can operate on 3D data through the use of convolutions in 3D space (3D CNNs) Wu et al. (2015); Wang et al. (2017); Choy et al. (2019); Wang et al. (2020). ANNFASS' architectural component recognition relies on an 3D CNN with sparse 3D convolutional layers, called MinkowskiNet Choy et al. (2019), adapted in our case to learn how to identify the various components of 3D building models.

In detail, the network is trained in a supervised manner, taking as input a set of 3D models and their ground truth labels (expert annotations), in order to learn patterns and extract features that appear frequently (fig. 10). The authors are aware that deep neural networks require a large number of data to learn effecting representations for part labeling. The Cypriot dataset on its own is not sufficient for this task. For this reason, another much larger 3D building dataset was developed at UMass, called BuildingNet Selvaraju et al. (2020), which consists of 513,087 annotated building mesh components across 2000 building models of different types (e.g., residential, religious), and 31 unique semantic labels (e.g., tower, wall).

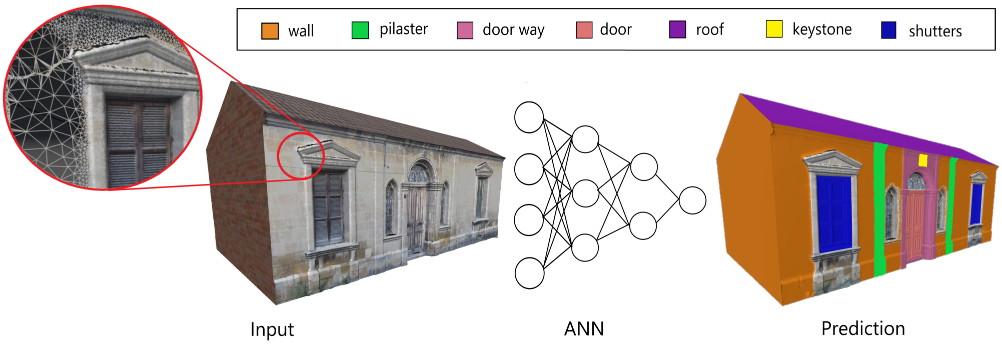


Figure 10: Architectural component recognition 3D CNN.

The aforementioned dataset, is used to train the 3D CNN on recognising the different structural components of buildings. Once the network is adequately trained on BuildingNet, a representative subset of the Cypriot monuments will be used for fine-tuning the network parameters. The fine-tuning phase aims to help the network identify specific to the Cypriot building labels (e.g., keystone), since only a few training examples are available in the training dataset. The remaining data will serve as test cases to evaluate the network's abilities on unseen data.

In the end, the trained network will apply the acquired knowledge (learned features) to classify components of 3D models uploaded by the platform users. It is not required for the uploaded models to be pre-segmented to components, as the used 3D CNN produces annotations using only the geometry information and textures of a model. Though, for purely aesthetic reasons, having this additional information can be used in post-processing steps to produce *smoother* results.

An example case is illustrated in fig. 11, where a toy example model (square) consists of a single component with four faces ($f1 - f4$) coloured based on two annotation strategies (face-based vs component-based labeling). In the first case, labels are derived for each mesh face, with faces $f1 - f3$ annotated as label-1 (orange) and $f4$ as label-2 (blue), resulting to an inhomogeneous component annotation. However, for the component-level labelling, an additional step is taken, that of averaging all component *child* faces label probabilities and assigning the predominant one to the whole component i.e., in this case label-1 (orange). Note that the only given information for this additional step was the component a face belongs to.

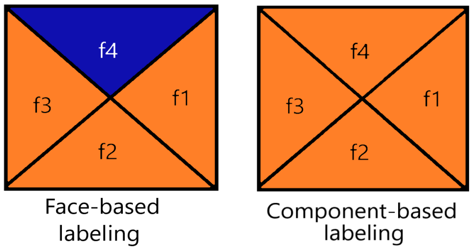


Figure 11: 3D CNN prediction projections on mesh on face and component level.

The inclusion of the 3D CNN in the platform allows a digital humanities expert to extract the architectural components of a building with minimal effort and in much less time, since the network will largely automate the classification task.

Automated suggestion of stylistic influences at building component / architectural feature level

The process of studying and understanding a heritage building in the context of the development and completion of the ANNFASS platform, continues after the steps presented by this article, with the detection of its stylistic influences. This task is proven to be more challenging than the architectural component labelling and recognition, since several styles can be concurrently present on a building - a common occurrence in the case of Cypriot heritage, with its characteristic hybridised architectural style. To tackle this issue, a similar approach to the previous task is pursued by the authors. That is, having another 3D CNN trained to recognise the architectural style influences of each component (if any) based on its appearance, and in the end of the process to present the various possible stylistic influences of the building identified by the tool to the user.

Once again, for the training of the 3D CNN a set of buildings with style-based annotations is required as input to the tool. To acquire this information another session with the architecture historians was conducted. This time, a period categorisation (Lusignan, Ottoman, etc.) was assigned to classify the various architectural components based on their appearance, instead of their functional aspects and type of architectural object (door, window, etc.). At the end of this process, and after the collection of historic architectural style labels and the training of the 3D CNN is completed, the user can load a new 3D model on the ANNFASS platform and have the 3D CNN recognise its stylistic influences by means of percentages (fig. 6). As this is an on-going task, a number of methods are being tested to determine the better performing one, before finalising the style recognition pipeline.

5. Conclusion

As stated before, this work is part of an on-going project (ANNFASS). The collection of data for the 3D model generation is currently completed, having 3D models for all selected Cypriot heritage buildings ready to be used for the training of the 3D CNNs. A variety of computation methods have been developed, tested and assessed regarding the reliability of the automated modules, which will be hosted on a server for better communication with the ANNFASS platform. The ANNFASS platform will be updated with multi-period monuments, built in Cyprus under different cultural influences, which in turn will be used to enrich the feature learning for the structure and style recognition 3D CNNs. A hidden aspect of the classification process of architectural elements is first to identify a building's stylistic influences and function by finding similarities that are common between monuments of the same historical period. By extension, the mapping of similarities makes it easier to identify the uncommon and the rarest samples that can be found in cultural heritage monuments. Monuments that are not able to be classified in a specific large group of canonical examples of a historical period can be viewed as uncommon or exquisite cases of architecture outside the norm / standard and established styles and could be examined as unique/ non-standard cases. In a sense, the ANNFASS platform strives to provide the capacity to identify what is hidden and uncommon in many examples of architectural elements, styles and buildings, contributing to scientific excellence in architectural history, education and research.

6. Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 739578 and under grant agreement No EXCELLENCE/1216/0352 and the Government of the Republic of Cyprus through the Directorate General for European Programmes, Coordination and Development. The work was also partially funded by Adobe, and project TIN2017-88515-C2-2-R from Ministerio de Ciencia, Innovación y Universidades, Spain.

Author contributions

Marissia Deligiorgi: Investigation (3D surveying & modelling, annotation); Data curation; Writing - original draft
Maria I. Maslioukova: Software (3D CNNs); Investigation (3D CNN training & testing); Writing - original draft
Melinos Avrekiou: Conceptualisation; Methodology; Supervision; Project coordination; Writing - review & editing
Andreas C.Andreou: Software (Platform); Writing - review & editing
Pratheba Selvaraju: Software (Annotation Tool, 3D CNN); Data curation (BuildNet dataset, mesh processing)
Evangelos Kalogerakis: Conceptualisation; Methodology; Writing - review & editing
Gustavo Patow: Discussions; Writing - review & editing
Yiorgos Chrysanthou: Resources; Writing - review & editing
George Artopoulos: Conceptualisation; Writing - review & editing

References

Baratin L, Bertozzi S, Moretti E, Spinella M. Gis and 3d models as support to documentation and planning of the baku historical centre (republic of azerbaijan). International Journal of Heritage in the Digital Era 2012;1:71–6. URL: http://journals.sagepub.com/doi/10.1260/2047-4970.1.0.71. doi:10.1260/2047-4970.1.0.71.

Bassier M, Genechten BV, Vergauwen M. Classification of sensor independent point cloud data of building objects using random forests. Journal of Building Engineering 2019;21:468–77. doi:10.1016/j.jobe.2018.04.027.

Cantoro G. Speeding up georeferencing with subpixel accuracy. In: Archaeology in the Digital Era. Amsterdam University Press; 2014. p. 752–60.

Canul-Ku M, Hasimoto-Beltran R, Jimenez-Badillo D, Ruiz-Correa S, Roman-Rangel E. Classification of 3d archaeological objects using multi-view curvature structure signatures. IEEE Access 2018;7:3298–313. doi:10.1109/ACCESS.2018.2886791.

Charalambous P, Artopoulos G. Enabling virtual collaboration in digital cultural heritage in the seem region. Scalable Computing 2018;19:161–74. doi:10.12694/scpe.v19i2.1348.

Choy C, Gwak J, Savarese S. 4D Spatio-Temporal ConvNets: Minkowski Convolutional Neural Networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2019;2019-June:3070–9. URL: http://arxiv.org/abs/1904.08755. arXiv:1904.08755.

Dallas RW. Measured surveys of historic buildings: User requirements and technical progress. Journal of Architectural Conservation 2003;9:58–81. doi:10.1080/13556207.2003.10785343.

Dhali MA, Jansen CN, de Wit JW, Schomaker L. Feature-extraction methods for historical manuscript dating based on writing style development. Pattern Recognition Letters 2020;131:413–20. doi:10.1016/j.patrec.2020.01.027.

Dirk J, Stucker C, Richard A, Wegner JD, Photogrammetry KS, Sensing R, Zürich E. Supervised outlier detection in large-scale mvs point clouds for 3d city modeling applications eth library supervised outlier detection in large-scale mvs point clouds for 3d city modeling applications. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. Copernicus Publications; 2018. p. 263–70. URL: https://doi.org/10.3929/ethz-b-000271685. doi:10.3929/ethz-b-000271685.

Engel C, Mangiafico P, Issavi J, Lukas D. Computer vision and image recognition in archaeology. In: AIDR '19: Proceedings of the Conference on Artificial Intelligence for Data Discovery and Reuse. 2019. p. 1–4. doi:10.1145/3359115.3359117.

Eramian M, Walia E, Power C, Cairns P, Lewis A. Image-based search and retrieval for biface artefacts using features capturing archaeologically significant characteristics. Machine Vision and Applications 2017;28:201–18. doi:10.1007/s00138-016-0819-x.

Feng M, Zhang L, Lin X, Gilani SZ, Mian A. Point attention network for semantic segmentation of 3d point clouds. Pattern Recognition 2020;107:107446. URL: http://www.sciencedirect.com/science/article/pii/S0031320320302491. doi:https://doi.org/10.1016/j.patcog.2020.107446.

Fiorucci M, Khoroshiltseva M, Pontil M, Traviglia A, Bue AD, James S. Machine learning for cultural heritage: A survey. Pattern Recognition Letters 2020;133:102–8. doi:10.1016/j.patrec.2020.02.017.

Georgopoulos A, Ioannidis C. Photogrammetric and surveying methods for the geometric recording of archaeological monuments photogrammetric and surveying methods for the geometric recording of archaeological monuments. In: International Federation of Surveyors (FIG). 2004. p. 22–7.

Grilli E, Remondino F. Classification of 3d digital heritage. Remote Sensing 2019;11:38121. URL: www.mdpi.com/journal/remotesensing. doi:10.3390/RS11070847.

Harrower N, Maryl M, Biro T, Immenhauser B, E-Humanities AWG. Sustainable and fair data sharing in the humanities: Recommendations of the allea working group e-humanities. 2020. URL: https://repository.dri.ie/objects/tq582c863/doi/tq582c863. doi:10.7486/DRI.TQ582C863.

Hillier B, Hanson J, Graham H. Ideas are in things: an application of the space syntax method to discovering house genotypes. Environment and Planning B: Planning and Design 1987;14:363–85. URL: http://epb.sagepub.com/lookup/doi/10.1068/b140363. doi:10.1068/b140363.

Jiménez-Badillo , Ruíz-Correa , García-Alfaro . Developing a recognition system for the retrieval of archaeological 3d models. In: Fusion of Cultures: Proceedings of the 38th. Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Granada, Spain. 2010. p. 325–31.

Kalogerakis E, Averkiou M, Maji S, Chaudhuri S. 3d shape segmentation with projective convolutional networks. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2017;:6630–9.

Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Commun ACM 2012;60(6):84–90. URL: https://doi.org/10.1145/3065386. doi:10.1145/3065386.

Lachambre S, Lagarde S, Jover C. Unity Photogrammetry Workflow. Technical Report; Unity Technologies; 2017.

LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to handwritten zip code recognition. Neural Computation 1989;1:541–51. doi:10.1162/neco.1989.1.4.541.

Llamas J, Lerones PM, Medina R, Zalama E, Gómez-García-Bermejo J. Classification of architectural heritage images using deep learning techniques. Applied Sciences (Switzerland) 2017;7. doi:10.3390/app7100992.

Malinverni ES, Pierdicca R, Paolanti M, Martini M, Morbidoni C, Matrone F, Lingua A. DEEP LEARNING for SEMANTIC SEGMENTATION of 3D POINT CLOUD. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. Copernicus Publications; volume 42; 2019. p. 735–42. doi:10.5194/isprs-archives-XLII-2-W15-735-2019.

Mathias M, Martinovic A, Weissenberg J, Haegler S, Gool LV. Automatic architectural style recognition. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. Copernicus Publications; volume XXXVIII-5/W16; 2012. p. 171–6. doi:10.5194/isprsarchives-xxxviii-5-w16-171-2011.

Messaoudi T, Véron P, Halin G, Luca LD. An ontological model for the reality-based 3d annotation of heritage building conservation state. Journal of Cultural Heritage 2018;29:100–12. doi:10.1016/j.culher.2017.05.017.

Michaelides D. Historic Nicosia. Rimal Publications, 2012.

Morbidoni C, Pierdicca R, Quattrini R, Frontoni E. GRAPH CNN with RADIUS DISTANCE for SEMANTIC SEGMENTATION of HISTORICAL BUILDINGS TLS POINT CLOUDS. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. Copernicus Publications; volume 44; 2020. p. 95–102. doi:10.5194/isprs-archives-XLIV-4-W1-2020-95-2020.

Pierdicca R, Paolanti M, Matrone F, Martini M, Morbidoni C, Malinverni ES, Frontoni E, Lingua AM. Point Cloud Semantic Segmentation Using a Deep Learning Framework for Cultural Heritage. Remote Sensing 2020;12(6):1005. URL: https://www.mdpi.com/2072-4292/12/6/1005. doi:10.3390/rs12061005.

Poux F, Neuville R, Billen R. POINT CLOUD CLASSIFICATION of TESSERAE from TERRESTRIAL LASER DATA COMBINED with DENSE IMAGE MATCHING for ARCHAEOLOGICAL INFORMATION EXTRACTION. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. Copernicus Publications; volume 4; 2017. p. 203–11. doi:10.5194/isprs-annals-IV-2-W2-203-2017.

Prasomphan S, Jung JE. Mobile application for archaeological site image content retrieval and automated generating image descriptions with neural network. Mobile Networks and Applications 2017;22:642–9. doi:10.1007/s11036-016-0805-6.

Qi CR, Yi L, Su H, Guibas LJ. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Advances in Neural Information Processing Systems 2017;:5100–9.

Selvaraju P, Nabail M, Loizou M, Maslioukova M, Averkiou M, Chaudhuri S, Kalogerakis E. BuildingNet: Learning to Label 3D Buildings, UM-CS-2020-003, December 21 2020. Technical Report; University of Massachusetts at Amherst; 2020.

da Silva IN, Spatti DH, Flauzino RA, Liboni LHB, dos Reis Alves SF. Artificial neural network architectures and training processes. Artificial Neural Networks 2017;:21–8doi:10.1007/978-3-319-43162-8_2.

Stathopoulou EK, Remondino F. SEMANTIC PHOTOGRAMMETRY - BOOSTING IMAGE-BASED 3D RECONSTRUCTION with SEMANTIC LABELING. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. Copernicus Publications; volume 42; 2019. p. 685–90. doi:10.5194/isprs-archives-XLII-2-W9-685-2019.

Su H, Jampani V, Sun D, Maji S, Kalogerakis E, Yang MH, Kautz J. Splatnet: Sparse lattice networks for point cloud processing. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2018;:2530–9URL: http://arxiv.org/abs/1802.08275.

Wang PS, Liu Y, Guo YX, Sun CY, Tong X. O-cnn: Octree-based convolutional neural networks for 3d shape analysis. ACM Transactions on Graphics 2017;36. doi:10.1145/3072959.3073608.

Wang PS, Yang YQ, Zou QF, Wu Z, Liu Y, Tong X. Unsupervised 3D Learning for Shape Analysis via Multiresolution Instance Discrimination. ACM Trans Graphic 2020;.

Wang Y, Sun Y, Liu Z, Sarma SE, Bronstein MM, Solomon JM. Dynamic graph CNN for learning on point clouds. CoRR 2018;abs/1801.07829. URL: http://arxiv.org/abs/1801.07829. arXiv:1801.07829.

Wu Z, Song S, Khosla A, Yu F, Zhang L, Tang X, Xiao J. 3D ShapeNets: A Deep Representation for Volumetric Shapes. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2015;07-12-June-2015:1912–20. URL: http://arxiv.org/abs/1406.5670.

arXiv:1406.5670.