Learning from synthetic point cloud data for historical buildings semantic segmentation

CHRISTIAN MORBIDONI, Universitá Politecnica delle Marche (DII), ITALY
ROBERTO PIERDICCA, Universitá Politecnica delle Marche (DICEA), ITALY
MARINA PAOLANTI, Universitá Politecnica delle Marche (DII), ITALY
RAMONA QUATTRINI, Universitá Politecnica delle Marche (DICEA), ITALY
RAISSA MAMMOLI, Universitá Politecnica delle Marche (DICEA), ITALY

Historical heritage is demanding robust pipelines for obtaining HBIM (Heritage Building Information Modeling) models that are fully interoperable and rich in their informative content. The definition of efficient *Scan-to-BIM* workflows represent a very important step towards a more efficient management of the historical real estate, as creating structured 3D models from point clouds is complex and time-consuming. In this scenario, semantic segmentation of 3D Point Clouds is gaining more and more attention, since it might help to automatically recognize historical architectural elements. The way paved by recent Deep Learning approaches proved to provide reliable and affordable degrees of automation in other contexts, as road scenes understanding. However, semantic segmentation is particularly challenging in historical and classical architecture, due to the shapes complexity and the limited repeatability of elements across different buildings, which makes it difficult to define common patterns within the same class of elements. Furthermore, as Deep Learning models requires a considerably large amount of annotated data to be trained and tuned in order to properly handle unseen scenes, the lack of (big) publicly available annotated point clouds in the historical building domain is a huge problem, which in fact blocks the research in this direction. On the other hand, creating a critical mass of annotated point clouds by manual annotation is very time-consuming and impractical. To tackle this issue, in this work we explore the idea of leveraging synthetic point cloud data to train Deep Learning models to perform semantic segmentation of point clouds obtained via Terrestrial Laser Scanning (TLS). The aim is to provide a first assessment of the use of synthetic data to drive Deep Learning based semantic segmentation in the context of historical buildings. To achieve this purpose, we present an improved version of the Dynamic Graph Convolutional Neural Network (DGCNN) named RadDGCNN. The main improvement consists on exploiting the radius distance. In our experiments, we evaluate the trained models on synthetic dataset (publicly available) about two different historical buildings: the Ducal Palace in Urbino, Italy, and Palazzo Ferretti in Ancona, Italy. RadDGCNN yields good results, demonstrating improved segmentation performances on the TLS real datasets.

CCS Concepts: • **Computing methodologies** → **Scene understanding**; **Neural networks**; • **Applied computing** → **Architecture (buildings)**;

Authors' addresses: Christian Morbidoni, Universitá Politecnica delle Marche (DII), Via Brecce Bianche, Ancona, IT, 60131, ITALY, c.morbidoni@staff.univpm.it; Roberto Pierdicca, Universitá Politecnica delle Marche (DICEA), Via Brecce Bianche, Ancona, IT, 60131, ITALY, r.pierdicca@staff.univpm.it; Marina Paolanti, Universitá Politecnica delle Marche (DII), Via Brecce Bianche, Ancona, IT, 60131, ITALY, m.paolanti@staff.univpm.it; Ramona Quattrini, Universitá Politecnica delle Marche (DICEA), Via Brecce Bianche, Ancona, IT, 60131, ITALY, r.quattrini@staff.univpm.it; Raissa Mammoli, Universitá Politecnica delle Marche (DICEA), Via Brecce Bianche, Ancona, IT, 60131, ITALY, r.mammoli@pm.univpm.it.

XXXX-XXXX/N/A/3-ARTN/A $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Additional Key Words and Phrases: Deep Learning, Point Cloud Semantic Segmentation, Synthetic Point Cloud, Cultural Heritage, Historical Building, Dynamic Graph Convolutional Neural Network, Radius Distance, Scan-to-BIM.

ACM Reference Format:
Christian Morbidoni, Roberto Pierdicca, Marina Paolanti, Ramona Quattrini, and Raissa Mammoli. N/A. Learning from synthetic point cloud data for historical buildings semantic segmentation. *ACM J. Comput. Cult. Herit.* N/A, N/A, Article N/A (March N/A), 17 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In the context of Cultural Heritage (CH) management and preservation, creating accurate and rich digital representations of historical buildings is of primary importance. Several research works investigate how to obtain accurate and reliable information of a heritage building through reality capture and through 3D modeling for supporting restoration purposes or knowledge-based applications. Historical heritage is demanding robust pipelines for obtaining HBIM (Heritage Building Information Modeling) models that are fully interoperable and rich in their informative content. The availability of efficient workflows could represent a very important step towards a more efficient management of the historical real estate.

The Heritage documentation process usually starts from the acquisition of 3D point clouds through Terrestrial Laser Scanning (TLS) or photogrammetric techniques that produce accurate and precise point clouds, up to millions of points. Photogrammetry methods and TLS tools allow to generate a great amount of geometrical and 3D coordinates of a scene. However, while point clouds are useful to visualize a given scene, they are mainly unstructured representations and lack of information about the single objects, such as surfaces, dimensions and semantics. For this reason, point clouds are often used, within dedicated BIM software environments, as a starting point to build parametric 3D representations, which incorporates semantic information and where architectural elements are identified and properly labeled. Such a process is often referred to as *Scan-to-BIM*.

Creating structured 3D models from point clouds is a manual and time-consuming process. It requires domain experts to visually inspect the point cloud identifying different kind of architectural elements of interest, to then model the objects, correctly position each element on the scene and annotating it according to domain thesauri and/or ontologies. Therefore, there is a strong need for partially automatizing the *Scan-to-BIM* process, developing methodologies and tools that are able to assist domain experts, making the whole process easier and faster. An important task in this direction is that of automatically identifying different kinds of architectural elements in the point cloud. This task is referred to as *semantic segmentation* and involves the classification of each point in the cloud as belonging to a particular kind of object (e.g. wall, roof, column, vault, etc.). Speeding up the *Scan-to-BIM* process, which is currently carried out manually, is crucial, but the automatic segmentation of 3D data is yet a bottleneck. While machine and Deep Learning techniques proved to be successful in semantically segmenting 2D objects (i.e. pictures), 3D point clouds are geometric structures of irregular nature, characterized by the lack of a grid, with a high variability of density, unordered and invariant to transformation and permutation, which makes the task more difficult. Recently, promising results have been achieved by the application of Deep Learning techniques specifically designed to handle point clouds [27] [28], [36]. Such approaches was applied to segmentation of indoor scenes [3], i.e offices, and evaluated on standard benchmarks, as the Stanford 2D-3D-Semantics Dataset(http://buildingparser.stanford.edu/dataset.html).

However, semantic segmentation is particularly challenging in historical and classical architecture, due to the shapes complexity and the limited repeatability of elements across different buildings, which make it difficult to define common patterns within the same class of elements [18],

[26]. Even if the shape is repeatable, perhaps for a given architectural style, the objects are still unique as they are handcrafted and not serialized. Another problem is that the objects are ancient and subject to erosion and decay.

Furthermore, as Deep Learning models requires a considerably large amount of annotated data to be trained and tuned in order to properly handle unseen scenes, the lack of (big) publicly available annotated point clouds is a huge problem, which in fact blocks the research in this direction. On the other hand, creating a critical mass of annotated point clouds by manual annotation is very time-consuming and impractical.

To tackle this issue, in this work we explore the idea of leveraging synthetic point cloud data to train state of the art Deep Learning models to perform semantic segmentation of real TLS point clouds. To the best of our knowledge this is the first attempt in this direction. We derive synthetic annotated point clouds from existing 3D models with the aim of obtaining a critical mass of data to learn from, without the need of time-consuming manual annotation of huge amounts of TLS data. While synthetic point clouds are relatively easy to obtain, they are inherently different from real TLS data, as they have a more regular structure and typically capture a lower level of details. Thus, the aim of the present research is to provide a first assessment of the use of synthetic data to drive Deep Learning based semantic segmentation in the context of historical buildings.

The main contributions of this paper are as follows:

- We provide the first study in assessing the use of synthetic point cloud data in the context of built heritage semantic segmentation, evaluating trained DGCNN models on real TLS data.
- We release a dataset composed of synthetic annotated point clouds derived from curated 3D models of 10 different historical buildings and from 159 3D models of single architectural elements collected from open web archives. The dataset [1] is available upon request to the authors and can be used to further advance research in this field.
- We propose a variation of the DGCNN architecture presented in [36] (see Section 2), based on the use of radius distance, which was proven to increase performances on our test datasets.

The paper is organized as follows: Section 2 provides an overview of the previous work concerning the recent efforts in point clouds semantic segmentation based on Deep Learning and the generation and utilisation of synthetic data. In Section 3 we describe the data and the methods used in this work. In Section 4 and 5 we present our experiments and their results respectively. Finally, in Section 6 we draw conclusions.

2 RELATED WORKS

In literature, the use of synthetic dataset involves different fields of application. The first work dates back to 1975 when Miller [20] generated appropriate test-case data to have the software quality enhancement in the near term using a methodology that relates functional software test-cases with its formal specifications to achieve a correspondence between the software and its specifications. Myers et al. [22] show that the introduction of simulated dataset brings several advantages. They developed a software for studying the human genome and a simulator (Celsim) that allows to describe and stochastically generate a target DNA sequence with different repeated structures. Multidimensional synthetic data generation was also applied to clustering and outlier analysis in [24].

Synthetic thermal images are used to augment an existing dataset of visible images in the work of Kniaz [16]. For the augmentation of the existing large dataset with synthetic thermal images, they developed a deep convolutional neural network, named Thermalnet, inspired by the colorisation deep neural networks. Departing from a single 2D-dimensional image, Fan et al. [9] reconstruct

[1] http://vrai.dii.univpm.it/content/synthetic-dataset-point-cloud-semantic-segmentation

the 3D geometry of the complete object. They use a Deep Learning algorithm to generate the 3D point cloud representation, predicting the real 3D shape of the object.

Li et al. [17] propose an approach based on supervised learning to classify 3D real urban scenes acquired from TLS, VLS and ALS. The training phase of the random forest network uses an automatically generating dataset of 3D point clouds. In [32] a synthetic images dataset of urban scene, automatically extracted from virtual urban scenes, is presented and shown to improve 2D semantic segmentation performances when added to manually annotated benchmark dataset during training phase. The works of [37], realistic synthetic point clouds of road scenes have derived from a popular videogame (Grand Theft Auto) and used to augment standard a benchmark dataset (KITTI [11]), improving performances of CNN based semantic segmentation.

In [34] virtual LiDAR (Light Detection and Ranging) sensors are used to acquire synthetic point cloud data from street scenes, simulating a variety of point cloud acquisition instrumentation. Experiments in semantic segmentation in the context of autonomous driving shows performances boost with by synthetic augmentation of manually annotated data. According the authors [15] a synthetic dataset of text generated using an engine is an optimal solution to replace real data, providing an enormous amount of data used for training. They use the synthetic dataset to train a convolutional neural network, then, in order to test the validity of the approach, they consider publicly available dataset. Similar is the work of [35] that trains a convolutional neural network using synthetically generated examples of text. The work of [13] uses a Deep Learning approach for automatic classification of 3D point cloud data. Due to the necessity of a great amount of training data, the authors generate SynthCity, a synthetic dataset of point clouds in a Urban/Suburban environment using the Blensor plugin for Blender. Their aim is to demonstrate that a network trained using synthetic data is able to well generalize.

In the context of cultural heritage synthetic dataset is not often used; however, some meaningful works are here presented [25]. [19] shows an overview of training approaches for the recognition of optical character recognition in the historical documents. Since the main problem is the lack of the annotated data, they summarize different ways to prepare synthetic data. They train a convolutional recurrent neural network classifier using synthetic dataset, and validate their approach with a real annotated dataset. The work of [38] uses a big, synthetically rendered dataset that consists of scenes with generated shapes and real-world to train Relight-Net, a deep convolutional neural network. An approach that uses a synthetic dataset is that proposed by [33], where is presented a comparison between a building information model and a point cloud of an indoor environment. They use synthetic dataset and ISPRS reference dataset to evaluate the performance in the detection of the differences between the two representation.

Thought it is not the purpose of this manuscript to provide a literature review of DL approaches for semantic classification of dense point clouds, it is worth to mention some architectures proposed recently. Interested readers can find a broader discussion of the topic in [12]. State-of-the-art deep neural networks are specifically designed to deal with the irregularity of point clouds, directly managing raw point cloud data, rather than passing to an intermediate regular representation. Among the others, PointNet [27] was the pioneer of this approach. This network obtains permutation invariance of points by independently operating on each point, and subsequently by the application of a symmetric function to accumulate features. Extensions of PointNet [28] analyse neighbourhoods of points, rather than acting on each separately. These permit the exploitation of local features, thus improving upon the performance of the basic model. These methods consider points at local-scale to preserve permutation invariance. PCNN (Point Convolutional Neural Network) [4] is a DL framework for applying CNN to point clouds. The point cloud convolution is defined by a pull-back of the Euclidean volumetric convolution via an extension-restriction mechanism. This consists of two operators, extension and restriction, that map point cloud functions to volumetric functions.

PCNN generalises image CNNs and allows adaptation of their architectures to point cloud settings. DGCNN [36], Dynamic Graph Convolutional Neural Networks, introduce the EdgeConv operation. EdgeConv is a module that creates edge features that describe the relationships between a point and its neighbours, rather than generating point features directly from their embeddings. This module is designed to be invariant to the ordering of neighbours, and it is permutation invariant. DGCNN has been shown to provide state of the art results in indoor scene segmentation. In our experiments we base on DGCNN and we evaluate the trained models on two TLS point clouds acquired from two different historical buildings: the Ducal Palace in Urbino, Italy, and Palazzo Ferretti in Ancona, Italy.

3 MATERIALS AND METHODS

The methodology developed to perform our experiments is based on few steps, summarized in Figure 1: after the dataset collection and organization step, the synthetic training dataset is created, starting from architectural objects. Afterwards, pre-processing actions are performed, as to provide the RadDGCNN with data suitable for the subsequent segmentation and validation phases. In the following subsections, these steps are detailed.

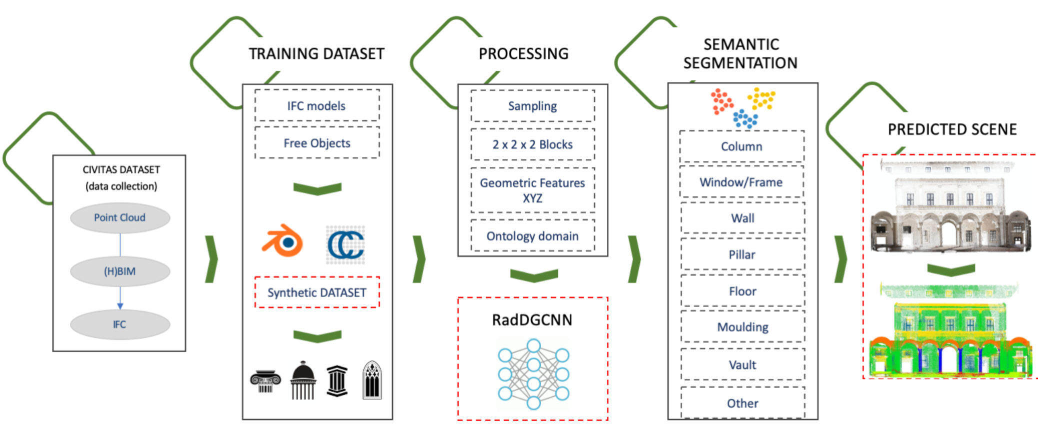

Fig. 1. The proposed semantic segmentation workflow

3.1 Classes of architectural elements

The first fundamental step for the methodology here illustrated is to identify classes of architectural elements to be considered for the semantic segmentation task. The classes we selected for our experiments are as follows: Column, Window/Door, Wall, Pilaster, Floor, Moulding, Vault, Other. Classes and their Level Of Detail (LOD) were chosen taking into account three aspects: coherence with existing thesaurus, consistency with the historical period and its language and correspondence with the data acquired through TLS. Concerning the chosen detail, reference was made to the ontological subdivisions proposed by robust thesauri, available online, related to art and architecture field. For this work the main reference is the vocabulary of classical architectural elements and their hierarchies developed by the Getty Museum - Research [2]. The hierarchical level of the division into classes is set up between structural elements and decorative components. The grade of the building breakdown allows to have a full match with this first test conducted. In fact, the expected accuracy

[2]http://www.getty.edu/research/tools/vocabularies/aat/

of the point cloud, in accordance with the reference literature despite used in other domains, is in compliance with the general level of detail of architectural elements. The classes chosen, moreover, are coherent with the main elements of the classical architectural language, massively used in Renaissance buildings (taken as reference case studies). For some semantic classes it was necessary to implement the available dataset with some 3D models online. The architectural classes are also compliant with the data acquired through TLS. In relation to this survey methodology and in order to maintain the highest uniformity to acquired data, some classes (as roof) were not taken into account even though they are in the vocabularies and historical architectural language. As in the case of any historical architecture of great value and complexity, despite the extensiveness of the classes, some elements were not included in any one of these. All the architectural elements not matching any of the selected classes were grouped into the *Other* class.

3.2 Synthetic Training Dataset

In this section we describe the dataset used to in our experiments to train the deep neural model. The dataset is made of point clouds derived from existing 3D models and is composed by two parts. The first includes points clouds extracted from 3D models of real historical buildings. In section 3.2.1 we provide details about about the *Scan-to-BIM* modeling process carried out by experts as part of their research activity. The second part includes point clouds derived from a collection of 3D objects of single architectural elements, harvested from open online repositories, detailed in section 3.2.2.

In both cases, we used the open source software CloudCompare [3] to convert the 3D models, in OBJ format, to point clouds, by randomly sampling points from surfaces with a density of 1000 points per m^2. The same sub-sampling was applied to TLS point clouds used to test the neural networks.

3.2.1 3D models from traditional Scan-to-BIM workflows. The main challenge in the HBIM field is to incorporate the huge knowledge about historical architecture available from various domains in reliable and accurate models, exploiting existing platforms and their potentials in semantic intelligence, parametric information and glossary feeding. The approach proposed in this article builds on previous semantic structuring and standardization of a significant set of buildings, thanks to the robust know-how of the research group in digitization of historical architecture and its representation [30].

An effective BIM-based approach for the documentation of architectural heritage [2] was applied in last years for the development of different case studies [29], clearly demonstrating a substantial coherence between the classical way of building [5] and the possibility of representing it in semantic models [7] [31]. An interesting methodological approach shows a semantic description of architectural elements based on theoretical reflections and research experiences . A common and validated concept is to extract knowledge of classical architecture from treatises as well as from accurate survey and, then, to develop parametric modeling at different object level [6]. Variants grow out from the parametric models by editing or regrouping parameters based on grammars or the awareness of operators.

The first approach to the HBIM in the Ducal Palace was a basic and already assessed modelling phase, based on Revit editing tools such as loadable or local families. However, the parameterization and adoption of robust taxonomies were intended as fundamental features for informed models. In order to develop an in-deep analysis and to develop all parts in a rigorous way, a first nucleus of the Courtyard of Honor has been modeled. The rhythmic scanning and the modularity of the space designed by Luciano Laurana includes all elements able to synthesize the whole Palace. The

[3]https://www.danielgm.net/cc/

outlines of the architectural elements were created starting from the integration of the survey data with the classical literature and with more recent references, obtaining also a reference in order to approach the issue of different levels of detail (LOD). Regarding the vertical structures, the main element is a Corinthian order with sudvision that currently was not considered in the following clustering. While all openings were semantically distributed in "window" and "frames". Regarding the loadable families "column", "window" and "pilasters", the objects have been modelled setting geometric constraints able to develop a parameterization. The corners of the Courtyard needed "ad hoc" local families, according to the innovative design of a “corner solution” well know in historiography and literature. Another important step in modelling, inoculating rich feature in the present framework, is the definition of LOD, allowing different levels of representation of reality as well as to extract point clouds shaped according to correct details. The 3D model obtained is informed, annotated and detailed.

In the case of Palazzo Ferretti, the procedure had foreseen to build a parametric model of the building in a Revit environment integrating several sources: Terrestrial Laser Scanner (TLS) survey served, two-dimensional surveys and photogrammetric survey for detailed information, as well as a continuous comparison with treatises rules. The main work regards the moldings of main façades with the realization of loadable families (*.rfa) incorporating semantics and geometric constraints in coherence with the practice of classical architecture [21]. The most significant results since now for Palazzo Ferretti concern the realization of parametric libraries for the external and internal building moldings, now those molding can feed our synthetic data set.

Considering the need to enlarge the semantically structured dataset, a new use of models coming from the Palladio Library[4] was performed. The idea was mainly derived from the concept of the Palladio Library, a geo-database [1], [10] constituting a sort of ante litteram HBIM. The group of models developed for the Palladio Library included 13 models, in particular two bridges and various villas and/or barchessa. For the present work, only the villas were considered because they presented classes consistent with those of the point clouds that were to be segmented. A main feature of those models, in the light of the present work, was the ‘shape-grammar’ adopted in the 3D models construction using a pre-established set of tree-shaped formal rules that indicate a clear purpose and evident structure. This organization was extended, in our cases, over several hierarchical levels, allowing now to find out matching with previous clustering. In fact, the models from Palladio Library were stored in 3DM file format incorporating a taxonomy thanks to layer naming, allowing an export in OBJ format and the following creation of synthetic semantically structured point cloud. From all above mentioned models, engineered by domain experts, it has been possible to obtain the core of the synthetic dataset that is robust enough to instantiate the subsequent workflow. In Figure 2 four samples of 3D models of existing historical buildings included in the dataset are illustrated. For each one of them we show the original form and the generated annotated point cloud.

3.2.2 3D Objects from online archives. While some classes, e.g. wall and floor, are extensively present in the 3D models described above, the some does not hold for other classes of architectural elements. For example not all the buildings have columns or vaults, and, when these elements are present, they are usually very similar within the same building. On the contrary, to learn generalizable features, we need to improve shape variability within the same class. In an attempt

[4]The digital documentation campaign and models construction of Andrea Palladio’s villas and bridges were part of the CISA Andrea Palladio project Palladio Library, funded by Arcus. Principal Investigator: Marco Gaiani, scientific responsible of the Universitá Politecnica delle Marche project unit: Paolo Clini.

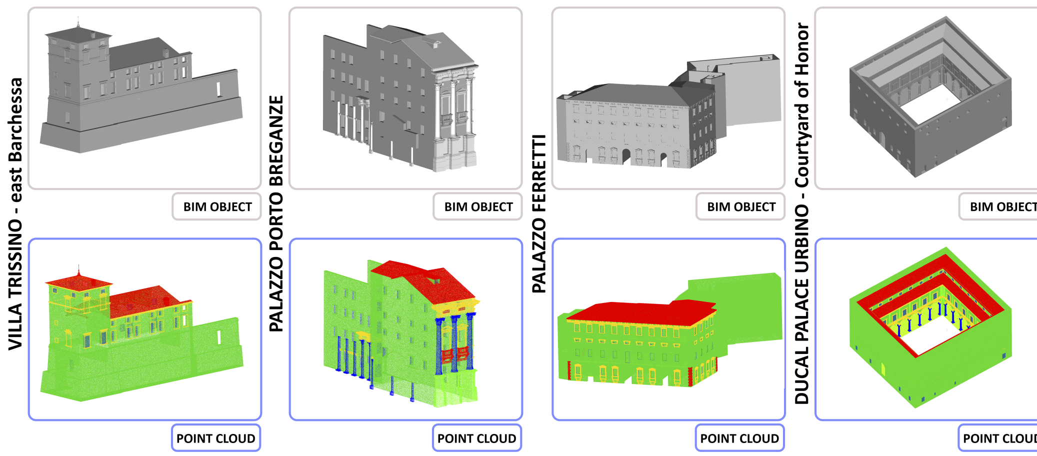

Fig. 2. Examples of 3D models of historical buildings and the derived annotated point clouds.

class	#3D models	#points
column	53	448297
Window/Door frame	41	1041552
moulding	31	442809
pilaster	24	558380
vault	10	3743896

Table 1. Overview of the 3D objects collected from online repositories

to address such issues, we extended our synthetic dataset with 3D models of single architectural elements, by selecting publicly available 3D objects from a number of online repositories. [5]

In Table 1 we report, for each class, the number of 3D models collected, which might include one or more architectural elements, and the total number of points obtained by transforming such models into point clouds.

As shown in Figure 3, the types of objects were chosen as a refinement of the classes that are lacking, prioritizing models of complexity and quality comparable to those modelled by the authors.

3.3 Survey Point Clouds

As detailed in [8, 23] the 3dimensional digitization of the Ducal Palace at Urbino is a first outcome allowed by the Strategic Project CIVITAS. In this framework a complete survey campaign, mainly based on TLS and photographic data capturing, was carried out. At the current stage, the whole numerical model of the Palace consists in 1.790 mln of points, while the experiment deals with a part of the point cloud related to the Courtyard of Honour (Table 2).

[5] http://www.cgtrader.com
https://www.3dcadbrowser.com
https://www.turbosquid.com
https://free3d.com
https://archive3d.net https://3dwarehouse.sketchup.com

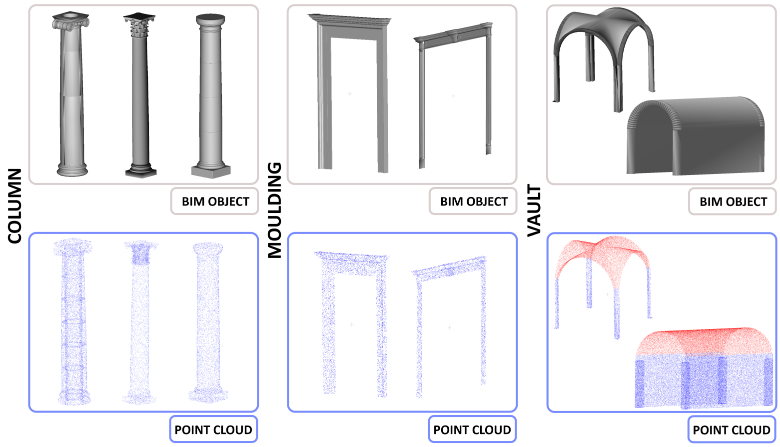

Fig. 3. Examples of 3D models of single architectural elements of different classes and the corresponding synthetic point clouds.

The first point cloud conceived as a frame is a comprehensive model for this large-complex building: a main reference data set was provided by the Leica Backpack acquisition. For the current work, instead, we prefer to use point cloud carried out thanks to the TLS (Leica ScanStation C10 and a Leica ScanStation P40) mounted on a tripod. This acquisition was performed by setting different levels of resolution according to the complexity of the rooms and their decorative elements, optimizing scanning times. Other acquisition are currently on going with Karta and UAVs in order to complete the numerical model.

The TLS survey operations of Palazzo Ferretti complex date back to 2014 and was carried out with 69 stand points: 49 stations (58 scans) for the external survey of the building and 20 stations and scans for the inner survey, involving the ground and the first floor. These scans were acquired with a resolution of 1 cm at 100 m, some parts in which greater details were needed are took with more dense point cloud (0,5 cm at 100 m). The final point cloud obtained by the alignment of the clouds of all the 78 scans is made up of 1.2 billion points (see table 2). For this research a part of the point cloud, referring to the Renaissance age part of the building was exploited in order to cover the classes already available in the Ducal Palace of Urbino. Only a part of the whole point cloud was then annotated (Table 2).

Finally, in Figure 4 we show the distribution of points among the selected classes in our synthetic training dataset and in the two annotated survey point clouds used in our experiments.

3.4 Deep neural network model

Dynamic Graph Convolutional Neural Networks (DGCNN), recently proposed in [36], was shown to provide state of the art semantic segmentation performances on standard banchmarks, as the S3DIS dataset [3]. In this paper we experimented with both DGCNN and a novel modified version (RadDGCNN) introduced in this work. We then combine the two approaches to further increase performances.

3.4.1 Architecture. The DGCNN model is based on the EdgeConv operation, introduced in [36], which learns hidden features of a point based on its neighbour points. Despite its name, the EdgeConv is implemented with a Multilayer Perceptron (MLP) fed with so called *edge features* and by a subsequent max-pooling operation on the learned features. In our case, the edge features

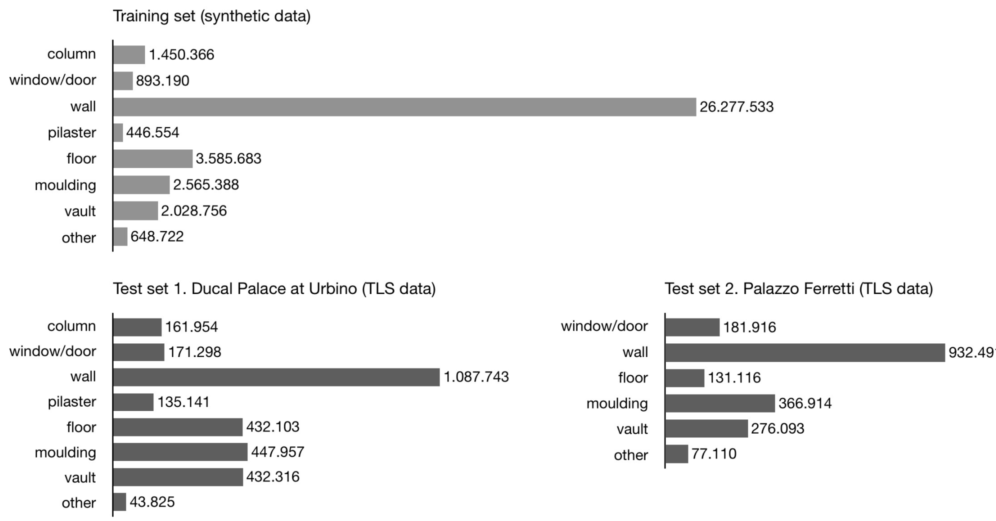

Fig. 4. Number of points per classes in the training set (synthetic set) and in the test set (from TLS survey).

Ducal Palace at Urbino

	Whole point cloud	Annotated point cloud (Courtyard of Honor)
Number of points	1.224 mln	17,3 mln
max alignment error	0,006 m	0,008 m
min alignment error	0,001 m	0,003 m
RMS	0,005 m	0,004 m

Palazzo Ferretti

	Whole point cloud	Annotated point cloud
Number of points	1.273 mln	8,4 mln
max alignment error	0,008 m	0,005 m
min alignment error	0,003 m	0,001 m
RMS	0,007 m	0,007 m

Table 2. Number of points, max/min alignment error and RMS of the two TLS point clouds used in our experiments.

of a point are composed by the 3-dimensional vector representing the point coordinates and the distance vectors between the point and its neighbour points.

The EdgeConv operation is initially performed on the input point cloud and then, repeatedly, on the output of the previous EdgeConv layer, thus learning hierarchical local hidden features for each point. The output hidden features of all the points learned by the EndgeConv layers are concatenated and fed to an MLP network to learn global point cloud features. Finally, the global features and the local features are concatenated and fed to an other MLP that outputs the predicted category for each point.

The network architecture used in our experiments is depicted in Figure 5. In this figure, N represents the number of points for each input point cloud and P is the number of classes considered for point-based classification, and K is the number of neighbors point considered for each point. Three EdgeConv layers are used to learn local point features. The learned features by each EdgeConv

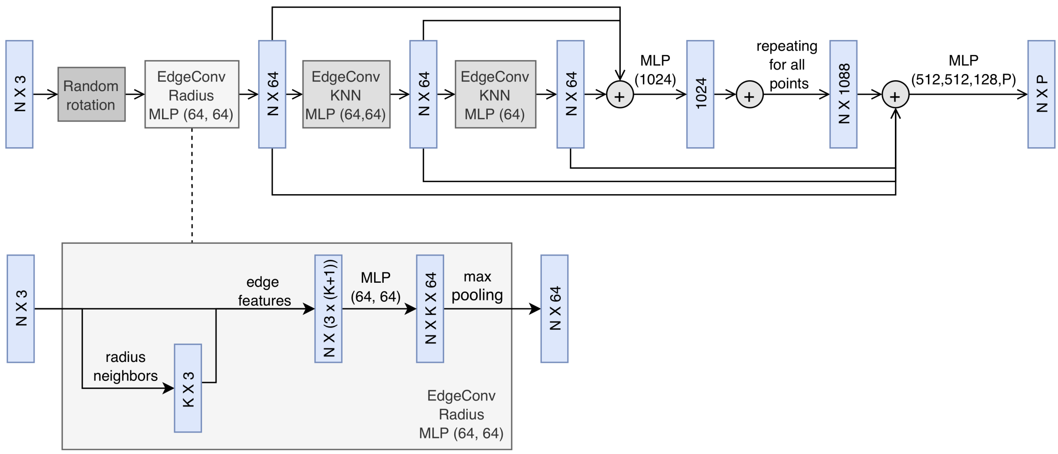

Fig. 5. The architecture of the deep neural network adopted in our experiments.

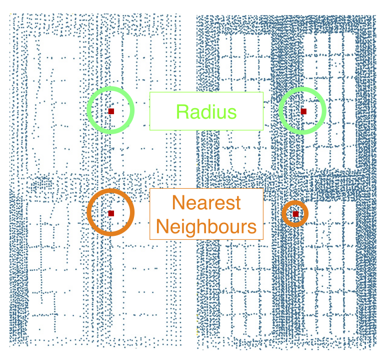

Fig. 6. KNN vs. Radius

layer for each point are then concatenated and global features are learned with a three layer MLP. We adopted the same settings, size of MLPs and number of layers, used in [36].

3.4.2 Points neighborhood. One important decision to be made in implementing a DGCNN is how to define the neighbourhood of a point. In [36], the K-Nearest Neighbours (KNN) of a point are considered to calculate edge features at each subsequent EdgeConv operation. However, as suggested by recent studies [14], KNN might not be the best method when dealing with non uniformly dense point clouds. This is indeed our case, as laser scan acquisition generally produce non uniform point clouds. In Figure 6 we show a fragment of the TLS point cloud of the Ducal Palace of Urbino, where two similar windows were captured with very different levels on density. As illustrated, in the case of KNN, the spatial extension of the neighborhood considered for the two points (marked in red) can be very different, depending on the local density of the point cloud. This in turn, might lead to different edge features for two points that belong to the same class (window). On the other hand, using radius distance, the neighbour points cover the same area.

Driven by this consideration, we experimented with a variation of the DGCNN architecture (called *RadDGCNN* in this paper), where the radius distance is used in the first EdgeConv operation in place of the KNN (see Figure 5). In practice, for each point, we first select all the points within the distance radius R from the point itself, then we randomly sample K of such points and provide them as input to the MLP to extract the edge features. Such a sampling operation allows us to always provide the same number of neighbours points (and thus the same number of edge features) as input to the MLP.

4 EXPERIMENTAL SETTING

We trained the DGCNN and the RadDGCNN models on our synthetic dataset by splitting each point cloud into blocks of equal dimensions (2 meters X 2 meters X 2 meters) each containing 4096 randomly sampled points. We did the same for the test set. The point cloud segmentation, which is also used in previous works on scene segmentation via neural networks, is a mandatory pre-processing step as providing the entire point clouds as input would be computationally unfeasible. Randomly sampling points for each block allows us to obtain inputs of equal dimensions, which is a requirement of the DGCNN architecture.

In our experiments, each point is represented by a 6-dimensional vector, where the first 3 elements are the absolute XYZ coordinates of the point in the original point cloud, while the following 3 values represent the normalized coordinates within the block (in the [0-1] range). To allow fair comparison, we adopted the same network hyper-parameters as in [36], with the only difference in the size of mini-batches, that we set to 16. More specifically, initial learning rate was set to 0.001 and learning rate decay of 0.5 was used, the momentum for batch normalization was set to 0.9 and the we used the Adam optimization algorithm for back-propagation. Finally, a drop-out regularization with keep probability of 0.5 was applied in the last two fully connected layers.

Ragarding the RadDGCNN approach, radius distance was calculated on the normalized coordinates, R was experimentally set to 0.1. In both models the value of K was set to 20, following the indications provided in [36]. Differently from other studies, we do not use color features (e.g. RGB or HSV) as the synthetic dataset is derived from 3D models, where the colors do not usually match the real color of the architectural elements.

Finally, before feeding a block into the network, we randomly rotate the points around the up direction. In this way we augment the dataset to take into account the fact that architectural elements can have different orientation. While additional data augmentation techniques could be applied to possibly increase the ability of the network in learning complex shapes, a.f. adopting approaches like the one described in [25], we leave this for future investigation.

Two experiments on two different case studies were performed in this study. In the first one we used the TLS point cloud of the courtyard of the Ducal Palace at Urbino to evaluate the trained models, trying to detect 8 different architectural element classes. In order to address a real world scenario, we removed from our training set the synthetic point cloud derived from the BIM model of the Ducal Palace of Urbino. In the second experiment, we evaluated the trained models on the the TLS point cloud of the Palazzo Ferretti. In this case we removed from the training set the BIM model of Palazzo Ferretti. As two of the selected architectural element classes (column and pilaster) are not present in the building, in this case we attempt at recognizing the remaining 6 classes.

Finally, in both experiments we combined the two models implementing a simple multi-classifier system. For each point in the test TLS clouds, we calculate the probability score assigned to each class by the DGCNN and by the RadDGCNN, selecting the class with the higher score.

5 RESULTS AND DISCUSSION

In Table 3, we report the results of the first experiment, obtained on the test scene (Courtyard of the Ducal Palace at Urbino) using DGCNN, RadDGCNN and the combined approach. The results of the second experiment, evaluated on the Palazzo Ferretti, are shown in Table 4. For each class we report precision, recall, F1-score and Intersection over Union (IoU). In both tables, the best results

	DGCNN [36]				RadDGCNN				Combined			
classes	precision	recall	f1-score	IoU	precision	recall	f1-score	IoU	precision	recall	f1-score	IoU
Column	0.8493	0.8387	0.8440	0.7301	**0.9807**	0.8414	**0.9057**	**0.8277**	0.8647	**0.8958**	0.8800	0.7857
Window/Door	0.2530	0.4067	0.3119	0.1848	0.4950	**0.5864**	0.5368	0.3669	**0.5679**	0.5540	**0.5609**	**0.3897**
Wall	0.7458	0.7057	0.7252	0.5689	0.7650	0.8499	0.8052	0.6739	**0.7680**	**0.8590**	**0.8110**	**0.6821**
Pilaster	**0.2066**	**0.0120**	**0.0226**	**0.0114**	0.0626	0.0006	0.0012	0.0006	0.0762	0.0004	0.0007	0.0004
Floor	**0.9957**	0.7221	0.8371	0.7198	0.9107	0.9266	0.9186	0.8494	0.9305	**0.9414**	**0.9359**	**0.8795**
Moulding	0.5057	0.5870	0.5433	0.3730	0.7046	0.6266	0.6633	0.4962	**0.7173**	**0.6409**	**0.6770**	**0.5117**
Vault	0.7094	**0.9378**	0.8077	0.6775	**0.7820**	0.8684	0.8229	0.6991	0.7780	0.8948	**0.8323**	**0.7128**
Other	0.0092	0.0074	0.0082	0.0041	0.0273	**0.0264**	0.0268	0.0136	**0.0377**	0.0220	**0.0278**	**0.0141**
micro avg			0.6714				0.7619				**0.7747**	
macro avg	0.5343	0.5272	0.5125	0.4087	0.5910	0.5908	0.5851	0.4909	**0.5925**	**0.6010**	**0.5907**	**0.4970**
weighted avg	0.6812	0.6714	0.6650		0.7322	0.7619	0.7436		**0.7363**	**0.7747**	**0.7518**	

Table 3. Results of the semantic segmentation of 8 classes obtained on the TLS data from Ducal Palace at Urbino.

	DGCNN [36]				RadDGCNN				Combined			
classes	precision	recall	f1-score	IoU	precision	recall	f1-score	IoU	precision	recall	f1-score	IoU
Window/Door	0.3443	**0.4992**	0.4075	0.2559	**0.4160**	0.4306	0.4232	0.2684	0.4135	0.4772	**0.4431**	**0.2846**
Wall	0.8032	0.7591	0.7805	0.6400	0.8276	0.7826	0.8045	**0.6729**	**0.8408**	**0.8304**	**0.8356**	**0.7175**
Floor	0.9833	**0.7344**	0.8408	0.7254	0.9913	0.7315	0.8418	0.7268	**0.9944**	0.7326	**0.8437**	**0.7296**
Moulding	**0.6016**	0.6829	**0.6397**	**0.4702**	0.4770	0.7063	0.5694	0.3980	0.5558	**0.7146**	0.6253	0.4548
Vault	**0.9282**	0.6348	0.7539	0.6050	0.9145	0.6215	0.7401	0.5874	0.9171	**0.6587**	**0.7667**	**0.6217**
Other	0.1845	**0.2807**	**0.2226**	**0.1253**	0.1273	0.1120	0.1192	0.0634	**0.2127**	0.1878	0.1995	0.1108
micro avg			0.6830				0.6834				**0.7202**	
macro avg	0.6408	0.5985	0.6075	0.4703	0.6256	0.5641	0.5830	0.4528	**0.6557**	**0.6002**	**0.6190**	**0.4865**
weighted avg	0.7284	0.6830	0.6981		0.7197	0.6834	0.6919		**0.7444**	**0.7202**	**0.7259**	

Table 4. Results of the semantic segmentation of 6 classes obtained on the TLS data from Palazzo Ferretti.

are in bold, while underline values indicate the best results between DGCNN and RadDGCNN. Qualitative results are illustrated in Figure 7, where we focused on two fragments of the point clouds of the two historical buildings, showing the original TLS data, the manually annotated point cloud (used as ground truth in the evaluation) and the output of the trained RadDGCNN model (prediction). Finally, the *other* class, including a huge variety of different things (e.g. stairs, rustication and different decorative elements), are clearly difficult to automatically categorize and are often confused with walls, mouldings or windows.

While in the first experiment, RadDGCNN sensibly outperforms the KNN based approach, in the second experiment they provide comparable overall results. In both the experiments, the combination of KNN and radius based networks provide the best overall results, reaching 0.77 and 0.72 of overall accuracy in the segmentation of the Ducal Palace and Palazzo Ferretti, respectively. This suggests that the two neighbouring functions are in fact complementary and can be effectively used together to increase performances. In both cases, floors, walls and vaults are relatively well segmented, with F1-scores ranging from 0.77 to 0.93, and IoU ranging from 0.62 and 0.87. Columns, which are only present in the Ducal Palace, are also precisely detected (0.88 F1-score and 0.78 IoU). Slightly worse results are obtained for windows/doors and mouldings. This was somehow expected as both classes present a high shape variability and thus are more difficult to segment. In fact, as mouldings also includes those framing windows and doors, they are often confused with windows and doors as the two element generally have a similar shape. This is visible also by inspecting the confusion matrices, shown in Figure 8. Results also point out how pilasters are very difficult to detect, as they are almost indistinguishable from walls (see Figure 8 a)), if not for the presence of the capital, which is almost always classified as moulding. Overall, we think that the obtained results demonstrates that the use of synthetic data can effectively drive automatic segmentation of

Fig. 7. Visualization of the result of the automatic segmentation of Ducal Palace at Urbino (*a*) and Palazzo Ferretti (*b*).

TLS point clouds and should be further investigated to support the *Scan-to-BIM* process of historical buildings.

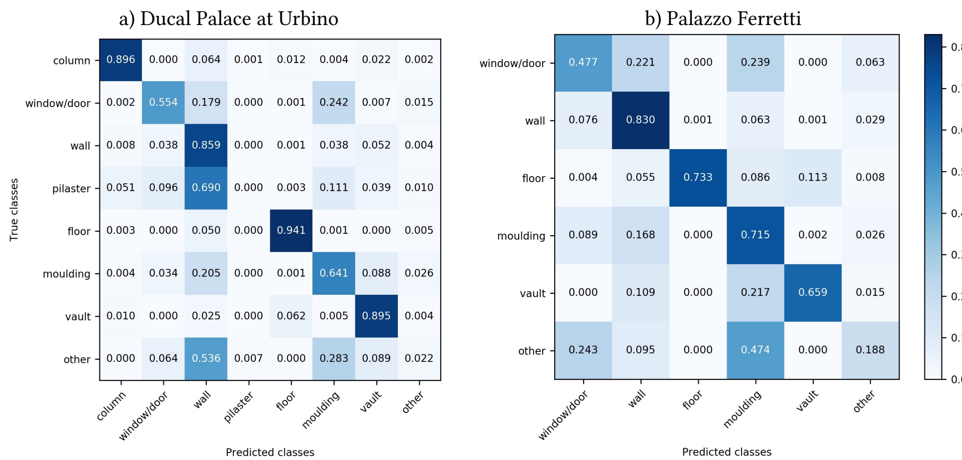

Fig. 8. Confusion matrix of the semantic segmentation of TLS data from Ducal Palace at Urbino (*a*) and Palazzo Ferretti (*b*), obtained with the combined method.

6 CONCLUSIONS AND FUTURE WORKS

The semantic segmentation of point clouds is a relevant task in CH, as it facilitates the automatic identification of historical architectural elements, possibly speeding up the Scan-to-BIM process. Recent improvements in Deep Learning approaches provided encouraging results in semantic segmentation of indoor and outdoor scenes. However, in the context of historical buildings, this task is made particularly challenging by the complexity and high variability of the objects to be detected, and by the lack of available annotated datasets. To tackle this research challenge, in this paper we provided the first study in assessing the use of synthetic data to train Deep Learning models to semantically segment TLS point clouds from historical buildings, with the final aim of helping to automate the *Scan-to-BIM* process, which, in the CH domain, is mainly performed manually. Results obtained on two different case studies, using both a state of the art method (DGCNN) and a novel modified version (RadDGCNN), provide encouraging results, reaching good levels of accuracy. In this study we based on a relatively small synthetic training dataset, and we believe that results could be significantly improved by increasing the amount of synthetic data. This, in turn, potentially provides an incentive to share curated BIM models within the Digital Cultural Heritage community. In future works we plan to put in place additional data augmentation strategies beyond simple random rotation, e.g. using the tool presented in [25]. An other valuable line of research will be the combination of both synthetic and real dataset, in order to enlarge the training set. Finally, further interesting research directions include designing and evaluating more complex edge features, e.g. by considering higher-order relationships between set of points instead of pairwise distance.

ACKNOWLEDGMENTS

This research is partially funded by the CIVITAS (ChaIn for excellence of reflectiVe societies to exploit dIgital culTural heritAge and museumS) project [8]. The authors would like to acknowledge the Professor Paolo Clini for providing the survey data and models of Urbino's Ducal Palace and Palazzo Ferretti, as well as the previous 3D models carried out in the Palladio Library.

REFERENCES

[1] Fabrizio Ivan Apollonio, Simone Baldissini, Paolo Clini, Marco Gaiani, Caterina Palestini, and Camillo Trevisan. 2013. The palladiolibrary geo-models: An open 3D archive to manage and visualize information-communication resources about palladio. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 40, 5W2 (2013), 49–54.

[2] Fabrizio Ivan Apollonio, Marco Gaiani, and Zheng Sun. 2012. BIM-based modeling and data enrichment of classical architectural buildings. *SCIRES-IT-SCIentific RESearch and Information Technology* 2, 2 (2012), 41–62.

[3] Iro Armeni, Ozan Sener, Amir R. Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, and Silvio Savarese. 2016. 3D Semantic Parsing of Large-Scale Indoor Spaces. In *Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, Vol. 2016-December. IEEE Computer Society, Las Vegas, USA, 1534–1543. https://doi.org/10.1109/CVPR.2016.170

[4] Matan Atzmon, Haggai Maron, and Yaron Lipman. 2018. Point Convolutional Neural Networks by Extension Operators. *CoRR* abs/1803.10091 (2018). arXiv:1803.10091 http://arxiv.org/abs/1803.10091

[5] Paul F. Aubin. 2014. *Renaissance Revit: creating classical architecture with modern software.* G3B Press, CreateSpace Independent Publishing Platform.

[6] Carlo Bianchini, Carlo Inglese, Alfonso Ippolito, Daniele Maiorino, and Luca James Senatore. 2017. Building Information Modeling (BIM): Great Misunderstanding or Potential Opportunities for the Design Disciplines? In *Handbook of Research on Emerging Technologies for Digital Preservation and Information Modeling*. IGI Global, 67–90.

[7] Fabio Bianconi, Marco Filippucci, and Federica Magi Meconi. 2018. Parametrical Vitruvius. Generative Modeling of the Architectural Orders. *SCIRES-IT* 8, 2 (2018), 29–48. https://doi.org/10.2423/i22394303v8n2p29

[8] Paolo Clini, Ramona Quattrini, Paolo Bonvini, Romina Nespeca, Renato Angeloni, Raissa Mammoli, Aldo Franco Dragoni, Christian Morbidoni, Paolo Sernani, Maura Mengoni, Alma Leopardi, Mauro Silvestrini, Danilo Gambelli, Enrico Cori, Marco Gallegati, Massimo Tamberi, Fabio Fraticelli, Maria Cristina Acciarri, and Serena Mandolesi. 2020. Digit(al)isation in Museums: Civitas Project - AR, VR, Multisensorial and Multiuser Experiences at the Urbino's Ducal Palace. *Virtual and Augmented Reality in Education, Art, and Museums.* (2020), 194–228. https://doi.org/10.4018/978-1-7998-1796-3.ch011

[9] Haoqiang Fan, Hao Su, and Leonidas J Guibas. 2017. A point set generation network for 3d object reconstruction from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* Institute of Electrical and Electronics Engineers Inc., Honolulu, USA, 605–613.

[10] Marco Gaiani, Fabrizio I Apollonio, Paolo Clini, and Ramona Quattrini. 2015. A mono-instrumental approach to high-quality 3D reality-based semantic models application on the PALLADIO library. In *Proceedings of the 2nd Digital Heritage International Congress Digital Heritage*, Vol. 2. Institute of Electrical and Electronics Engineers Inc., Granada; Spain, 29–36.

[11] Andreas Geiger, Philip Lenz, and Raquel Urtasun. 2012. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* Institute of Electrical and Electronics Engineers Inc., Providence, Rhode Island, 3354–3361. https://doi.org/10.1109/CVPR.2012.6248074

[12] David Griffiths and Jan Boehm. 2019. A Review on deep learning techniques for 3D sensed data classification. *Remote Sensing* 11, 12 (2019), 1499.

[13] David Griffiths and Jan Boehm. 2019. SynthCity: A large scale synthetic point cloud. *arXiv preprint arXiv:1907.04758* (2019).

[14] Pedro Hermosilla, Tobias Ritschel, Pere-Pau Vázquez, Àlvar Vinacua, and Timo Ropinski. 2018. Monte Carlo convolution for learning on non-uniformly sampled point clouds. *ACM Transactions on Graphics* 37, 6 (2018). https://doi.org/10.1145/3272127.3275110

[15] Max Jaderberg, Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. 2014. Synthetic data and artificial neural networks for natural scene text recognition. *arXiv preprint arXiv:1406.2227* (2014).

[16] Vladimir Kniaz, V.S. Gorbatsevich, and Vladimir Mizginov. 2017. Thermalnet: A deep convolutional network for synthetic thermal image generation. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 42, W4 (2017), 41.

[17] Zhuqiang Li, Liqiang Zhang, Ruofei Zhong, Tian Fang, Liang Zhang, and Zhenxin Zhang. 2016. Classification of urban point clouds: A robust supervised approach with automatically generating training data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10, 3 (2016), 1207–1220.

[18] Eva Savina Malinverni, Roberto Pierdicca, Marina Paolanti, Massimo Martini, Christian Morbidoni, Francesca Matrone, and Andrea Lingua. 2019. Deep Learning for Semantic Segmentation of 3D Point Cloud. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 42, 2/W15 (2019), 735–742. https://doi.org/10.5194/isprs-archives-XLII-2-W15-735-2019

[19] Jiří Martínek, Ladislav Lenc, and Pavel Král. 2019. Training Strategies for OCR Systems for Historical Documents. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, Hersonissos, Crete, Greece, 362–373.

[20] Edward F. Miller and Richard A. Melton. 1975. Automated generation of testcase datasets. *ACM SIGPLAN Notices* 10, 6 (1975), 51–58. https://doi.org/10.1145/390016.808424

[21] A Moreira, R Quattrini, G Maggiolo, and R Mammoli. 2018. HBIM Methodology as a Bridge Between Italy and Argentina. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 42, 2 (2018).

[22] G. Myers. 1999. A dataset generator for whole genome shotgun sequencing.. In *Proceedings of the International Conference on Intelligent Systems for Molecular Biology, ISMB*. AAAI, Heidelberg, Germany, 202–210.

[23] R. Nespeca. 2018. Towards a 3D digital model for management and fruition of Ducal Palace at Urbino. An integrated survey with mobile mapping. *SCIRES-IT* 8, 2 (2018), 1–14. https://doi.org/10.2423/i22394303v8n2p1

[24] Yaling Pei and Osmar Zaïane. 2006. *A synthetic data generator for clustering and outlier analysis*. Technical Report. University of Alberta, Edmonton, Canada.

[25] Roberto Pierdicca, Marco Mameli, Eva Savina Malinverni, Marina Paolanti, and Emanuele Frontoni. 2019. Automatic generation of point cloud synthetic dataset for historical building representation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11613 LNCS (2019), 203–219. https://doi.org/10.1007/978-3-030-25965-5_16

[26] Roberto Pierdicca, Marina Paolanti, Francesca Matrone, Massimo Martini, Christian Morbidoni, Eva Savina Malinverni, Emanuele Frontoni, and Andrea Lingua. 2020. Point cloud semantic segmentation using a deep learning framework for cultural heritage. *Remote Sensing* 12, 6 (2020). https://doi.org/10.3390/rs12061005

[27] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. 2017. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Institute of Electrical and Electronics Engineers Inc., Honolulu; United States, 652–660.

[28] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *Advances in neural information processing systems*. 5099–5108.

[29] Ramona Quattrini, Paolo Clini, Romina Nespeca, and Ludovico Ruggeri. 2016. Measurement and Historical Information Building: challenges and opportunities in the representation of semantically structured 3D content. *Disegnarecon* 9, 16 (2016), 14–1.

[30] Ramona Quattrini, Roberto Pierdicca, and Christian Morbidoni. 2017. Knowledge-based data enrichment for HBIM: Exploring high-quality models using the semantic-web. *Journal of Cultural Heritage* 28 (2017), 129–139. https://doi.org/10.1016/j.culher.2017.05.004

[31] Ramona Quattrini, Roberto Pierdicca, Christian Morbidoni, and Eva Savina Malinverni. 2017. Conservation-Oriented HBIM. The BIMEXPLORER Web Tool. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 42, 5W1 (2017), 275–281. https://doi.org/10.5194/isprs-Archives-XLII-5-W1-275-2017

[32] German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, and Antonio M. Lopez. 2016. The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2016-December. IEEE Computer Society, Las Vegas; United States, 3234–3243. https://doi.org/10.1109/CVPR.2016.352

[33] Ha Tran and Kourosh Khoshelham. 2019. Building Change Detection Through Comparison of a Lidar Scan with a Building Information Model. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 42, 2/W13 (2019), 889–893. https://doi.org/10.5194/isprs-archives-XLII-2-W13-889-2019

[34] Fei Wang, Yan Zhuang, Hong Gu, and Huosheng Hu. 2019. Automatic Generation of Synthetic LiDAR Point Clouds for 3-D Data Analysis. *IEEE Transactions on Instrumentation and Measurement* 68, 7 (2019), 2671–2673. https://doi.org/10.1109/TIM.2019.2906416

[35] Tao Wang, David J Wu, Adam Coates, and Andrew Y Ng. 2012. End-to-end text recognition with convolutional neural networks. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*. IEEE, 3304–3308.

[36] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. 2019. Dynamic graph Cnn for learning on point clouds. *ACM Transactions on Graphics* 38, 5 (2019). https://doi.org/10.1145/3326362

[37] Bichen Wu, Xuanyu Zhou, Sicheng Zhao, Xiangyu Yue, and Kurt Keutzer. 2019. SqueezeSegV2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a LiDAR point cloud. In *Proceedings - IEEE International Conference on Robotics and Automation*, Vol. 2019-May. Institute of Electrical and Electronics Engineers Inc., Montreal, Canada, 4376–4382. https://doi.org/10.1109/ICRA.2019.8793495

[38] Zexiang Xu, Kalyan Sunkavalli, Sunil Hadap, and Ravi Ramamoorthi. 2018. Deep image-based relighting from optimal sparse samples. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–13.

Received N/A; revised N/A; accepted N/A