

# Application of Semantic Segmentation of Clouds of Points for Preservation of Cultural Heritage

Nataliya Boyko<sup>1</sup> and Mariia Rizhko<sup>1</sup>

<sup>1</sup>Lviv Polytechnic National University, Profesorska Street 1, Lviv, 79013, Ukraine

## Abstract

Artificial intelligence is evolving and emerging in many new areas, but a literature analysis suggests that 3D and AI-based technologies to monitor cultural heritage have not been studied enough. Cultural heritage requires non-stop detailed observation and protection as buildings become older and ruin through time. This process is critical and requires much time, financial and human resources. Since it is not always possible to provide these resources, Computational and Information Technologies are needed to build a risk-informed system that will analyze and notify about cultural heritage changes in a timely manner. Therefore, the contribution of this document is potentially essential for this area. The study examines ArCH datasets and best techniques for segmenting 3D point clouds: Point-wise MLP with PointNet, PointNet ++ and RandLA-Net, Point Convolution with PointCNN, RNN-based with RSNet, Graph-based with DGCNN. The paper examines the efficiency of the semantic segmentation models PointNet, PointNet ++, RandLA-Net, PointCNN, RSNet, DGCNN on S3DIS, ScanNet, Semantic3D, and SemanticKITTI datasets. The efficiency of semantic segmentation models PointNet, PointNet ++, RandLA-Net, PointCNN, RSNet, DGCNN on S3DIS, ScanNet, Semantic3D, and SemanticKITTI datasets are compared.

## Keywords

Artificial Intelligence, Point Cloud, Semantic Segmentation, Monitoring, Cultural Heritage, Risk-Informed Systems, Information Technologies

## 1. Introduction

Cultural heritage plays a vital role in preserving the memory and knowledge of the past. Moreover, its preservation is essential in developing modern infrastructure, constructing new cities, roads, and railways. At the same time, we must not forget about the development of tourist services, an adaptation of old buildings to modern needs, illegal archeological excavations, and other potential risks related to destroying cultural heritage.

Preserving cultural heritage has three main risks. First of all, it is a time-consuming process that is required to be done repeatedly. If it is impossible to do so, cultural heritage will get damaged and require renovation, or in some cases, it can even be lost forever. The second concern is financial resources. Non-stop monitoring takes much time and human resources; thus, it takes

---

<sup>1</sup>CITRisk'2021: 2nd International Workshop on Computational & Information Technologies for Risk-Informed Systems, September 16–17, 2021, Kherson, Ukraine

EMAIL: nataliya.i.boyko@lpnu.ua (N.Boyko); mariia.rizhko.knm.2018@lpnu.ua (M.Rizhko)

ORCID: 0000-0002-6962-9363 (N.Boyko); 0000-0003-3885-4661 (M.Rizhko)



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

much money. No monitoring results in cultural heritage damage, and restoration takes even more money. The last but not least risk are people working with cultural heritage. They do a monotonous job checking cultural heritage for destruction. Instead, they could spend time on research and renovation tasks.

Nowadays, cultural heritage monitoring is managed by cultural organizations, which are constantly confronted with a large amount of data that needs to be processed and small resources that they can use. The solution is to create a risk-informed system to automate data monitoring and analysis based on 3D and AI technologies. By automating the processes in collecting and analyzing information, it is possible to achieve significant cost savings, both human and financial.

The work aims to systematize approaches to directions and technologies of 3D and AI technologies to analyze and recognize cultural heritage, developing a system for practical application.

The solution of the following tasks is required:

1. Review of existing 3D and AI solutions for monitoring and analysis of cultural heritage preservation.
2. Research of requirements, methods, and algorithms to get a decision for the task.
3. Select and collect the necessary data of cultural heritage to be analyzed.
4. Development of architecture for monitoring and analysis of cultural heritage preservation.
5. Creating an application program - a system of semantic segmentation for cultural heritage.

The study's practical importance is to create a new risk-informed system for functional, objective, and cost-effective monitoring of cultural heritage changes that will monitor many facilities and act quickly and promptly.

## **2. Related Works**

Risk analysis is one of the essential tools for preserving cultural heritage. It is used for decision-making in the process of cultural heritage asset management and maintenance. For this purpose, both quantitative and qualitative analysis is used [21].

Although risk categorization plays a vital role in risk management in other disciplines, it has yet to be successfully applied to cultural heritage studies [22].

The importance of information searching and systematizing in the modern world has led to the thematic modeling of text document collections in this study. Therefore, thematic models are used to identify trends in scientific publications or news streams for classifying and categorizing image documents and video streams, information retrieval, including multilingual, tagging web pages, detecting spam, recommendation, and other applications.

3D scanning - building a computer model of a material object. It is studied by many researchers [5-8]. Currently, there are two leading technologies of 3D scanning - laser scanning and photogrammetry.

Laser scanning is a technology for obtaining information about terrain and objects using a laser. This method has been studied by scientists [5-8]. The result of the laser scan is a cloud of laser reflection points.

There are two types of laser scanning: mobile and stationary. During mobile scanning, continuous measurement is performed while the vehicle is driving. The device is established

motionlessly during stationary scanning, and measurement is carried out from several standing points.

Photogrammetry is a science that studies the appearances, shapes, and positions of various objects in space, objects, and their shapes by measuring their photographic image. It was studied by researchers [5-8].

No special devices are required to use this method. It is enough to have a camera on a modern phone.

When choosing photogrammetry for 3D scanning, an important question is what and how affects the accuracy of 3D models. Accuracy can depend on many factors [1]: optical and digital camera characteristics, spatial distribution, and ground control points.

Also, shooting with the help of remote control systems, such as drones, has been studied. This method has an advantage over standard photogrammetry due to the aerial view [2].

In recent years, interest in preserving cultural heritage has begun to grow, so more and more data is being digitized, which is very important for artificial intelligence, as the training of models is based on data. However, collecting a large enough amount of data is still a problem because it is time-consuming and requires a human factor to mark the correct elements.

Machine learning technologies have become popular not only in computer science but in other fields as well. One of the reasons for the growth, in particular, is the successful application of deep learning methods for image classification [3], in which convolutional neural networks (CNN) exceed the human ability to analyze objects [4].

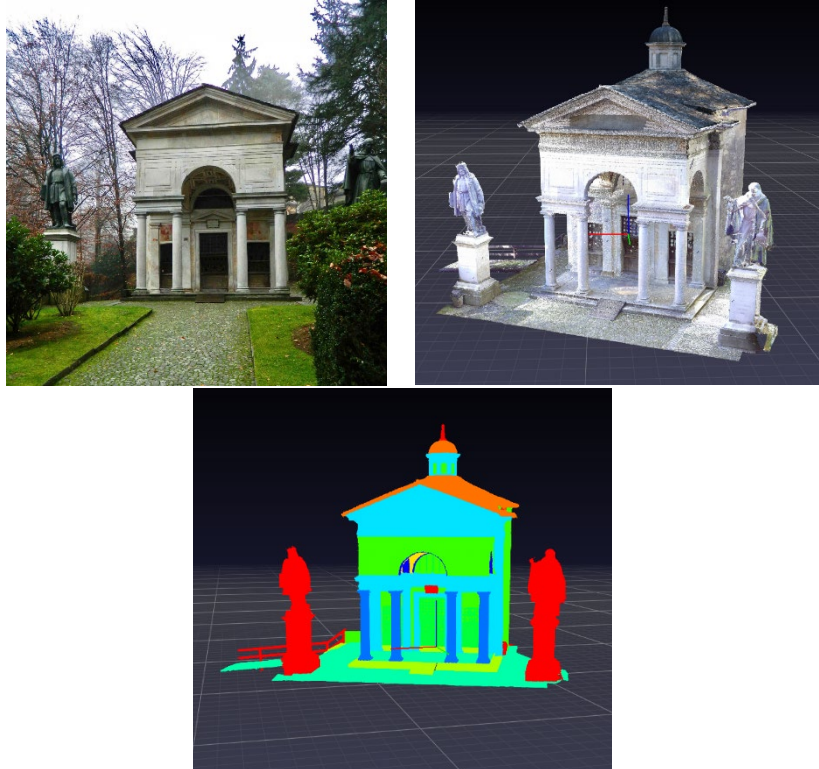
The potential of deep learning technologies for three-dimensional image analysis achieved a remarkable breakthrough in 2012 when the AlexNet model [5] showed excellent analysis results during the ImageNet competition. In 2014, GoogLeNet [6] won the ImageNet competition, achieving 93.3% accuracy of semantic segmentation. In 2016, Microsoft Networks ResNet [7] won the ImageNet competition, achieving 96.4% accuracy.

### **3. Materials and Methods**

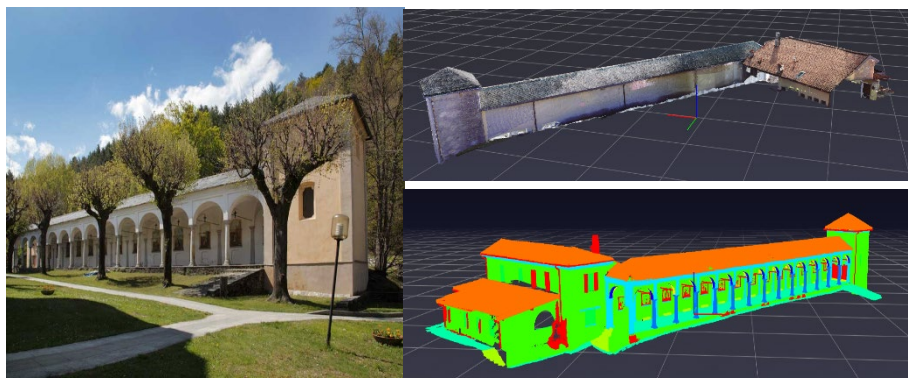
Learning on point clouds is attracting more and more attention with the development of augmented and virtual reality, their wide application in computer vision, autopilot, robot development. Deep learning is well researched to solve 2D problems, but for 3D cultural heritage data, it is only evolving and needs further research and the creation of new datasets to train effective models.

In this paper, 3D data are presented in point clouds, and the task is their semantic segmentation. For this purpose, the rights to use the ArCH dataset (Architectural Cultural Heritage point clouds for classification and semantic segmentation) were obtained. [8] The dataset consists of 17 annotated scenes, each point of which belongs to one of 9 classes: "arch": 0, "column": 1, "moldings": 2, "floor": 3, "door\_window": 4, "wall": 5, "stairs": 6, "vault": 7, "roof": 8, "other": 9. Some of these scenes belong to the UNESCO heritage. Others are part of the historical heritage and represent different historical periods and architectural styles.

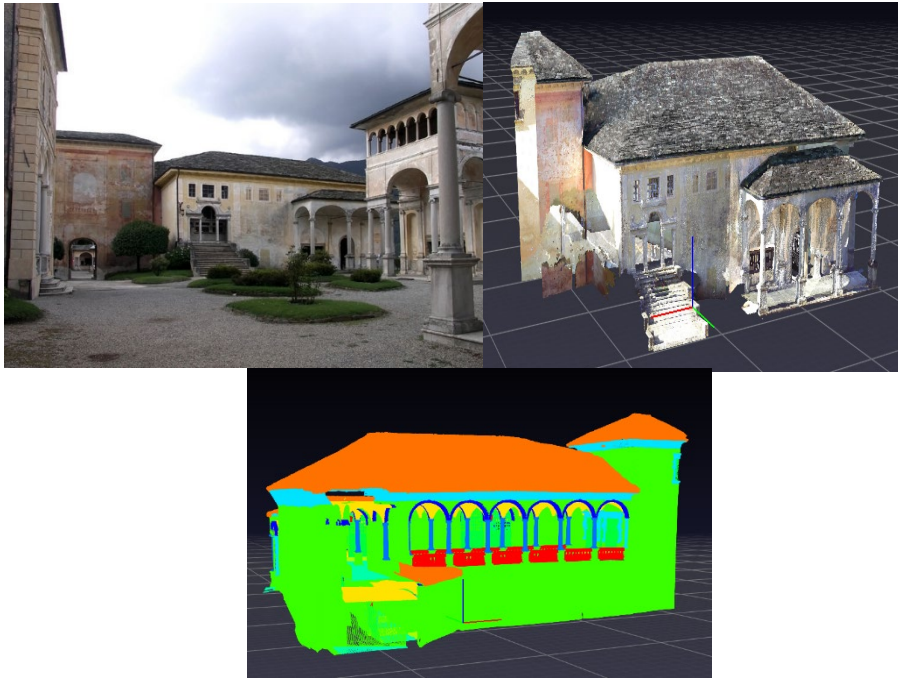
Fifteen scenes are used for training and two for testing models. Scenes for training include churches, chapels, porticos, loggias, pavilions, and monasteries. Two test scenes have different characteristics. The first represents a simple, almost symmetrical one-level building with standard and repetitive geometric elements. The second scene represents a complex, asymmetrical building with two levels, shot both inside and outside, with different types of vaults, stairs, and windows.



**Figure 1:** Photography, 3D color image, and semantic segmentation of a training object



**Figure 2:** Photography, 3D color image, and semantic segmentation of the first testing object



**Figure 3:** Photography, 3D color image, and semantic segmentation of the second testing object

Figures 1-3 show a visualization of one of the training objects and two objects used to test the quality of the models. Objects are represented as clouds of points with corresponding r / g / b values for each point to indicate color and class. Data were obtained using various sensors (cameras, scanners) and platforms (UAV and others). Preprocessing included spatial translation, subsampling, and feature selection.

**Table 1**

Information about training objects

Name	Number of points	Scene	Getting data	Class number	Subsampling (cm)
1_TR_cloister	15,740,229	Indoor/ Outdoor	TLS + UAV	8/9	1
2_TR_church	20,862,139	Indoor	TLS	8/9	1
3_VAL_room	4,188,066	Indoor	TLS	6/9	1
4_CA_church	4,850,807	Outdoor	TLS + UAV	6/9	1
5_SMV_chapel_1	3,783,412	Outdoor	TLS + UAV	9/9	1
6_SMV_chapel_2t	6,326,871	Indoor/Outd oor	TLS + UAV	9/9	1
7_SMV_chapel_24	3,571,064	Outdoor	TLS + UAV	9/9	1
8_SMV_chapel_28	3,156,753	Outdoor	TLS + UAV	9/9	1
9_SMV_chapel_10	2,193,189	Indoor/Outd oor	TLS + UAV	6/9	1

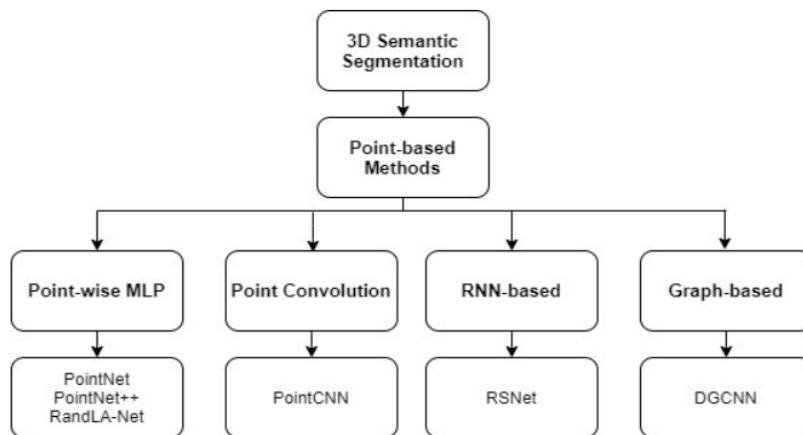
10_SStefano_portico_1	3,783,699	Outdoor	Terrestrial photogrammetry	8/9	1
11_SStefano_portico_2	10,047,392	Outdoor	Terrestrial photogrammetry	8/9	1
12_KAS_pavillion_1	598,384	Indoor/Outdoor	TLS	4/9	1
13_KAS_pavillion_2	325,822	Indoor/Outdoor	TLS	4/9	1
14_TRE_square	9,409,239	Outdoor	Terrestrial photogrammetry	8/9	1.5
15_OTT_church	13,302,903	Indoor/Outdoor	TLS	9/9	1.5

**Table 2**  
Information about testing objects

Name	Number of points	Scene	Getting data	Class number	Subsampling (cm)
A_SMG_portico	17,798,012	Outdoor	TLS + UAV	9/9	1
B_SMV_chapel_27to35	16,200,442	Indoor/Outdoor	TLS + UAV	9/9	1

Tables 1 and 2 provide more information about training and testing objects. The total number of points for training is 102,139,969, for testing 33,998,454.

Point-based Networks are used for the segmentation (Fig. 4).



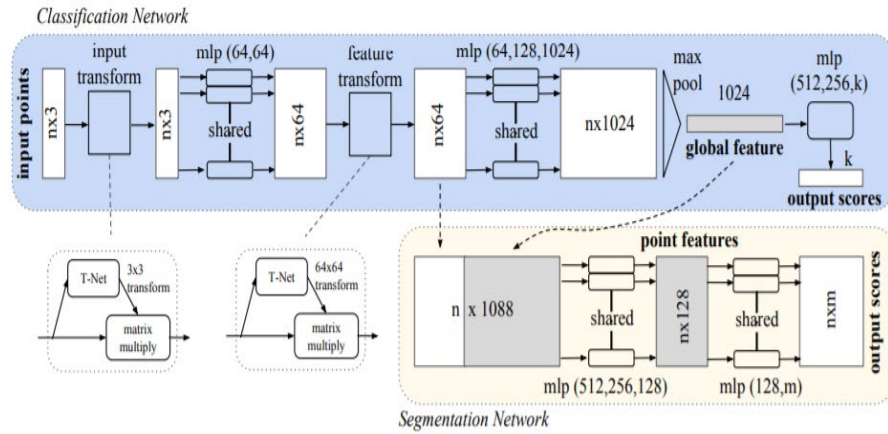
**Figure 4:** Point-based methods for clouds of points

The semantic segmentation task is to divide a cloud of points into parts according to the semantic meaning of the points. This section will describe the best semantic segmentation techniques nowadays: Point-wise MLP on PointNet [9], PointNet ++ [10] and RandLA-Net [11], Point Convolution on PointCNN [12], RNN-based on RSNet [13], Graph-based on DGCNN [14].

Pointwise MLP methods typically use common MLP (Multi-Layer Perceptron) as the central unit in their network for its high efficiency. However, point functions obtained with MLP cannot cover local geometry in point clouds and interactions between points. Therefore, various methods have been proposed, including PointNet, PointNet ++, and RandLA-Net, to provide more context for each point and explore deeper local structures.

Convolutional networks require highly structured data to obtain scales and other optimizations. Because the point cloud is not standard, the data must be transformed into a voxel grid or image collection before transmitting it for learning. However, this transformation makes the obtained data excessively voluminous and can also change the nature of the data. That is why PointNet accepts a cloud of points without transformations.

The PointNet architecture (Fig. 5) consists of three main modules: a max-pooling layer as a symmetric function for aggregating information from all points, a structure for combining local and global information, two networks for combining entry points and point features.



**Figure 5:** PointNet architecture

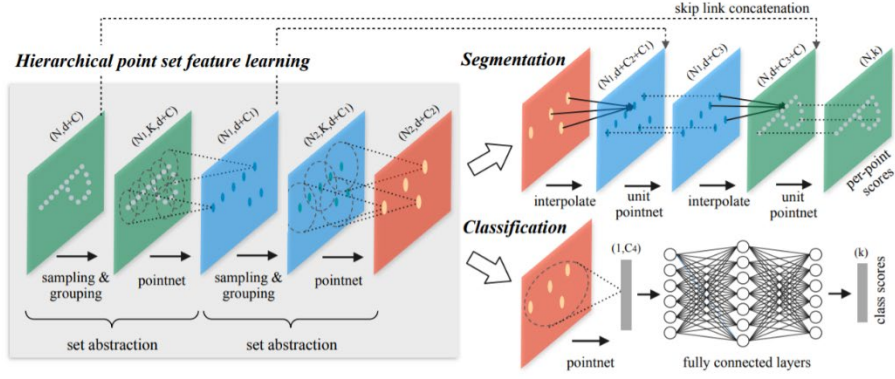
The idea of this model is to approximate the general function defined on the set of points through the application of a symmetric function on the transformed elements in the network (Formula 1):

$$f(\{x_1, \dots, x_n\}) \approx g(h(x_1), \dots, h(x_n)), \quad (1)$$

where  $f: 2^{\mathbb{R}^N} \rightarrow \mathbb{R}$ ,  $h: \mathbb{R}^N \rightarrow \mathbb{R}^K$ ,  $g: \mathbb{R}^K \times \dots \times \mathbb{R}^K \rightarrow \mathbb{R}$  – is a symmetric function.

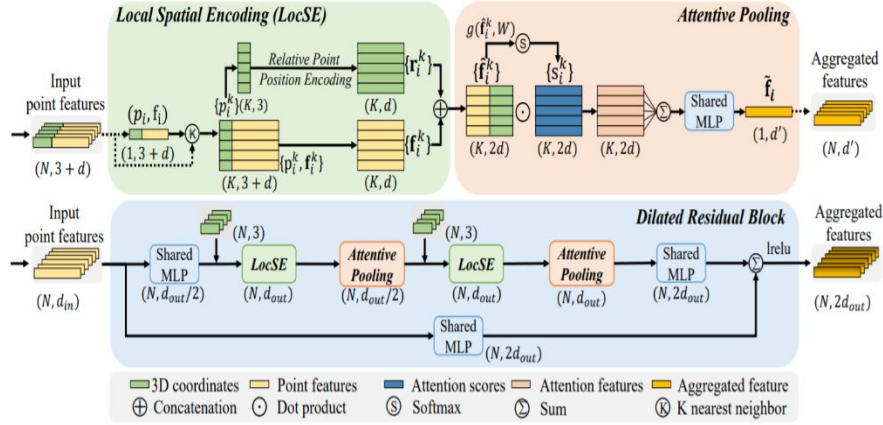
H is approximated through the MLP network and g through the composition by a function of one variable and max pooling function.

PointNet does not cover local structures because of the metric space in which the points are located, limiting the ability to recognize small patterns and generalize complex scenes. PointNet ++ is a hierarchical neural network that applies PointNet recursively to a set of entry points. Using distances, PointNet ++ can study local features with increasing contextual scale (Fig. 6).



**Figure 6:** PointNet++ architecture

RandLA-Net is a lightweight neural architecture (Figure 7) that can handle large-scale point clouds 200 times faster than other architectures, as most existing architectures use time-consuming preprocessing and post-processing techniques. PointNet is computationally efficient but does not capture the contextual information of each point. RandLA-Net handles large 3D point clouds in one pass without requiring any pre/post-processing steps, such as voxelization, block separation, or graphing. RandLA-Net relies only on random sampling within the network and therefore requires much less memory and computation.



**Figure 7:** RandLA-Net architecture

The first step, Local Spatial Encoding (LocSE), is finding adjacent points. For each point, its neighboring points are searched for by a simple K-nearest neighbors (KNN) algorithm based on the point-wise Euclidean distance.

The next step is Relative Point Position Encoding. For each nearest  $K$  point  $\{p_i^1 \dots p_i^k \dots p_i^K\}$  of the central point  $p_i$ , the following is considered (Formula 2):

$$r_i^k = MLP(p_i \oplus p_i^k \oplus (p_i - p_i^k) \oplus \|p_i - p_i^k\|), \quad (2)$$



where  $p_i$  and  $p_i^k$  - is x-y-z coordinates of points,  $\oplus$  - concatenation operation,  $\| \|$  - calculates the Euclidean distance between adjacent and central points.

The last step in this part is Point Feature Augmentation. For each adjacent point  $p_i^k$  corresponding to  $r_i^k$  concatenation is made with the corresponding point features  $f_i^k$  and the resulting vector  $\hat{f}_i^k$  is received.

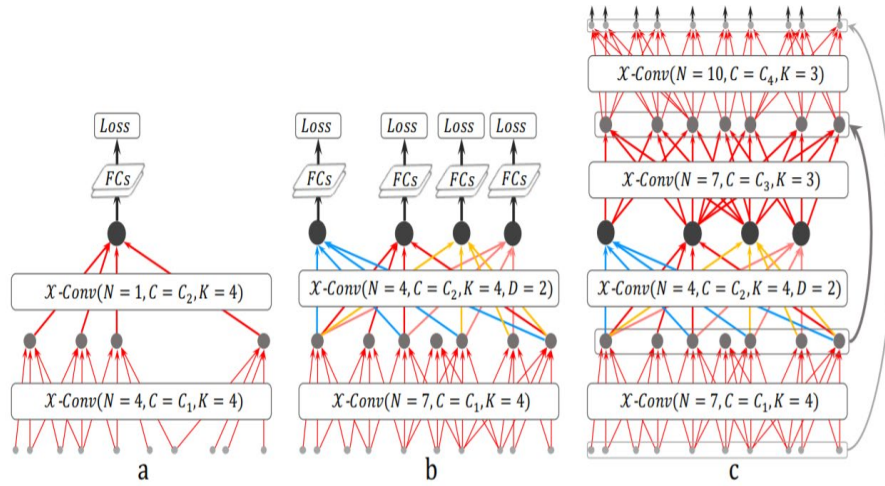
The next step is Attentive Pooling, which consists of Computing Attention Scores and Weighted Summation. Computing Attention Scores (Formula 3):

$$s_i^k = g(f_i^k, W), \quad (3)$$

where  $W$  - is MLP weights

$$\text{Weighted Summation: } \bar{f}_i = \sum_{k=1}^K \hat{f}_i^k \cdot s_i^k.$$

Point Convolution uses spatial-local correlation in data presented densely in grids and provides a basis for studying features from point clouds. One example of such an architecture is PointCNN (Figure 8).



**Figure 8:** PointCNN architecture for classification (a i b) and segmentation (c)

PointCNN learns the transformation of input points to weigh the input features associated with the points and rearrange the points in the canonical order. The PointCNN architecture contains two designs: Hierarchical Convolution and  $\chi$ -Conv Operator.

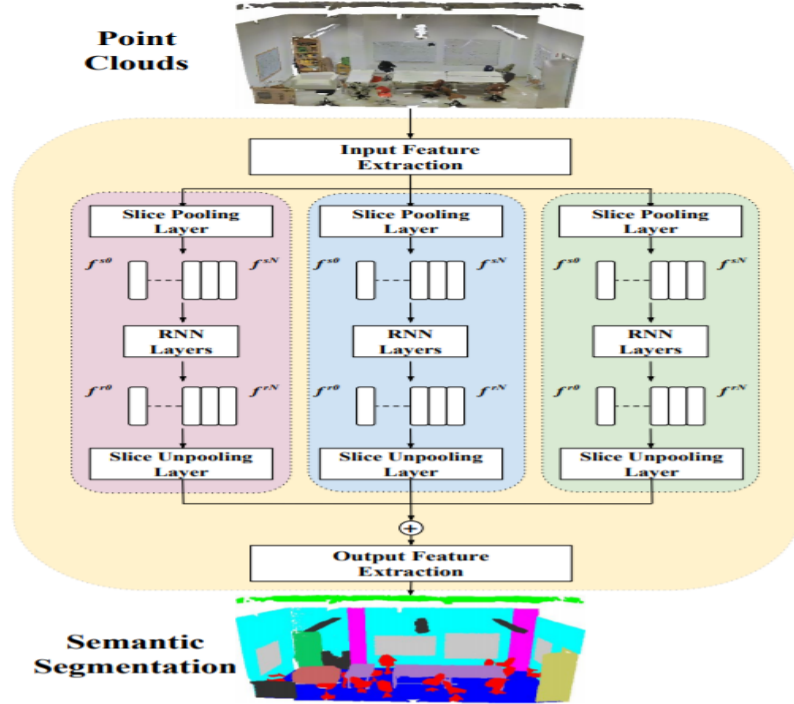
Hierarchical Convolution is recursively applied to local parts of the grid, often reducing data to fewer representative points but with more saturated information.

The  $\chi$ -Conv operator works in local parts, accepts connected points as input data, and makes convolution. Neighboring points are transformed into local coordinate systems of representative points, and later these local coordinates are individually combined with the corresponding features (Formula 4).

$$F_p = \chi - \text{Conv}(K, p, P, F) = \text{Conv}(K, \text{MLP}(P - p) \times [\text{MLP}_\sigma(P - p), F]) \quad (4)$$

where  $\text{MLP}_\sigma$  is used separately for each point, as in PointNet.

Most other semantic networks do not model the necessary relationships between point clouds. RNN-based models will be presented on the example of RSNet (Fig. 9). A key component of the RSNet architecture is a highly efficient module of local dependence between points. RSNet takes as input clouds of not preprocessed points and returns semantic labels for each of them.



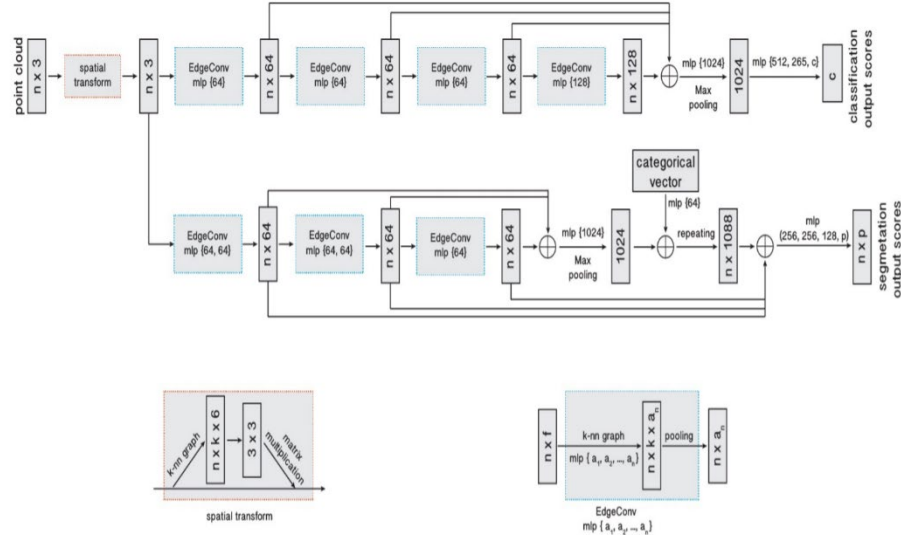
**Figure 9:** RSNet diagram

Input and output feature blocks are used to generate features independently. In the middle of them, the local dependency module is located. The input function block receives entry points and generates attributes, and the output blocks receive processed input attributes and return final predictions for each point. Both blocks use a sequence of multiple layers to create independent representations of the features for each point. The local dependency module combines an aggregation layer, a bidirectional recurrent neural network (RNN) layer, and a separation layer. The problem of the local context is solved first by projecting disordered points on ordered features and then applying traditional learning algorithms.

Graph-based networks are used to capture the shapes and geometric structures of three-dimensional point clouds. First, a point cloud is represented as a set of simple interconnected shapes and super points, then a graph of super points is used to capture the structure and context of information. After that, the large-scale problem of cloud point segmentation is divided into three subtasks: geometrically homogeneous distribution, the embedding of super points, and context segmentation.

One example of a Graph-based architecture is DGCNN (Figure 10). DGCNN is an EdgeConv that is suitable for CNN complex point cloud tasks, including classification and segmentation.

EdgeConv operates on graphs that are dynamically computed at each level of the network. It covers the local geometric structure, preserving the invariance of the permutation. Instead of generating point features directly from their embeddings, EdgeConv generates edge features that describe the relationships between a point and its neighbors. EdgeConv is designed to be invariant to the ordering of neighbors and therefore is an invariant of permutation.



**Figure 10:** DGCNN architecture

Because EdgeConv creates a local graph and learns embeddings for edges, the model can group points in Euclidean and semantic space. Instead of working on individual points, as in PointNet, DGCNN uses local geometric structures to construct a local graph of adjacent points and apply operations on the edges connecting adjacent pairs of points.

## 4. Experiments

The results of the analyzed methods in the previous section are compared on the datasets S3DIS [15], ScanNet [16], Semantic3D [17], and SemanticKITTI [18]. For this purpose, mean class accuracy (mAcc), overall accuracy (oAcc), and mean class intersection over union (mIoU) metrics are used. The data are taken from the articles of the corresponding algorithms and datasets.

**S3DIS:** all point clouds are obtained without manual intervention using a Matterport scanner. Dataset consists of 271 rooms, which belong to 6 large-scale internal scenes from 3 different buildings, with 6020 sq.m. These areas mainly include offices, training and exhibition spaces, conference rooms.

**ScanNet:** annotations contain expected calibration parameters, camera positions, three-dimensional surface reconstructions, textured grids, dense object-level semantic segmentation, CAD models. The dataset contains annotated RGB-D environment scans. In total, there are 2.5M images in 1513 scans obtained in 707 different locations.

**Semantic3D:** includes about 4 billion 3D points obtained using static ground-based laser scanners, covering up to 160x240x30 meters of space. Point clouds belong to 8 classes (urban and rural) and contain coordinates, RGB information, and intensity.

SemanticKITTI: an extensive outdoor dataset containing a detailed point annotation of 28 classes. The dataset contains labels for the whole horizontal 360-degree field of view of the rotating laser sensor.

**Table 3**

Effectiveness evaluation of semantic segmentation models on S3DIS, ScanNet, Semantic3D, and SemanticKITTI datasets

	S3DIS		ScanNet		Semantic3D		SemanticKITTI	
	mAcc	mIoU	oAcc	mIoU	oAcc	mIoU	mAcc	mIoU
PointNet	48.98	41.09	-	14.69	-	-	29.9	17.9
PointNet++	-	50.04	71.4	34.26	-	-	-	-
RandLA-Net	-	-	-	-	94.8	77.4	-	53.9
PointCNN	63.86	57.26	85.1	45.8	-	-	-	-
RSNet	59.42	56.5	-	39.35	-	-	-	-
DGCNN	-	56.1	-	-	-	-	-	-

Table 3 shows the results achieved in the original articles of methods and datasets. As can be seen from the results presented in this table, the best metrics in the RandLA-Net model are on the Semantic3D and SemanticKITTI datasets, in the PointCNN model - on the S3DIS, ScanNet datasets. However, there are quite a few unknown results on various datasets. Accordingly, we can assume that the RandLA-Net or PointCNN models will work best on the ArCH dataset. However, due to the omitted values, it may turn out that the other models will still be better than the previous two.

## 5. Results

The experiments part demonstrates semantic segmentation models' performance on S3DIS, ScanNet, Semantic3D, and SemanticKITTI datasets. The presented models were of different types: PointNet, PointNet ++, and RandLA-Net are Point-wise MLP models, PointCNN is Point Convolution, RSNet is RNN-based, DGCNN is Graph-based. Furthermore, these methods were tested on different types of environment: offices, training and exhibition spaces, conference rooms, cities and towns, open and closed space.

The results were shown in Table 1. They show that different models were better on different datasets for different metrics. So for the S3DIS dataset, the best model is PointCNN with mAcc equal to 63.86 while PointNet has mAcc equal to 48.98, RSNet - 59.42. Also, PointCNN shows the best results on mIoU metric, which is equal to 57.26, while it is 41.09 in PointNet, 50.04 in PointNet++, 56.5 in RSNet, and 56.1 in DGCNN.

For the ScanNet dataset, the best model is also PointCNN with oAcc equal to 85.1, while it is 71.4 in PointNet++. Also, PointCNN shows the best results on the mIoU metric, which is equal to 45.8 as opposed to 14.69 in PointNet, 34.26 in PointNet ++, 39.35 in RSNet.

For the Semantic3D dataset, the best model is RandLA-Net, which shows high results with oAcc = 94.8 and mIoU = 77.4. In the two previous datasets, the maximum value of oAcc was 85.1 and mIoU 57.26.

The SemanticKITTI dataset is also poorly researched. The mAcc metric is shown for PointNet only and is 29.9. The mIoU metric is presented for PointNet and RandLA-Net and is 17.9 and 53.9, respectively.

Therefore, if we compare the presented methods PointNet, PointNet ++ and RandLA-Net, PointCNN, RSNet, and DGCNN for the S3DIS, ScanNet, Semantic3D, and SemanticKITTI datasets, we can get the following conclusions. Further research is needed on the methods presented on the relevant datasets, as not all possible options have been considered in the official articles in which the models and datasets were first presented. Further research should be performed on the same metrics, finding mAcc, oAcc, and mIoU in all combinations of models and datasets. After the same metrics are received, it will be possible to compare which datasets on which models work best. The last step will be to check the models presented on the ArCH dataset against the same metrics mAcc, oAcc, and mIoU.

## 6. Conclusion

The protection of cultural heritage in the urbanization epoch and city development time is critical for preserving history. However, this is quite an enormous task with many risks connected to time, financial, and human resources. Therefore, a solution for automating the monitoring and analysis of data by creating semantic segmentation of point clouds was presented. A risk-informed system based on computational and information technologies will reduce risks and increase the efficiency of using these resources.

The existing solutions were considered, the methods and datasets that correspond to the goal were analyzed, and their results on different metrics were collected and analyzed.

The following steps in continuing this study will be: conducting experiments on the presented methods for the respective datasets; comparison of experimental results on the same metrics; verification of the presented methods on the ArCH dataset.

## References

- [1] M.Bolognesi, A.Furini, V.Russo, A.Pellegrinelli, P.Russo, Accuracy of cultural heritage 3Dmodels by RPAS and terrestrial photogrammetry, *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch*, 40, 2014, pp. 113–119
- [2] E.Karachaliou, E.Georgiou, D.Psaltis, E.Stylianidis, UAV for mapping historic buildings: From 3D modeling to BIM, *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, XLII-2, 2019, pp. 397–402
- [3] A.Krizhevsky, I.Sutskever, G.Hinton, Imagenet classification with deep convolutional neural networks. *NIPS*, 1, 2012, pp. 1097–1105
- [4] F.Radenović, G.Tolias, O.Chum, CNN image retrieval learns from BoW: Unsupervised finetuning with hard examples. *Eur. Conf. Comput. Vis. ECCV*, 2016, pp. 1–17
- [5] A.Krizhevsky, I.Sutskever, G.Hinton, Imagenet classification with deep convolutional neural networks. *NIPS* 1, 2012, pp. 1097–1105
- [6] C.Szegedy, W.Liu, Y.Jia, P.ermanet, S.Reed, D.Anguelov, D.Erhan, V.Vanhoucke, A.Rabinovich, Going Deeper with Convolutions, *CVPR*, 2015, p. 1–9
- [7] K.He, X.Zhang, S.Ren, J.Sun, Deep Residual Learning for Image Recognition, 2015, 1–12
- [8] F.Matrone, A.Lingua, R.Pierdicca, E. S.Malinverni, M.Paolanti, E.Grilli, F.Remondino, A.Murtiyoso, T.Landes, A benchmark for large-scale heritage point cloud semantic

- segmentation, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B2-2020, 2020, pp. 1419–1426
- [9] R.Q.Charles, S.Hao, M.Kaichun, J.G.Leonidas, Deep learning on point sets for 3d classification and segmentation, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660
- [10] Ch.R.Qi, L.Yi, H.Su, J.G.Leonidas, Pointnet++: Deep hierarchical feature learning on point sets in a metric space, *Advances in neural information processing systems*, 30, 2017, pp. 5099–5108
- [11] Q.Hu, B.Yang, L.Xie, S.Rosa, Y.Guo, Z.Wang, N.Trigoni, A.Markham, RandLA-Net: Efficient semantic segmentation of large-scale point clouds, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11108–11117
- [12] Y.Li, R.Bu, M.Sun, W.Wu, X.Di, B.Chen, Pointcnn: Convolution on x-transformed points. *Advances in neural information processing systems*, 31, 2018, pp. 820–830
- [13] Q.Huang, W.Wang, U.Neumann, Recurrent slice networks for 3d segmentation of point clouds, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 2626–2635
- [14] Y.Wang, Y.Sun, Z.Liu, S.E.Sarma, M.Bronstein, M.Justin, Dynamic graph cnn for learning on point clouds. *Acm Transactions On Graphics (tog)*, 38, 5, 2019, pp. 1–12
- [15] I.Armeni, O.Sener, A.R.Zamir, H.Jiang, I. Brilakis, M. Fischer, S. Savarese, 3d semantic parsing of large-scale indoor spaces, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1534–1543
- [16] A.Dai, A.Chang, M.Savva, M.Halber, T.Funkhouser, Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5828–5839
- [17] T.Hackel, N.Savinov, L.Ladicky, J.D.Wegner, K.Schindler, M.Pollefeys, Semantic3d. net: A new large-scale point cloud classification benchmark, *arXiv preprint arXiv:1704.03847*, 2017
- [18] J.Behley, M.Garbade, A.Milioto, J.Quenzel, S.Behnke, C.Stachniss, J.Gall, SemanticKITTI: A dataset for semantic scene understanding of lidar sequences, in: *Proceedings of the IEEE International Conference on Computer Vision*. IEEE, 2017, pp. 9297–9307,
- [19] N.Boiko, The issue of access sharing to data when building enterprise information model, in: *IX International Scientific and Technical conference, Computer science and information technologies (CSIT 2014)*, Lviv, Ukraine, 2014, pp. 23-24
- [20] N.Boyko, R.Hlynka, Application of Machine Algorithms for Classification and Formation of the Optimal Plan, in: *Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021)*, Vol. 1: Main Conference Lviv, Ukraine, April 22-23, 2021, pp. 1853-1865.
- [21] V.Rajcic, Risks and resilience of cultural heritage assets, in: *International Conference: Europe and the Mediterranean: Towards a Sustainable Built Environment At: Malta*, Vol. 1, 2016. [https://www.researchgate.net/publication/299395298\\_Risks\\_and\\_resilience\\_of\\_cultural\\_heritage\\_assets](https://www.researchgate.net/publication/299395298_Risks_and_resilience_of_cultural_heritage_assets)
- [22] R.Sharifi, Risk Characterization for Preserving Cultural Heritage Assets, 2016. [https://www.chnt.at/wp-content/uploads/eBook\\_CHNT22\\_Sharifi.pdf](https://www.chnt.at/wp-content/uploads/eBook_CHNT22_Sharifi.pdf)