The contribution of Deep Learning to the semantic segmentation of 3D point-clouds in urban areas

Z.BALLOUCH[1], R.HAJJI[1], M.ETTARID[1]
College of Geomatic Sciences and Surveying Engineering
IAV Hassan II
Rabat, Morocco
Corresponding authors, e-mail: ballouchzouhair@gmail.com, r.hajji@iav.ac.ma, m.ettarid@iav.ac.ma

Abstract— **Semantic segmentation in a large-scale urban environment is crucial for a deep and rigorous understanding of urban environments. The development of Lidar tools in terms of resolution and precision offers a good opportunity to satisfy the need of developing 3D city models. In this context, deep learning revolutionizes the field of computer vision and demonstrates a good performance in semantic segmentation. To achieve this objective, we propose to design a scientific methodology involving a method of deep learning by integrating several data sources (Lidar data, aerial images, etc) to recognize objects semantically and automatically. We aim at extracting automatically the maximum amount of semantic information in a urban environment with a high accuracy and performance.**

Keywords: Lidar, Deep Learning, Semantic Segmentation, Urban environment

I. INTRODUCTION

Currently, there is an increasing need for representing and analyzing the third dimension of urban space. Beyond exploration and visualization purposes, 3D models are useful for many urban applications such as efficient urban planning [1], construction and simulations (noise, energy, emergency, etc). They also provide interesting support in many other fields such as mobile telephony, where engineers can determine the coverage areas of the network with propagation models, as well as in archaeology and cultural heritage, where they allow the documentation and conservation of sites and monuments. 3D models are also necessarily applied in civil engineering, for the production of realistic scenes when designing major construction projects.

3D models are basically created using satellite imagery, photogrammetry or Lidar technology. Recent development in the field of airborne Lidar technology allows acquiring data in high spatial resolution and accuracy, where the main advantage of Lidar technology is the speed of data acquisition and the large amount of data provided by this system.

Two main steps are essential to automatically build a 3D city model from Lidar data: 1) segmentation of the 3D point cloud and 2) 3D modeling of the resulting classes. In computer vision, segmentation is the task of segmenting images or point clouds to understand what the different segments are [2]. Semantic segmentation is based on dividing an image or a point cloud into semantically meaningful parts where each one is semantically labeled into one of the predefined classes [2] (figure1).

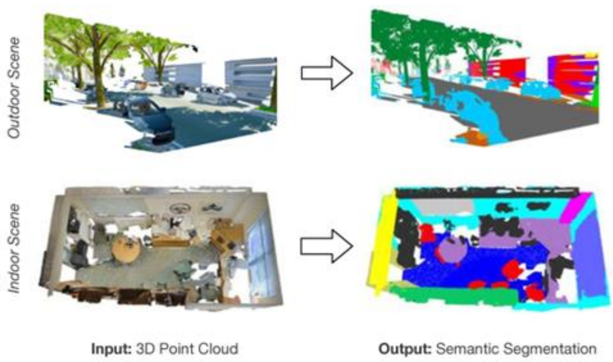

Fig. 1. Illustration of the process of semantic segmentation [3]

Three families of approaches exist to perform a semantic segmentation of Lidar point clouds. The first one is based on the raw point cloud; the second one is based on a product derived from the cloud, mainly Digital Surface Model (DSM), while the third one combines original point clouds and other data sources (aerial image, land map, etc) [4].

With the rapid emerging of deep learning techniques, multiple types of deep learning frameworks have been developed and applied in visual recognition and classification tasks [5]. It has been shown that Convolutional Neural Network (CNN) is capable of recognizing high-level objects by automatically learning a set of abstract features [5]. Among the challenges facing contemporary research in the Lidar field that have encouraged us to conduct our research, is the automation of semantic segmentation of 3D point clouds. The research question we aim to answer is: “How to retrieve the maximum amount of semantic information from Lidar point clouds in an urban environment and how to do it automatically?”

978-1-7281-5806-8/20/$31.00 ©2020 IEEE

The present paper is structured as follows: In section 2, we present some relevant research achievements in the context of segmentation of Lidar data. Section 3 highlights the contribution of Deep Learning methods in semantic segmentation, while section 4 resumes and discusses the principal research findings and limits in the field of semantic segmentation. We finally present the first methodological guidelines of our proposed approach in section 5.

II. AUTOMATIC SEGMENTATION OF LIDAR 3D POINT CLOUDS

Segmentation can be defined as : « a processing which consists in creating a partition of image A in «Ai» regions, in such a way that no region is empty and the intersection between two regions is empty; so all the regions must cover the entire image. A region is a set of related pixels with common properties that differentiate them from neighboring pixels» [6].

Lidar data segmentation can be defined as an iterative grouping of point cloud sets according to the represented object; the grouping continues until there are only regions of homogeneous points remaining and with sufficiently large differences between them [4]. In the case of Lidar data, segmentation approaches can be divided into three categories [4]:

1. Direct approaches that directly address the point clouds.
2. Derived Product Based Approaches.
3. Approaches based on the combination of Lidar data and other sources.

The three approaches are explained in the following paragraphs.

A. Direct approaches

In the literature, a limited number of algorithms has been based on this approach. Among these methods, [7] have proposed a segmentation process based on 3D surface detection, specifically by using Lidar raw data directly without any prior interpolation. This method of segmentation allows automatic division of the point cloud into two classes: ground and buildings, considered as the main objects of an urban scene. The method proposed by [8] is based on the hierarchical decomposition of the cloud to the octree structure until the points contained in each node belong to the same plane. Octree is most often used to partition a three-dimensional space by recursively subdividing it into eight octants [8]. The disadvantage of this method is the disability to classify point clouds [4]. Later, [9] proposed an algorithm for segmentation of airborne Lidar data where the segmentation is performed by a cluster analysis in a feature space, then it's based on several 3D point cloud parameters including density, accuracy, horizontal and vertical point distribution to facilitate the computation of accurate and reliable attributes for the segmentation. This type of segmentation with attributes (step edges, layered surfaces, etc) reveals more information about 3D point clouds. [10] proposed a new 3D point cloud segmentation technique based on clustering attributes, which can be utilized for Lidar data acquired by a terrestrial or airborne mapping system. In this study, the neighborhood of each point is firstly established using an adaptive cylinder, taking into account the density of the acquired point clouds and surface trend, then the Octree structure is used for the detection and the extraction of clusters. The results show the effectiveness of this method for segmentation of point clouds acquired by different mapping systems (terrestrial, airborne, etc.) Afterwards, a clustering-based segmentation technique has been proposed by [11] in which the laser points are aggregated into homogenous regions based on the attributes defined by an adaptive neighborhood of individual points. Then, the segmentation approach is introduced and implemented through three main procedures: neighborhood definition, attribute computation, and clustering the points with similar attributes. Based on the obtained results, an approach was presented for terrain and off-terrain classification of laser scanning data. The experimental results prove that the introduced technique can provide a reliable solution for the classification of laser datasets from multiple sources (airborne and terrestrial).

To sum up, we can say that direct segmentation approaches have a major advantage which lies on direct consumption of raw point clouds without neither resampling operations nor referring to other complementary data sources; this preserves the original characteristics of the acquired data including the precision of the 3D point cloud, the topological relations …etc. However, the major drawback of these methods is the high requirement on memory consumption and calculation time [4].

B. Derived product based approaches

The majority of segmentation methods found in the literature are from this family. Among these methods, [12], [13], [14] have developed some algorithms for image segmentation, allowing for fine-writing of image contours and isolating each part (sky, garden, people... etc.) But the major problem is that these algorithms require a transformation of three-dimensional data into two-dimensional data (resampling), which leads to a loss of information as well as false data caused by the resampling step (one can find a pixel whose spatial information is false caused by extrapolation since the area in question does not contain raised points). Another segmentation method was proposed in [15]. It is based on the segmentation of a 2D image sequence into objects using information about the motion of the robot, which moves objects and observes their visual feedback. [16] proposed a method to improve the segmentation performance for urban scenes which was validated using the publicly accessible dataset (KITTI) with manually annotated images. Later, [17] proposed a new framework for segmentation and classification of 3D point clouds based on the combination of spatial, temporal and semantic information. The addition of semantic indices for the segmentation process would improve the overall performance of the object perception system.

To sum up, we can state that derivatives-based approaches obtained after a resampling operation, have a major advantage which lies on the ease of handling and the speed of 2D data processing, but the major drawback is loss of information due to the transformation of three-dimensional data into two-dimensional data (ie., (resampling shifts the point at the center of the pixel), which can negatively influence the segmentation results.

C. *Approaches based on the combination of Lidar data and other sources*

In the literature, a limited number of algorithms is based on this approach. Among these methods, [18] proposed a new technique that allows the automatic detection of buildings based on the combination of Lidar data and other sources (multispectral imagery). The results obtained by these methods show that the proposed technique gives a very high success rate either in the case of residential buildings or in the case of industrial ones. However, this technique has some limits especially in the case of areas with a very steep slope. [19] proposed an algorithm that combines multispectral imaging and Lidar data acquired by an airborne mapping system to write building boundaries in an urban environment; in the first phase, Lidar data is divided into two parts: soil and objects on the ground; the second phase begins with the process of "Connected Component Analysis" (CCA) where the number of objects present in the test scene are identified by initial boundary detection and labelling [19], then by point clouds. NDVI (Normalized Difference Vegetation Index) data and entropy image are used to extract vegetation classes, buildings, isolated building's occluded parts and finally building boundaries. The obtained results show that the proposed technique gives very satisfactory results that can go up to 100% (overall accuracy) for buildings of more than 50m2. Recently, [1] proposed a new approach based on the combination of Lidar and image data for semantic segmentation of Lidar point clouds; the results show a precision of 88%. Similarly, [20] proposed a semantic segmentation technique based on the combination of 3D point clouds and high resolution images to efficiently merge image data (2D) and Lidar data (3D) ; 3D point clouds are transformed into raster data (DSM and NDSM) to evaluate the effectiveness of the proposed method. The results are validated with the ISPPS dataset of Postdam and a region of Guangzhou in China that are already labelled by an existing database with a good performance in semantic segmentation. Another study of [21] focused on semantic segmentation of complex 3D scenes by merging 2D images and 3D point clouds which objective is to improve the rate of semantic segmentation and to reduce the computational load and memory consumption. To ensure the validity of the DVLSHR model, four benchmark datasets (PASCAL VOC12 [22], SIFT-Flow, CamVid and CityScapes) were utilized in the training and validation stages [21]. The last reviewed work [23] focus on the combination of a high spatial resolution image and airborne Lidar data for building segmentation. The results show that the proposed methodology provides an overall accuracy of 92%.

To conclude, we can say that approaches based on the combination of Lidar data and other planimetric data sources (satellite images, aerial photographs, etc) have more accurate results thanks to the altimeter accuracy of the 3D point clouds and the planimetric continuity of the images [18], but the hugest requirement in this approach is the need to have a minimal time gap between the acquisition of 3D point clouds and that of image data.

D. *Summary*

3D point cloud segmentation approaches can be grouped into direct approaches, post-processing approaches, and approaches based on the combination of data and other sources. The post-processing approaches are still the most used in practice thanks to the ease of handling and the speed of 2D data processing, but their major disadvantage lies on the loss of information caused by the operation of resampling and a lack of algorithms based on these approaches. Direct approaches retain the original characteristics of data but consume a lot of memory and calculation time. Approaches combining 3D points clouds and other sources (satellite images, aerial photos...etc) involve several types of data, which improve the reliability of results, but their major drawback is the need for additional information in real time requiring much more financial resources and material requirements.

Finally, several researches and developments have addressed the issue about segmentation of 3D Lidar data by adopting approaches presented in section2, however, investigations are still needed to improve the quality of the results and the efficiency of the methods (in terms of time of calculation, data required, memory consumption, etc).

The development of deep learning algorithms like Convolutional Neural Networks revolutionizes the field of computer vision and demonstrates good performance of results in many reference datasets for image classification [24]. Their potential in segmentation process would enhance the quality and the efficiency of the results. The next section tries to give a brief overview of researches addressing deep learning in semantic segmentation.

III. CONTRIBUTION OF DEEP LEARNING TO SEMANTIC SEGMENTATION

Semantic segmentation of 3D point clouds acquired in urban environment is the highest-skill task that allows extracting the maximum amount of semantic information present in a urban scene. Today, semantic segmentation is one of the main challenges in computer vision. The revolution of deep learning resulted in resolving many problems in this field processed by deep learning methods and often-by convolutional neural networks [25]. In this section, we examine some deep learning methods that are often used in semantic segmentation of 3D point clouds acquired in urban environments.

The first method called ECC (Labelling in Related Components) has been proposed for image segmentation (multicomponent, color, multispectral) by hierarchical analysis of 3D histograms [26]; the results show that the proposed technique with significant calculation times still need developments to reduce the complexity of this particular approach when the image entropy is high. Similarly, [27] extracted an architecture called VGG-16 composed of 16 layers, for spot localization of objects with an accuracy of 92.7%. Afterwards, a new approach to semantic labeling of

3D point clouds called SnapNet was proposed in [28], this approach applies the CNN to two 2D image views obtained after the resampling of the Lidar data: Red-Green-Blue (RGB) view and another one containing geometric characteristics. These pairs are used as inputs to neural network to give a semantic label to each corresponding pair of pixels from the two input images. The method was evaluated against the semantic data set and achieved better performance. A new semantic segmentation approach called “SegCloud” has been developed in [29]. This approach combines 3D-FCNN (3D Fully Convolutional Neural Networks), Trilinear Interpolation (TI) and CRF (Conditional Random Field). 3D data are voxelized and the resulting 3D grid are processed by the fully convolutional neural network (FCNN)[29]. The results show some noise which has been eliminated by a cleaner segmentation of cloud points with CRF, to improve 3D-FCNN-TI output and qualitative results by recovering object limits in a clear way. The authors evaluated also the approach using two 3D external data sets (KITTI, Semantic3d.net). The results show performance comparable or superior to the state of the art on all datasets. Moreover, an effective architecture called “SegNet”, which uses a fully convolutional neural network for the understanding of road scenes, has been proposed in [30]. The technical particularity of this approach is having no connection part that allows reducing considerably the number of parameters of the network and allowing an easier and faster drive. Furthermore, “SegNet” performs competitively and achieve high scores in understanding road scenes. [31] Developed a new large-scale urban semantic labelling framework by integrating multiple aggregation levels (point-segment-object) and contextual features for road recognition facilities for large-scale highway scene point clouds. This work achieved very satisfactory results with an object recognition accuracy of more than 90%. A new deep learning technique called “PointNet” has been proposed in [32] that directly consumes point clouds without any resampling operations, this preserves the original characteristics of the acquired data including: the precision of the 3D point clouds, the topological relations …etc, in order to improve the rate of semantic segmentation of point clouds through an overall accuracy of 78.62%.

Finally, another work about semantic segmentation approach based on deep learning called Superpoint Graph (SPG) has been proposed by [33]. The later approach distinguishes semantically different objects (grass, tree, building, cars, artifacts, landscaping etc.). This method defines a new state of the art for semantic segmentation of Lidar scans for both outdoor and indoor environments by achieving an overall accuracy of 94%.

Based on the previous analysis, we can raise some shortcomings and gaps that hinder the automatic extraction of maximum semantic information from 3D point clouds acquired in a large-scale urban environment with better precision, good segmentation quality, acceptable computation time, etc. (table1).

TABLE I. ADVANTAGES AND DISADVANTAGES OF THE MAIN DEEP LEARNING SEMANTIC SEGMENTATION METHODS FROM LIDAR DATA

Method	Advantages	disadvantages
ECC	-Structured .	-Expensive . -More adapted for 2D data (color images, multispectral images).
VGG-16	-Few learning parameters.	-More non-linearity functions (since more layers of convolution).
SnapNet	-Efficient	-Requires a transformation of 3D to 2D. - Information losses.
SegCloud	-Effective [34]	-A voxel-based approach which tends to discard small details of 3D shape.
SegNet	-Few parameters -Easy and fast drive.	-Low prediction for some objects (barriers, poles, pedestrians and cyclists [35]).
PointNet	-Accurate [34].	- Expensive [34] . -Uncapability for processing large 3D points at once [33].
SPGraph	- Applicable for large point clouds.	-Low prediction for some objects (low vegetation, hard-scape [33]).

IV. DISCUSSION

Semantic segmentation is an effective operation for a deep and rigorous understanding of urban environments. Tools and methods used for semantic segmentation of Lidar points cloud are numerous and very successful as they facilitate the extraction of semantic information for urban scenes understanding. Some of them are particularly flexible, allowing their adaptation to user needs. These tools are therefore interesting in several respects (robust, faithful), but have some limitations. Indeed, some approaches provide interesting semantic information but have low accuracy (the recognition rate of objects is sometimes low).While other approaches give a good precision but with very limited semantic classes which can cause a miss understanding of an urban scene. The PointNet approach is accurate but uncapable for processing large 3D points at once. Some work such as in [31] resulted in very satisfactory results with very high object recognition accuracy, but the experiment was conducted on a point cloud that presents fewer constraints (example highway point clouds). Some algorithms, such as the case of ECC, with significant computing times, show that efforts are still needed to reduce the complexity of this approach.

Based on the previous discussion, we can say that the SPGraph method is the preferred choice for semantic segmentation of point clouds captured in large-scale urban environments (applicable for large point clouds), because its principle is based on the splitting the point cloud into simple shapes, easier to classify than points. This choice was also confirmed by a study of [36]; which concluded that the SPGraph method is the preferred one over the PointNet and PointCNN techniques.

Based on our investigation of the state of the art of semantic segmentation and deep learning methods from recent literature, we propose to address some challenges about efficiency of the process in terms of semantic precision and automation, for a better understanding of urban scenes acquired by airborne Lidar. We so propose to design and test a methodology that allows both automatic extraction and better precision of semantic information present in a large-scale urban environment.

V. OUR PROPOSED APPROACH

The analysis of the previous works show that the semantic segmentation of Lidar point clouds, automatically and especially in urban areas, is one of the problems that remain even more difficult in computer vision creating ambiguity in semantic identification of objects. Hence, developing robust, faithful and flexible methods to recognize objects and associate them with semantics in order to make the best use of three-dimensional data of the mapping systems remains a challenging task.

For a better semantic segmentation of Lidar point clouds acquired in an urban environment, we propose a method of semantic segmentation based on deep learning and integrating several data sources (satellite images, aerial photos) in order to enhance the accuracy of the segmentation process.

A. Materials and methods

To test the semantic segmentation algorithm that will be developed as part of this research, we propose to use Lidar data collected in two areas. Each area has multiple sites with different locations and resolutions.

Zone 1: this zone covers three sites of the city of Khemisset (Morocco), where the Lidar data was acquired by the National Agency for the Land Conservation of Cadastre and Cartography.

Zone 2: this zone covers five sites in the region of San Francisco, California, USA, where the data were downloaded free of charge from the site: http://www.opentopography.org . These data were chosen because of the low point density in zone 1; the density of points in these five sites is greater than 4 points/ m2.

We also planify to acquire Lidar data for a third zone with high resolution and diverse urban objects.

B. Methodology

Our research sets as an objective: the proposal of a methodology based on the combination of Lidar data and other sources in conjunction with a deep learning method for semantic segmentation of large-scale urban Lidar point clouds. Such a methodology will ensure, among other things, the following functionalities:

- Integration of data from different sources for a given zone;
- Utilization of raw point cloud directly;
- Development of a deep learning method (neural networks);
- Consideration of geometric and radiometric information;
- Assistance in the deep understanding of large-scale urban scenes;
- Extraction of maximum semantic information in a large-scale urban environment;
- A high precision rate compared to the existing results in the literature;

To improve the semantic segmentation rate compared to the existing methods, we present a general workflow that summarizes the different steps to extract the maximum amount of semantic information (figure 2).

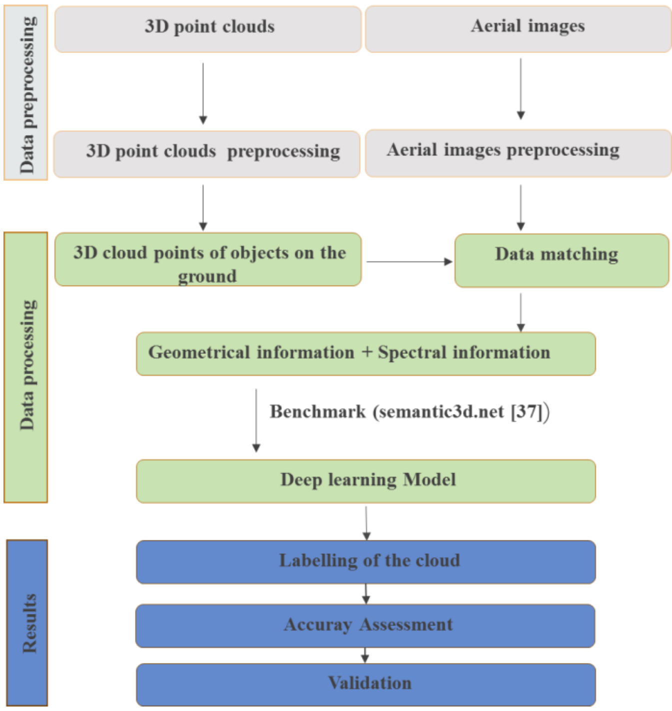

Fig. 2. Our proposed approach

Our approach is proposed after a deep bibliographical study which allowed us to well understand the limitations of the semantic segmentation methods mainly used in the literature. The second reason to choose this approach is based on the crucial needs of scientific research in this field. We could conclude which points are interesting and we decided to focus on them in order to obtain the maximum of semantic information available in an urban area with better precision. For example, our approach will take into consideration geometrical and radiometric aspects, it will not require any sampling method and it will be applied for large point clouds. Our methodology is expected to give better results in terms of precision and robustness to recognize 3D objects and associate them a rich semantic.

CONCLUSION:

Several deep learning architectures for semantic segmentation of 3D point clouds acquired in urban environment have been developed in the literature.

In this paper, we analyzed and discussed several scientific contributions in the field of semantic segmentation of large-scale point-clouds in urban areas , we can conclude that the extraction of all the semantic information necessary for a better understanding of urban scenes with better precision remains even more difficult in computer vision. This is the reason why we have proposed an approach involving a deep learning method, which integrates several data sources that will allow to extract automatically the maximum amount of semantic information in a urban environment and get around the limitations of the existing methods.

REFERENCES

[1] Haoyi. Xiu, Poliyapram. Vinayaraj, Kyoung-Sook. Kim,Ryosuke. Nakamura , Wanglin. Yan , "3D Semantic Segmentation for High-resolution Aerial Survey Derived Point Clouds using Deep Learning (Demonstration)," Information Systems (SIGSPATIAL '18), November 6–9, 2018, Seattle,WA, USA, Farnoush Banaei-Kashani and Erik Hoel (Eds.). ACM, New York, NY, USA,2018.

[2] Patrik. Tosteberg ," Semantic Segmentation of Point clouds using Deep Learning," Master of science Thesis in Electrical Engineering , Department of Electrical Engineering , Linkoping University , 2017.

[3] Francis. Engelmann,Theodora. Kontogianni,Jonas. Schult,Bastian. Leibe, "Know What Your Neighbors Do: 3D Semantic Segmentation of Point Clouds," RWTH Aachen University, Aachen, Germany,2018.

[4] Bellakaout. Abdelmounaim,"Extraction automatique des batiments, végétation et voirie à partir des données Lidar 3D ," Thèse de docteur de l'institut agronomique et vétérinaire Hassan II , Maroc , 2016.

[5] Zixiang. Zhou, Jie. Gong , "Automated residential building detection from airborne Lidar data with deep neural networks," Advanced Engineering Informatics 36 229–241 , 2018 .

[6] Capri. A ," caractérisation des objets dans une image en vue d'une aide à l'interprétation et d'une compression adaptée au contenu : application aux images échographiques ,"2007.

[7] Lee. I,Schenk. T, "Perceptual organization of 3D surface points," Photogrammetric computer vision. ISPRS comm . III , Graz , Austria .Vol .XXXIV,part 3A/B , ISSN 1682-1750,2002.

[8] Wang. M ,Tseng.Y. H,"Lidar data segmentation and classification based on octree structure," Int . International Archives of Photogrammetry and Remote Sensing , ISSN 1682-1750, Vol , XXXV,partB3,2004.

[9] Filin. S,Pleifer. N, " Segmentation of airborne laser scanning data using a slope adaptive neighbourhood ," ISPRS Journal of Photogrammetry and Remote Sensing 60 (2006) 71-80,2006.

[10] Lari. Z,A. Habib,E. Kwak, "An adaptive approach for segmentation of 3D laser point cloud ,"ISPRS Workshop Laser Scanning , Calgary,Canada ,2011.

[11] Lari. Z,A. Habib, "Segmentation–based classification of laser scanning data,"ASPRS 2012 Annual Conference , Sacramento , California ,2012.

[12] Heath. M,S. Sarkar,T. Sanocki,K. Bowyery , " Comparison of edge detectors a methodology and initial study, "Computer Vision and Image Understanding ,69,38-54,1998.

[13] Jiang. X,H. Bunke, "Edge detection in range images based on scan line approximation ," Computer Vision and Image Understanding , 73,183-199,1999.

[14] Sappa. A. D,M. Devy , "Fast range image segmentation by an edge detection strategy ," Proceedings of 3nd International Conference on 3-D Digital Imaging and Modeling,292-299,2001.

[15] Beale. D, Iravani. P , Hall. P , "Probabilistic Models for Robot-Based Object Segmentation," Robotics and Autonomous Systems, vol. 59, issue 12, pp. 1080–1089,2011.

[16] Jeonghyeon. Wang , Jinwhan. Kim, " Semantic Segmentation of Urban Scenes with a Location Prior Map Using Lidar Measurements ," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) September 24–28, 2017, Vancouver, BC, Canada , 2017 .

[17] Mehmet. Ali. Cagrı. Tuncer ,Dirk. Schulz , "A Framework for the Segmentation and Classification of 3D Point Clouds using Temporal, Spatial and Semantic Information," Cognitive Mobile Systems, Fraunhofer FKIE, Fraunhoferstr. 20, 53343 Wachtberg, Germany .

[18] Mohammad. Awrangjeb, Mehdi. Ravanbakhsh, Clive S. Fraser, "Automatic detection of residential buildings using LIDAR data and multispectral imagery," ISPRS Journal of Photogrammetry and Remote Sensing 65 (2010) 457-467, 2010.

[19] S. A. N. Gilani , M. Awrangjebb , G. Lu , " FUSION OF LIDAR DATA AND MULTISPECTRAL IMAGERY FOR EFFECTIVE BUILDING DETECTION BASED ON GRAPH AND CONNECTED COMPONENT ANALYSIS ," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-3/W2, 2015 PIA15+HRIGI15 – Joint ISPRS conference 2015, 25–27 March 2015, Munich, Germany , 2015.

[20] Ying. Sun, Xinchang. Zhang, Qinchuan. Xin, Jianfeng. Huang, "Developing a multi-filter convolutional neural network for semantic segmentation using high-resolution aerial imagery and Lidar data ," ISPRS Journal of Photogrammetry and Remote Sensing , 2018 .

[21] Rui. Zhanga ,Guangyun. Lia, Minglei. Lia,Li. Wanga," Fusion of images and point clouds for the semantic segmentation of largescale 3D scenes based on deep learning , "ISPRS Journal of Photogrammetry and Remote Sensing ,2018.

[22] Everingham. M,Van.Gool. L, Williams.C.K. I,Winn. J,Zisserman. A, Int. J. Comp. Vis. 111 (1), 98–136,2015.

[23] David. Griffiths , Jan. Boehm , " Improving public data for building segmentation from Convolutional Neural Networks (CNNs) for fused airborne Lidar and image data using active contours , "ISPRS Journal of Photogrammetry and Remote Sensing 154 (2019) 70–83 , 2019.

[24] Li. Y, "Deep learning for remote sensing image classification: A survey.Wiley Interdisciplinary Reviews," Data Mining and Knowledge Discovery. Vol. 8, pp. 1264,2018.

[25] Garcia-Garcia.A, "A Review on Deep Learning Techniques Applied to Semantic Segmentation," ArXiv press. 23 p,2017.

[26] S. OUATTARA,A. CLÉMENT, "Étiquetage d'histogrammes multidimensionnels compacts pour l'analyse d'images multicomposantes , " Colloque GRETSI, Troyes , 11-14 septembre 2007.

[27] Zisserman.A, Simonyan.K, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv print. 14p, 2014.

[28] A. Boulch, J. Guerry, B. L. Saux, and N. Audebert, "Snapnet: 3d point cloud semantic labeling with 2d deep segmentation networks," Computers Graphics, vol. 71, pp. 189–198, 2018.

[29] L. P. Tchapmi, C. B. Choy, I. Armeni, J. Gwak, and S. Savarese, "Segcloud: Semantic segmentation of 3d point clouds," 2017 International Conference on 3D Vision (3DV), pp. 537–547, 2017.

[30] Vijay.B, Kendall. A, Cipolla.R, "SegNet: A Deep Convolutional Encoder- Decoder Architecture for Image Segmentation," IEEE transactions on pattern analysis and machine intelligence. Vol. 39, pp. 2481-2495, 2017.

[31] Bisheng.Yang , Zhen. Dong, Yuan.Liu, Fuxun.Liang , Yongjun.Wang , " Computing multiple aggregation levels and contextual features for road facilities recognition using mobile laser scanning data," ISPRS Journal of Photogrammetry and Remote Sensing, Volume 126, pp. 180-194,2017.

[32] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," CoRR, vol. abs/1612.00593, 2016.

[33] Loic.Landrieu, Martin.Simonovsky, "Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs," The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4558-4567,2018.

[34] Loic. Landrieu,"Segmentation sémantique de grands nuages de points avec Superpoint Graph," MATIS – IGN,2018.

[35] Nicolas. Audebert, Alexandre. Boulch, Bertrand. Le Saux, Sébastien. Lefèvre,"Segmentation sémantique profonde par régression sur cartes de distances signées," Reconnaissance des Formes, Image, Apprentissage et Perception (RFIAP),Marne-la-Vallée, France. hal-01809991,Jun 2018.

[36] Chakri.Lowphansirikul, Kyoung-Sook .Kim, Poliyapram .Vinayaraj, and Suppawong.Tuarob ,"3D Semantic Segmentation of Large-Scale Point-Clouds in Urban Areas Using Deep Learning," International Conference on Knowledge and Smart Technology (KST),2019 . DOI: 10.1109/KST.2019.8687813

[37] T. Hackel, N. Savinov, L. Ladicky, J. D. Wegner, K. Schindler, and M. Pollefeys, "SEMANTIC3D.NET: A new large-scale point cloud classification benchmark," in ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. IV-1-W1,pp. 91–98,2017.