Archives of Photogrammetry, Cartography and Remote Sensing, vol. 33, pp. 13-22
ISSN 2083-2214, eISSN 2391-9477
DOI: http://doi.org/10.2478/apcrs-2021-0001

RESEARCH ARTICLE

DEEP LEARNING FOR AUTOMATIC LIDAR POINT CLOUD PROCESSING

Wojciech Dominik, Marcin Bożyczko, Karolina Tułacz-Maziarz

OPEGIEKA Sp. Z o. o., Elbląg, Poland

KEY WORDS: deep learning, LiDAR, point cloud, classification, automation

ABSTRACT: The paper presents the method of automatic point cloud classification that has been developed by OPEGIEKA. The method is based on deep learning techniques and consists of an in-house developed algorithm of point cloud transformation to a regular array accompanied by internally designed convolutional neural network architecture. The developed workflow as well as experiences from its application during the execution of the CAPAP project are described. Results obtained on real project data as well as statistics obtained on the ISPRS 3D semantic labelling benchmark with the use of OPEGIEKA's method are presented. The achieved results place OPEGIEKA in the top 3 of the classification accuracy rating in the ISPRS benchmark. The implementation of OPEGIEKA's solution into LiDAR point clouds classification workflow allowed to reduce the amount of necessary manual work.

1. INTRODUCTION

Classification of LiDAR point clouds is a very challenging task mainly because of the huge amount of data. Since the appearance of airborne laser scanning technology efforts have been made to automate this task to the maximum extent possible. This goal was achieved mainly for ground surface and buildings.

The most commonly used algorithms for automatic extraction of ground points are based on searching for lowest points in the point cloud and iteratively adding points to the ground class (Axelsson 2000, Kraus & Pfeifer 2001). Building classification is usually based on roof planarity analysis and plane fitting (Maas & Vosselman 1999, Rottensteiner & Briese 2004). For the last two decades, these algorithms were used as a standard in the LiDAR point cloud classification workflow. However automatic classification of objects with a more complex geometry is still a challenge. Still, many objects have to be classified manually.

Since the appearance of deep learning techniques trials have been made to use these techniques for LiDAR point cloud classification. However, the main limitation of applying machine learning techniques for LiDAR data classification is the irregular structure of the point cloud. Any type of artificial neural network requires the input data to be a vector of a fixed length containing values in a defined order or an array representing the spatial relations of the data. However, a point cloud is an unordered set of coordinates. The density and spatial distribution of points can vary a lot. Therefore, the data cannot be fed directly into the neural network. Some form of transformation of the point cloud has to be applied.

The most common types of neural networks used for LiDAR data classification are convolutional neural network (CNN). CNNs are applicable for all data that have spatial structure and can be represented by a regular grid. Therefore, these types of neural networks are perfectly suitable for geospatial data because they learn spatial patterns at different scales.

The most straightforward solution to apply a convolutional neural network to point cloud classification is to transform the unstructured 3D point sets into a regular 3D array of voxels (Huang & You 2016, Zhou & Tuzel 2018, Tchapmi et al. 2017, Maturana & Scherer 2015). This makes it possible to feed the data to a neural network constructed from 3D convolutional layers. However, such transformation is associated with the generalization of the point cloud due to voxel dimensions. Moreover, it results in an unnecessarily large representation of the point cloud, where the majority of the voxels remain empty. This drawback makes it difficult to apply the voxelization method to large areas covered by airborne laser scanning data.

Another possibility of feeding point clouds to a convolutional neural network is to transform the points into a 2D representation (Yang et al. 2017, Hu & Yuan 2016, Yang et al. 2018, Rizaldy et al. 2018a, Rizaldy et al. 2018b). In this type of methods for each point of the point cloud, an image of its surrounding is generated. This image contains features derived from raw point coordinates such as planarity, sphericity or roughness. The important drawback of these methods is the need of designing features which describe point cloud characteristics. Another limitation is that the data preparation is very time-consuming because an image has to be generated for every point of the point cloud.

Some works apply convolution directly to unstructured point sets (Qi et al. 2017a, Qi et al. 2017b, Yousefhussien et al. 2018, Su et al. 2018, Wen et al. 2020). These methods are mainly dedicated to a single object or indoor scene classification. The neural network is fed by a set of points of a fixed length. Applying a symmetric function makes the model invariant to the point order. Their main drawback in the context of airborne laser scanning data is that a limited number (usually a few thousand) of points can be fed to the network. Therefore, processing large datasets requires splitting the point clouds into sets of a predefined number of points. This makes it time-consuming when processing large areas. Another limitation is that the input sets of points represent only small parts of the point cloud and lack context information.

Since 2016 OPEGIEKA is carrying out research in order to implement deep learning techniques into LiDAR point cloud classification workflow. Some of the most promising solutions found in the literature have been tried. However, none of these were suitable, because of unsatisfying quality or limitations concerning its application to large data volumes. Therefore, OPEGIEKA has developed an in-house solution for LiDAR point cloud classification based on deep learning techniques.

This paper presents the method of automatic point cloud classification that has been used by OPEGIEKA during the execution of several projects since 2019. Its strengths, flaws and future potential have been described.

The method has been developed by OPEGIEKA's engineers as a result of R&D works executed using the company's own aircraft, remote sensing devices as well as in-house developed software and solutions commonly available on the market.

Sharing that experience and knowledge is our way of taking a stand in a current discussion about the future of manual work in processing remote sensing data. At OPEGIEKA we strongly believe in automating the processes of data analysis

The details of the most recent project in which the method was successfully launched are presented below. The data was gathered during the execution of the CAPAP project in Poland.

2. WORKFLOW

OPEGIEKA's solution for point cloud classification is based on a fully convolutional neural network. The core of the solution is the in-house developed algorithm of point cloud transformation to a regular array accompanied by internally designed convolutional neural network architecture. First the space occupied by the point cloud is divided into grid cells of 1 meter by 1 meter on the XY plane. Then images are generated layer by layer starting from the lowest points in each grid cell. The coordinates of the consecutive points in ascending elevation order are written as cell attributes of the consecutive images. Then the set of generated images is stacked together. This results in a 4-dimensional array. The first two dimensions of the array are the spatial width and height of the point cloud divided into a 1-meter grid. The third dimension is the ascending order of elevations of the points in a given grid cell. In the fourth dimension point coordinates are stored. The size of the array can be set arbitrarily but hardware limitations have to be taken into account. An array of 64 by 64 by 64 by 3 is a reasonable choice. If in a certain grid cell, the number of points exceeds the size of the third dimension points are randomly selected. The 4-dimensional array is then fed to a fully convolutional neural network constructed from 3D convolutional layers. During training the data is generated by a generator that randomly selects patches of point clouds and transforms them into the 4-dimensional array.

In order to carry out point cloud classification the neural network has to be trained using correctly classified data samples. The training sample should be classified according to the target class definition for the data to be classified. It is important that the classification of the training sample is consistent and as accurate as possible so that the neural network will avoid learning errors or misinterpreting the results.

For each project, a training sample must be selected and classified with the highest accuracy. The required amount of training data depends on the terrain type. For a LiDAR block representing a homogeneous terrain type a few square kilometers of representative data would be enough. If the terrain type varies, the optimal results are obtained by training dedicated models for different terrain types, such as urban, rural, forest, seacoast, flat, and mountainous.

The process of training the neural network from scratch takes a couple of days. When training models are dedicated for different terrain types, a pre-trained general model can be used which reduces the amount of time required for training further models. Furthermore, a model trained from another project can be used as a pre-trained model as well. However, if the number of classes is different retraining the model will take longer, causing the drop of significance of the amount of time saved in comparison with the identical class definition.

After training the neural network, the classification of the point cloud is carried out by prediction of the neural network. For practical reasons, the LiDAR data is divided into 500 by 500 meters tiles. The full process of classification of such tile takes about 3-4 minutes on a single machine equipped with a graphic card (GPU).

In order to obtain optimal results an automatic cleaning of the classification obtained from the neural network is carried out. That means applying a set of Terrascan macros, enabling deleting some of the classification errors of insignificant importance to the statistic accuracy but decreasing the visual quality of the classification. The core of this stage is applying several macros based on isolated points and point neighborhood analysis.

At the final stage, the point cloud is subjected to manual inspection to correct residual errors and objects that require human interpretation.

3. IMPLEMENTATION

The current automation method was implemented for the first time during the execution of the CAPAP project in 2019. CAPAP, standing for The Centre for Spatial Analysis of Public Administration, is the project that has been implemented in Poland by the Head Office of Geodesy and Cartography (GUGiK) since 2015. OPEGIEKA has been involved in the project implementation since the beginning by gathering and processing both LiDAR data and aerial imagery from different parts of the country.

The entry dataset required for conducting the training has been selected from the vicinity of Cracow and Warsaw. The data sample consisted of nearly 1000 tiles, with a single tile being a square with a side length equal to 500 meters. The values used for prediction in the CAPAP 2019 have been obtained after a few days of training the model. As a matter of fact, already then it has been noticed that using automation in such project could lead to achieving satisfying results. For further optimisation of the process, the approach towards the training and class distribution has been changed. In the CAPAP project, a few types of objects that differ with respect to their shape and spatial position were contained within a single class. In order to simplify the learning process, a new class division has been created, which allowed for more complex segmentation, i.e. cars, staircases, terraces, powerlines, and so on have been allocated separately. Subsequently, a few representatives, urban tiles have been selected and manually prepared as entry data for the training on previously changed classes. It is worth mentioning that the highest possible quality of the entry dataset consisted of manually prepared tiles is essential. Due to that approach, a few tiles were created, which the neural network has been trained on. The abovementioned process of the optimisation has led to expected effects resulting in greater accuracy of the classified objects.

The substantial advantage of implementing the automation processes in the CAPAP is achieving a high level of data coherence, which is especially difficult to accomplish using only manual methods. Apart from this, the increase in the accuracy of classifying points allocated for the objects that the model has already been familiar with was observed. Besides, the acceleration of the data processing was equally important. The best results have been observed within the urban area, where typical objects occur commonly, e.g. detached houses or cars. Insuch cases, the increase in efficiency amounted to several dozens of percent. Ingeneral, the prediction time for a single tile characterised by size 500 per 500 meters and density of 12 pts per square metre lasted approximately 2 minutes. Admittedly, the developed technology has great potential for optimisation. Due to that fact, the company is going to carry out the works regarding further improvements leading to even more accurate classification of irregular objects.

For current needs, the company usually applies in-house developed automated methods of classifying point clouds to the following groups: ground surface, vegetation, buildings and bridges, water surface, noise above and below ground, and all other manmade objects that are not buildings nor bridges (vehicles, fences, powerlines, power poles, street lanterns, bus stops, temporary ground repository, greenhouses, jetties etc.) and other commonly used classes. However, OPEGIEKA's experience shows that automatically performing more detailed classification is feasible. The general rule is that if a type of object is clearly and logically distinguishable from the point cloud based only on point cloud geometry (without any external source of information or general knowledge about the world that is inaccessible to the machine learning algorithm) it can be automatically detected with the use of described methods.

The company has tested and obtained outstanding results by classifying objects such as vehicles, building facades, stone walls on the fields, etc. to separate classes. The solution is flexible and allows for defining any number of classes to be classified. The only requirement is to have training data of good quality where the objects to detect are correctly classified as separate classes.

4. QUALITY ASSESSMENT

The quality of automatic classification of the point cloud is difficult to measure statistically. The percentage of correctly classified points does not always reflect the quantity of work that has to be done in order to achieve fully correct classification. It depends on the distribution of the misclassified points and the number of objects requiring manual correction of the classification. In the results obtained with the use of our methods the majority of objects that usually are subject to human intervention, such as buildings and vehicles, are classified correctly (fig. 1-3). The most common errors are related to rarely occurring or unconventional objects that "confuse" the neural network leading to mixing different classes. However, if this type of object is concerned, avoiding human intervention is nearly unattainable. The number of such objects is insignificant; therefore the need for human intervention is greatly reduced in comparison with conventional methods.

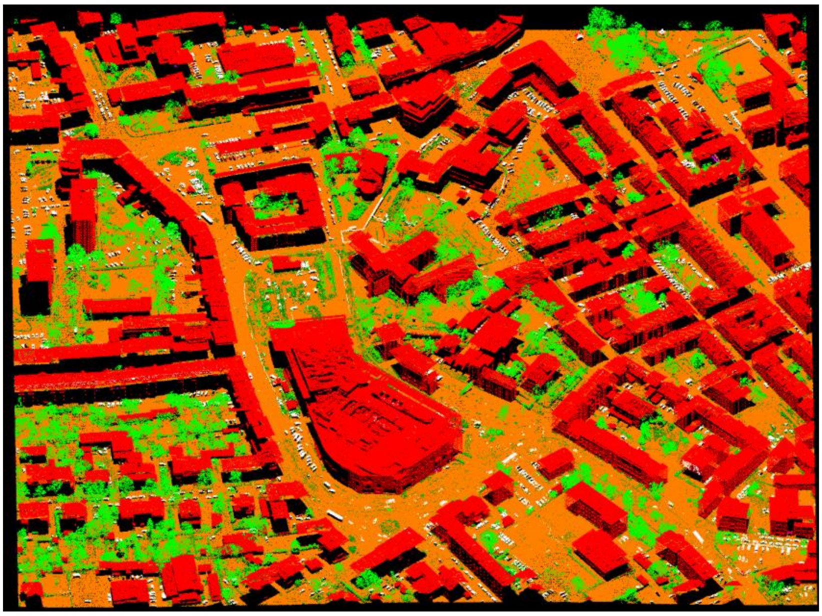

Fig. 1. LiDAR point cloud tile classified with the use of OPEGIEKA's method

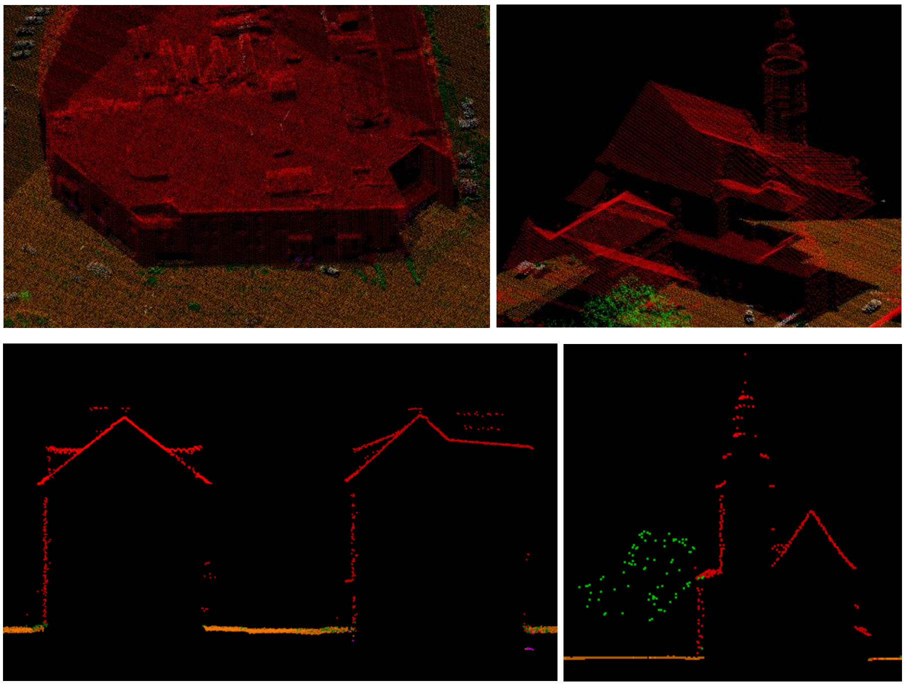

Fig. 2. Buildings classified with the use of OPEGIEKA's method

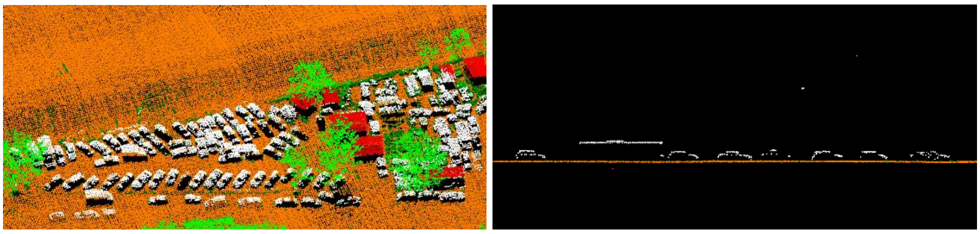

Fig. 3. Cars classified with the use of OPEGIEKA's method

The statistic accuracy depends on many factors and varies greatly from tile to tile depending on the objects that occur on a given tile. In order to present possibly the most objective statistics the company used the results obtained at the ISPRS 3D semantic labelling benchmark (Rottensteiner et al. 2012, ISPRS 2012). Achieved results place OPEGIEKA in the top 3 of the classification accuracy rating. It is worth mentioning that most of the authors use external tools for ground classification and use this information as input to their algorithm. However, OPEGIEKA's method uses only the information obtained directly from the .las files, such as points coordinate and optionally intensity return number and number of returns for a given pulse. Efficiency is a big advantage of our method, as it takes less than 60 seconds to classify the reference dataset of the benchmark.

Table 1. Results obtained on the ISPRS 3D semantic labelling benchmark with the use of OPEGIEKA's method

<table>
<tr><th>Class</th><th>F1 score [%]</th><th>Overall accuracy [%]</th></tr>
<tr><td>Powerline</td><td>50.4</td><td rowspan="9">82.6 %</td></tr>
<tr><td>Low Vegetation</td><td>81.3</td></tr>
<tr><td>Impervious Surfaces</td><td>91.1</td></tr>
<tr><td>Car</td><td>77.0</td></tr>
<tr><td>Fence/Hedge</td><td>27.9</td></tr>
<tr><td>Roof</td><td>93.2</td></tr>
<tr><td>Facade</td><td>56.0</td></tr>
<tr><td>Shrub</td><td>41.2</td></tr>
<tr><td>Tree</td><td>80.1</td></tr>
</table>

5. CONCLUSION

The proposed solution allows for minimizing the amount of manual work in comparison with conventional methods applied before that are based only on the Terrascan macros. The advantages of OPEGIEKA's method are the most clearly visible objects difficult or impossible to classify using state-of-the-art automatic methods, such as cars, building walls and details, powerlines etc.

The method requires having correctly classified sample data of the size of at least a few square kilometers. The sample is used as training data for the neural network. Consequently, the quality of the classification of the training data is crucial for the accuracy of the latter automatic classification. Moreover, the training data has to be representative for the data to be classified. It has to contain all types of objects that occur in the project area in a representative quantity. Therefore, the training samples have to be carefully selected from the whole dataset. This is an additional difficulty of the whole classification process compared to the conventional workflow.

To obtain optimal results the training data has to come from the same project. This requires a reorganization of the classification workflow. Firstly, some samples of representative data have to be classified correctly. Then, the training process, which takes a few days, is carried out. The necessity of spending that additional time is an important drawback of the workflow.

The company's experience shows that applying the method based on neural network enables achieving a significant reduction of time required for manual inspection, as well as increasing the final quality of the classification. Those are the main reasons making OPEGIEKA's engineers believe this is the best methodology at the moment.

LITERATURE

Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. *International archives of photogrammetry and remote sensing*, *33*(4), 110-117.

Hu, X., & Yuan, Y. (2016). Deep-learning-based classification for DTM extraction from ALS point cloud. *Remote sensing*, *8*(9), 730.

Huang, J., & You, S. (2016). Point cloud labeling using 3D convolutional neural network. In *2016 23rd International Conference on Pattern Recognition (ICPR))* (pp. 2670-2675). IEEE.

ISPRS. 2012: https://www2.isprs.org/commissions/comm2/wg4/benchmark/3d-semantic-labeling (access: 12/2020)

Kraus, K., & Pfeifer, N. (2001). Advanced DTM generation from LIDAR data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, *34*(3/W4), 23-30.

Maturana, D., & Scherer, S. (2015). Voxnet: A 3D convolutional neural network for real-time object recognition. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 922-928)*. IEEE.

Maas, H. G., & Vosselman, G. (1999). Two algorithms for extracting building models from raw laser altimetry data. *ISPRS Journal of photogrammetry and remote sensing*, *54*(2-3), 153-163.

Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017a). Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 652-660).

Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017b). Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*.

Rizaldy, A., Persello, C., Gevaert, C. M., & Oude Elberink, S. J. (2018a). Fully convolutional networks for ground classification from LiDAR point clouds. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 4(2).

Rizaldy, A., Persello, C., Gevaert, C., Oude Elberink, S., & Vosselman, G. (2018b). Ground and multi-class classification of airborne laser scanner point clouds using fully convolutional networks. *Remote sensing*, 10(11), 1723.

Rottensteiner, F., & Briese, C. (2003). Automatic generation of building models from LIDAR data and the integration of aerial images.

Rottensteiner, F., Sohn, G., Jung, J., Gerke, M., Baillard, C., Benitez, S., & Breitkopf, U. (2012). The ISPRS benchmark on urban object classification and 3D building reconstruction. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* I-3, Nr. 1, 1(1), 293-298.

Su, H., Jampani, V., Sun, D., Maji, S., Kalogerakis, E., Yang, M. H., & Kautz, J. (2018). Splatnet: Sparse lattice networks for point cloud processing. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2530-2539).

Tchapmi, L., Choy, C., Armeni, I., Gwak, J., & Savarese, S. (2017, October). Segcloud: Semantic segmentation of 3d point clouds. In *2017 international conference on 3D vision (3DV) (pp. 537-547).* IEEE

Wen, C., Yang, L., Li, X., Peng, L., & Chi, T. (2020). Directionally constrained fully convolutional neural network for airborne LiDAR point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 50-62.

Yang, Z., Jiang, W., Xu, B., Zhu, Q., Jiang, S., & Huang, W. (2017). A convolutional neural network-based 3D semantic labeling method for ALS point clouds. *Remote Sensing*, *9*(9), 936.

Yang, Z., Tan, B., Pei, H., & Jiang, W. (2018). Segmentation and multi-scale convolutional neural network-based classification of airborne laser scanner data. *Sensors* 18(10), 3347.

Yousefhussien, M., Kelbe, D. J., Ientilucci, E. J., & Salvaggio, C. (2018). A multi-scale fully convolutional network for semantic labeling of 3D point clouds. *ISPRS journal of photogrammetry and remote sensing*, *143*, 191-204.

Zhou, Y., & Tuzel, O. (2018). Voxelnet: End-to-end learning for point cloud based 3d object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4490-4499).

Details of authors:

PhD, Eng., Wojciech Dominik
e-mail: wojciech.dominik@opegieka.pl
tel: +48 (55) 237 60 00

M.Sc., Marcin Bożyczko
e-mail: marcin.bozyczko@opegieka.pl
tel: +48 (55) 237 60 00

M.Sc., Eng., Karolina Tułacz-Maziarz
e-mail: karolina.tulacz@opegieka.pl
tel: +48 (55) 237 60 00

GŁĘBOKIE UCZENIE W AUTOMATYCZNYM PRZETWARZANIU CHMURY PUNKTÓW SKANOWANIA LASEROWEGO

SŁOWA KLUCZOWE: głębokie uczenie, LiDAR, chmura punktów, klasyfikacja, automatyzacja

Streszczenie

W artykule przedstawiono metodę automatycznej klasyfikacji chmur punktów opracowaną przez firmę OPEGIEKA. Metoda opiera się na technice głębokiego uczenia i składa się z opracowanego przez autorów algorytmu transformacji chmury punktów do regularnej macierzy, któremu towarzyszy wewnętrznie zaprojektowana architektura konwolucyjnej sieci neuronowej. W tekście opisano opracowany ciąg technologiczny uwzględniający metodykę na przykładzie doświadczenia podczas realizacji projektu CAPAP. Przedstawiono wyniki uzyskane na rzeczywistych danych projektowych oraz statystyki uzyskane na benchmarku ISPRS dotyczącego etykietowania semantycznego z wykorzystaniem metody OPEGIEKA. Osiągnięte wyniki plasują OPEGIEKA w pierwszej 3 rankingu dokładności klasyfikacji w benchmarku ISPRS. Wdrożenie rozwiązania OPEGIEKA do przepływu pracy klasyfikacji chmur punktów LiDAR pozwoliło zmniejszyć ilość niezbędnej pracy manualnej.

Submitted 30.07.2021
Accepted 31.12.2021