QUALITY IMPROVEMENT FOR AIRBORNE LIDAR DATA FILTERING BASED ON DEEP LEARNING METHOD

T. Yotsumata[*], M. Sakamoto, T. Satoh

PASCO CORPORATION, 4-9-6 Aobadai Meguro-ku, 153-0042 Tokyo, Japan - (taotha7751, moittu9191, tuoost7017)@pasco.co.jp

Commission II, WG II/3

KEY WORDS: Airborne LiDAR, Filtering, Deep Learning, Point Cloud, Voxelization, CNN

ABSTRACT:

In this paper, we discuss how to improve the quality of classification results when deep learning is applied for the filtering of airborne LiDAR point cloud. We introduce the baseline method which utilizes convolutional neural network (CNN) based on voxelization, and then we propose three methods to improve the quality of classification result. The first method is data pre-processing that aims to exclude data in advance that is clearly not on the ground surface in order to efficiently extract the ground surface data. Data pre-processing can greatly reduce the number of target points and the subsequent processing can be performed efficiently. It also has the effect of preventing noise-like points floating in the air from being misclassified as the ground surface, as compared to the case without pre-processing. The second method is changing the network structure. In recent years, various networks have been proposed for classifying point clouds. In our study, the baseline is using very simple networks. In order to improve the classification result of the baseline method, the layer depth and the range size of convolution are changed, and we investigated about the improvements of the results. The current discussion can be used as a guidance when considering new networks. The third method is the integration of classification results from multiple networks. We integrated individual results from multiple networks with varying layer depths and convolution sizes, starting with the baseline, and investigated whether the results improved. We observed that even if the individual results were similar, the classification results can be improved by integrating the results.

1. INTRODUCTION

Airborne LiDAR surveying is a technology that uses a laser scanner and a GNSS/IMU unit mounted on an aircraft to acquire elevations and shapes of a wide area of ground surface as point cloud data. A laser scanner emits laser beam pointing towards the ground, and measures the distance between the scanner and ground objects from the time until the laser reflects off the ground and returns. The GNSS/IMU unit measures the position and orientation of the laser scanner.

This technology has been widely used in many fields, such as topographic mapping, forestry, flood control, and has become the most common method for acquiring extensive terrain data over the past few decades. Since the laser emitted from the laser scanner toward the ground reflects off everything on the ground, the obtained point cloud data (DSMs: Digital Surface Models) include anything like ground surface, vegetation, buildings, vehicles, and so on. Therefore, classification of airborne LiDAR point cloud data is essential for various applications. Especially, the ground-only point cloud (DTMs: Digital Terrain Models) are used in many applications and are very important data. DSMs are obtained directly by laser scanning, but filtering must be performed to obtain DTMs. This filtering process is the most time-consuming and manpower-intensive process in airborne LiDAR surveying. Even now, many operators are still performing manual task to obtain accurate DTMs. Therefore, it is necessary to improve the efficiency and accuracy of this process.

In order to improve the efficiency of the filtering process, we have been trying to utilise deep learning based techniques. In recent years, examples of applying deep learning to filtering laser data have emerged and are indicating good results. However, deep learning based methods are not perfect and can misclassify. In this paper, we propose several methods to improve the accuracy of deep learning based filtering, and report the results of our actual attempts to perform accordingly.

2. METHODOLOGY

2.1 Baseline method

In recent years, there have been a growing number of examples of deep learning applied to the filtering of airborne LiDAR point clouds. Since deep learning has advanced in the field of image classification, many methods have been used to extend images, such as transforming three-dimensional point clouds into two-dimensional images and applying imaging methods (Marmanis et al., 2015, Hu et al., 2016, Rizaldy et al., 2018, Gevaert et al., 2018), or voxelizing a point cloud (Hackel et al., 2017, Wang et al., 2018) to construct a regular neighborhood structure similar to the image.

In this study, a voxel-based CNN is employed as the baseline method. Voxel-based CNNs are simple and easy to understand. In addition, voxel-based CNNs are very suitable for extending various networks proposed for image recognition to three dimensions. A voxel-based deep learning method was adopted in this study which is a slightly improved version of large-scale point cloud classification (Hackel et al., 2017). The basic procedure described in this paper as follows.
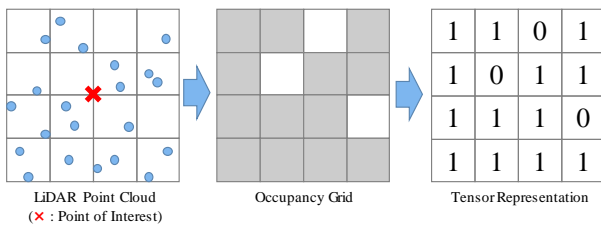
---

* Corresponding author

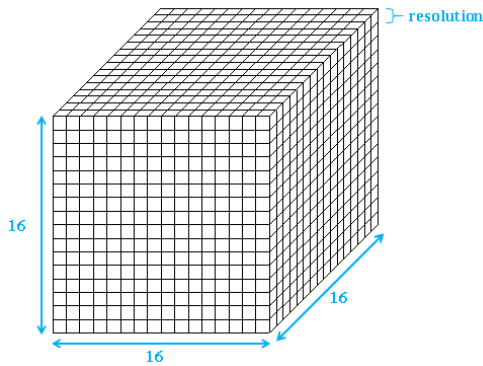Figure 1. Occupancy Grid and Tensor Representation

Figure 2. 16 x 16 x 16 voxels and resolution

At each point in the airborne LiDAR point cloud, the distribution of other points around the point of interest in 3D space was represented as a tensor using an occupancy grid. This tensor was trained by a convolutional neural network (CNN) together with the label of the point of interest (ground surface or other), and a classification model was generated. The number of occupancy grid was 16 x 16 x 16 voxel area (Figure 2). The size of the occupancy grids (resolution) was set to five types (0.2m, 0.4m, 0.8m, 1.6m, and 3.2m), and the tensors generated with each resolution were connected by a fully connected layer. Figure 2 shows the basic network configuration. We considered whether the quality of classification result could be improved by adding some ideas to the baseline method. To improve the quality, we adopted three methods: data pre-processing, changing the network structure, and integrating classification results from different networks.

2.2 Data pre-processing

Our first method is to pre-process the data and consider pre-empting points that is clearly not at the ground surface. The points on the ground surface (DTM) are usually located at the lowest level in the airborne LiDAR point cloud. Therefore, the upper points can be excluded at the beginning as non-ground points. The points floating in the air, such as power lines, are easily misclassified as points on the ground surface, but it can be prevented by excluding the upper points in advance. At the same time, the number of points that are subject to subsequent deep learning (training and classification) process can be significantly reduced, and the process can be more efficient. The pre-processing flow to exclude the upper points of the airborne LiDAR point cloud is described as below.

1) To separate airborne LiDAR point cloud into grids of specific size (here, 1m x 1m) using the XY coordinates.
2) To extract the lowest points in each grid and create a grid lowest point cloud.
3) To retain the points within a certain distance (2 m in our case) from each point in the grid lowest point cloud, and exclude other points.

2.3 Changing the network structure

The second method is to change the network structure used in the baseline method. The baseline method of this study is a VGG-like network. In the image-targeted VGG-Net, it is reported that classification accuracy increases as the number of convolutional layers increases (Shimonan et al., 2015). Also, in general, the wider the range of convolution (i.e., the larger the receptive field), the wider the range of information can be taken into account, and the classification accuracy can be improved. In this study, we adopted this idea and tried to improve the classification accuracy for three-dimensional point clouds by changing the number of convolutional layers and the size of the convolutions among the network structures.

The baseline network used in this study is shown in Figure 3. For the part enclosed by the dashed line, we created and used multiple networks with different number of convolution layers and convolution sizes. The number of layers and convolution size of the used network are shown in Table 1, A-G. In this case, A is the baseline method.
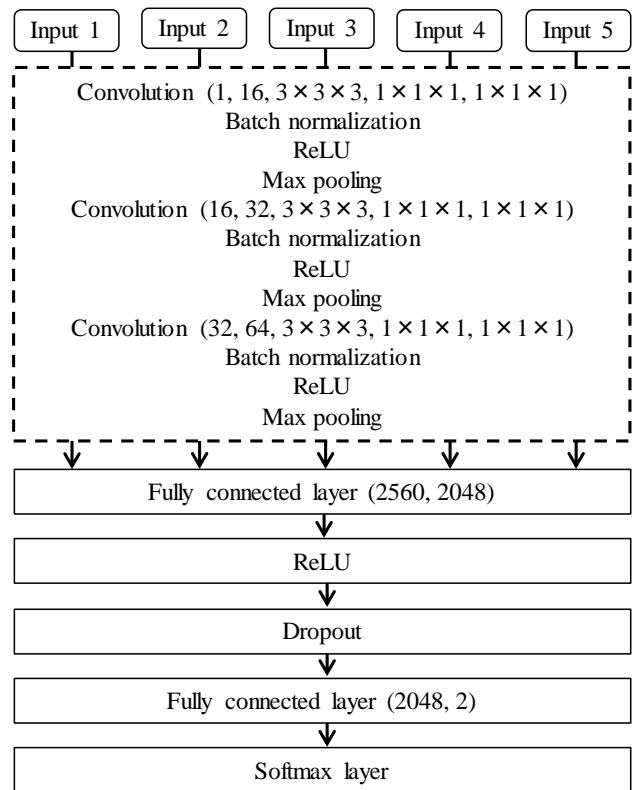

Figure 3. Network Structure

ID	Size of Convolution	No. of Layers
A	3×3×3	3
B	3×3×3	6
C	3×3×3	9
D	5×5×5	3
E	5×5×5	6
F	7×7×7	3
G	7×7×7	6

Table 1. Hyperparameters for the Networks

In each network, the same dataset (labeled airborne LiDAR point cloud) was used to train and classify the data differently from the data for training. Here, we used unprocessed airborne LiDAR point cloud to compare the results.

2.4 Integrating classification results from different networks

The third method is to integrate the classification results across multiple networks. We integrate individual results from multiple networks with varying layer depths and convolution sizes to see if the classification results improve. Subsequently, we use the results of the classification by the seven networks mentioned in section 2.3. The classification results of each network were integrated into three types of ground surfaces.

1) Ground surface having one or more of the seven networks were classified as the ground surface (H)
2) Ground surface with more than half (four or more) of the seven networks were classified as the ground surface (I)
3) Ground surface of all seven networks were classified as the ground surface (J).

These three types of ground surfaces are compared to the result of individual networks (A-G).

3. EXPERIMENTAL RESULTS

3.1 Dataset

The method described in the previous section was applied to the actual airborne LiDAR point cloud data. The data used in this study was airborne LiDAR point cloud acquired for the forest area. The data was divided into two parts, one for training and the other for classification. The data for training and classification have the same topographical features. An overview of the data for the classification is shown in Figure 4.
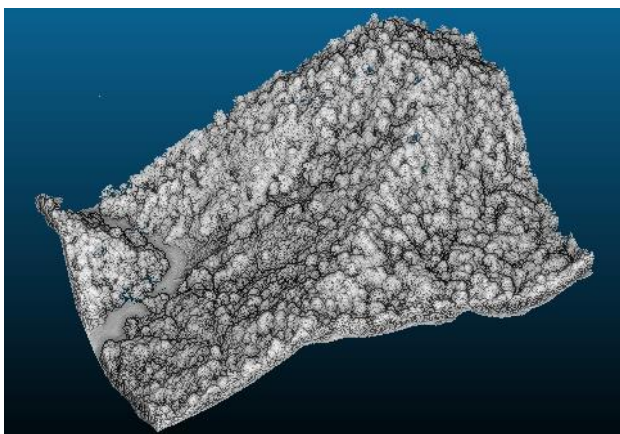

Figure 4. Overview of the classification data

3.2 Data pre-processing

Data pre-processing presented in section 2.2 was performed. The airborne LiDAR points with and without data pre-processing is shown in Figure 5. It can be seen that the pre-processed point cloud exists only in the lower part of original point cloud. The number of points in pre-processed point cloud has been reduced to about 27%.

Next, we trained on the baseline network using the point cloud with and without pre-processing. Afterwards, we used the

model of each network to classify the point cloud. The ground surface of the classified point clouds is shown in Figure 6, and the indices (Precision, Recall, F-measure, and Accuracy) are shown in Table 2.
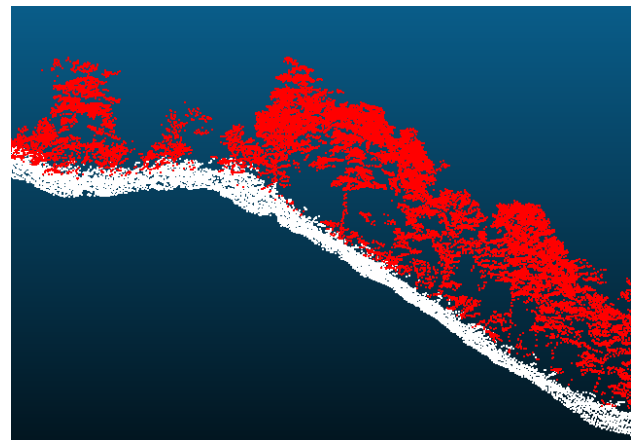

Figure 5. Pre-processed points (white) and original point (red)

	Precision	Recall	F-measure	Accuracy
Baseline	0.769	0.993	0.867	0.947
Pre-processing	0.649	0.669	0.659	0.554

Table 2. Indices of baseline method and pre-processing

Figure 6. Resulting ground surface of baseline method (above) and process with pre-processing (below)

	A	B	C	D	E	F	G	H	I	J
Convolution Size	3×3×3	3×3×3	3×3×3	5×5×5	5×5×5	7×7×7	7×7×7	-	-	-
No. of Layers	3	6	9	3	6	3	6	-	-	-
Precision	0.769	0.776	0.771	0.772	0.770	0.771	0.783	0.758	0.773	0.790
Recall	0.993	0.991	0.993	0.993	0.993	0.993	0.989	0.996	0.993	0.986
F-measure	0.867	0.870	0.868	0.869	0.867	0.868	0.874	0.861	0.869	**0.877**
Accuracy	0.947	0.948	0.947	0.948	0.947	0.947	0.950	0.944	0.948	**0.952**

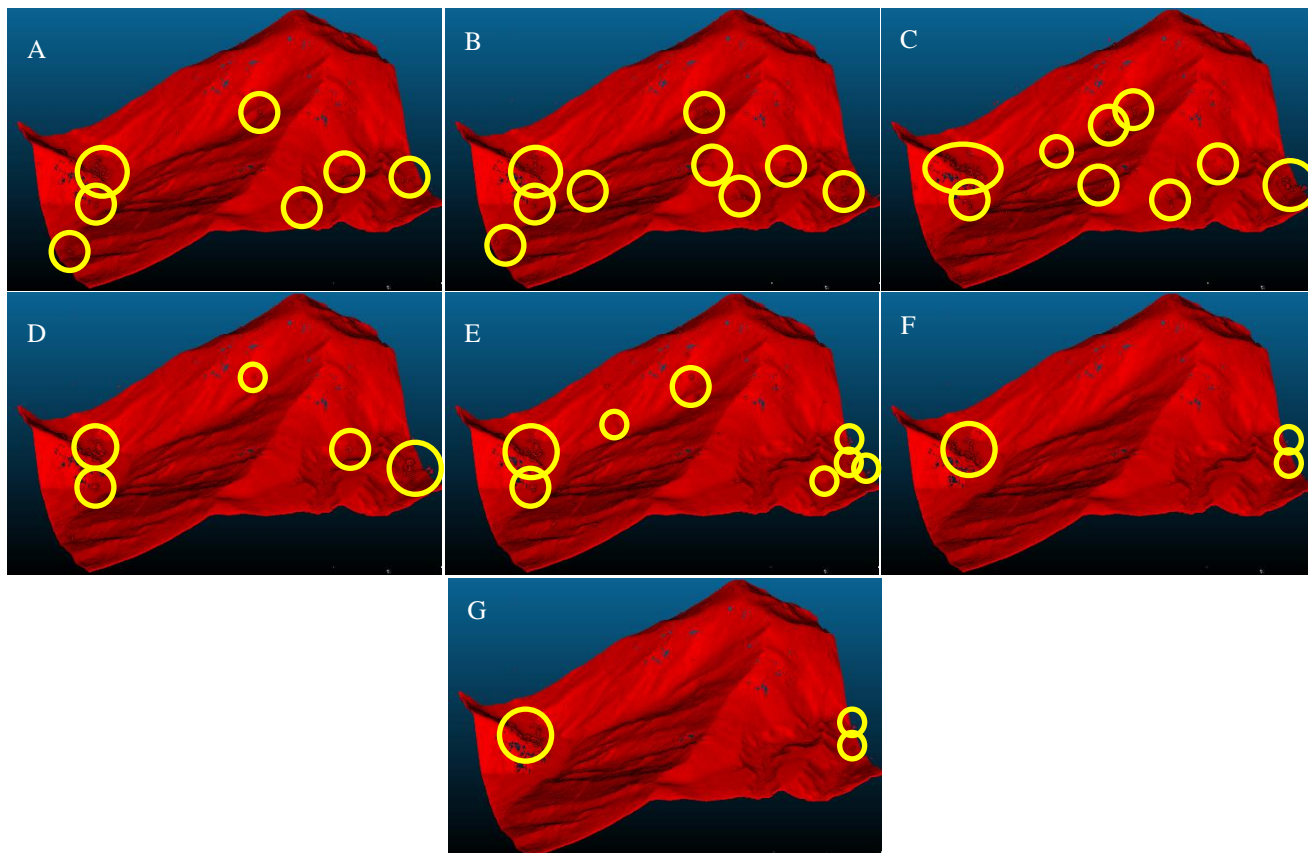Table 3. Classification results by individual network (A-G) and integration (H-J)

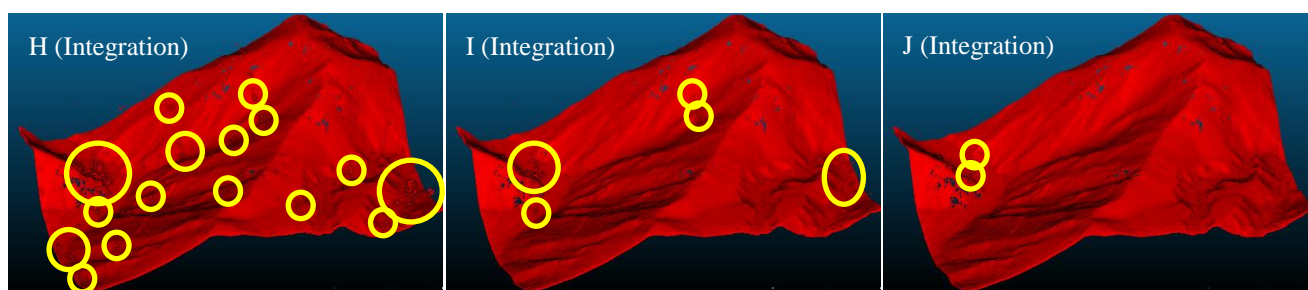Figure 7. Resulting ground surface by each network (A-G)

Figure 8. Resulting ground surface by integration of multiple network results

3.3 Changing the network structure

All networks (A-G) shown in Figure 3 and Table 1 were used to train and classify on a common dataset. The resulting ground surface is shown in Figure 7, and the indices are shown in Table 3. The yellow circles in Figure 7 are notable examples of misclassification that other points (e.g., trees) being misclassified as ground surfaces. This is the same expression described in section 3.4.

3.4 Integrating classification results from different networks

Finally, we aggregated the seven classification results in the previous section and integrated them to create the three types of point clouds presented in section 2.4. The resulting ground surface is shown in Figure 8 and the indices are shown in Table 3 (the table is merged with the results of the previous section).

4. DISCUSSION

The data pre-processing has completely removed the noise floating in the air (Figure 6). In addition, an overview of the ground surface shows that it is satisfactorily classified. However, observing the indices of classification results, the results are not as good as the baseline method. This may be due to the effect of the change in the distribution of points in the laser point cloud. To obtain better results, further consideration is needed.

The classification results by each network show that accuracy ranges from 0.947 to 0.950 and F-measure ranges from 0.867 to 0.874, all of which classify the ground surface significantly. Comparing the results with the number of layers in the networks, there is a slight increase in accuracy as the number of layers increases. Comparing the results with the size of the convolution, it can be observed that the accuracy slightly improved with increasing the size of the convolution. An overview of the resulting ground surface shows that the misclassification of other points as ground surfaces has almost disappeared in cases such as G, where the number of layers is large and the convolution size is large.

When the A-G results were combined to create three different types of ground surfaces, the best results (J) were obtained by utilizing only the points classified as ground surfaces in all networks, Figure 8. The reason for this result is that each network has different classification features and therefore different points of misclassification are observed, but only those points that are reliable due to the integration of all the results that are present on the ground surface. An overview of the resulting ground surface shows that misclassification where other points that are classified as ground surface has almost disappeared. This suggests that the quality of ground surface data can be improved by integrating the results of different networks.

5. CONCLUSION

In this study, we proposed several methods to improve classification accuracy by performing deep learning methods to the filtering process of airborne LiDAR surveying data. We applied these methods to the actual classification of airborne LiDAR point clouds and reported the results.

In the first method, we pre-processed the data by limiting the target point cloud for deep learning to the underlying point cloud. Although the indices of classification accuracy did not improve, an overview of the resulting ground surface data confirmed that misclassifications were removed noticeably. In addition, it was confirmed that the pre-processing can significantly reduce the number of target points and thus improve the efficiency of subsequent process. A suitable network for the point cloud of preprocessing results will be investigated in the future study.

In the second method, we tried several patterns that changed the structure of the network, i.e., the number of convolutional layers and the size of the convolution. As a result, it was confirmed that a tendency to increase the number of convolution layers and the size of convolution, which are generally referred to in deep learning for images, improves the accuracy for 3D point clouds as well.

Finally, in the third method, we tried to integrate the results of classification by multiple networks to produce ground surface data. Three different methods of integrating the results were attempted, and highly satisfactory results were obtained when only points classified as ground surface in all networks were used. The results were better than all the individual classification results, therefore we affirm that the integration of the classification results improve the quality of filtering.

REFERENCES

Gevaert, C. M., Persello, C., Nex, F., Vosselman, G., 2018. A deep learning approach to DTM extraction from imagery using rule-based training labels. ISPRS Journal of Photogrammetry and Remote Sensing, Volume 142, August 2018, pp.106-123

Hackel, T., Savinov, N., Ladicky, L., Wenger, J. D., Schindler, K., Pollefeys, M., 2017. Semantic3D.net: A new Large-scale Point Cloud Classification Benchmark. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-1-W1, pp.91-98.

Hu, X., Yuan, Y., 2016. Deep-Learning-Based Classification for DTM Extraction from ALS Point Cloud. Remote Sens. 2016, 8, 730.

Janssens-Coron, E., Guilbert, E., 2019. Ground Point Filtering from Airborne LIDAR Point Clouds Using Deep Learning: a Preliminary Study. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W13, 2019, pp.1559-1565, https://doi.org/10.5194/isprs-archives-XLII-2-W13-1559-2019, 2019.

Marmanis, D., Fathalrahman, A., Datcu, M., Esch, T., Stilla, U., 2015. Deep Neural Networks For Above-ground Detection in Very High Spatial Resolution Digital Elevation Models. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, II-3/W4, pp.103-110.

Maturana, D., Scherer, S., 2015. VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition. Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on, IEEE, pp. 922–928.

Rizaldy, A., Persello, C, Gevaert, C. M., Oude Elberink, S. J., 2018. Fully Convolutional Networks for Ground Classification from LiDAR Point Clouds. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2.

Shimonyan, K., Zisserman, A., 2015. Very Deep Convolutional Network for Large-Scale Image Recognition. Proceedings of ICLR 2015.

Wang, L., Huang, Y., Shan, J., He, L., 2018. MSNet: Multi-Scale Convolutional Network for Point Cloud Classification. Remote Sens. 2018, 10, 612.

Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X. and Xiao, J., 2015. 3D ShapeNets: A Deep Representation for Volumetric Shapes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1912–1920.