The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-6-2025
ISPRS, EARSeL & DGPF Joint Istanbul Workshop "Topographic Mapping from Space" dedicated to Dr. Karsten Jacobsen's 80th Birthday
29–31 January 2025, Istanbul, Türkiye

Evaluate the Impact of Class Granularity in Point Cloud Semantic Segmentation on DTM Accuracy

Haval AbdulJabbar Sadeq [1]

[1] Geomatics (Surveying) Engineering Department, Salahaddin University-Erbil, Erbil, Iraq - haval.sadeq@su.edu.krd

Keywords: Airborne LiDAR, digital terrain model, DTM, Ground Filtering, deep learning, semantic segmentation

Abstract

The Digital Terrain Model (DTM) is considered an essential component in various applications, including road design, urban planning, terrain analysis, and, environmental monitoring. LiDAR data is known to have very high accuracy therefore it is considered the most reliable source for DTM generation. However, accurately filtering the LiDAR data for the ground classification remains a challenge. This study explores the impact of class granularity on semantic segmentation and its effect on the accuracy of DTM generation sourced to LiDAR data. The RandLA-Net has been used for semantic segmentation, and for the training, the DALAS dataset which comprises various terrains and structures is used. The process is comprised of training the deep learning models on datasets that are classified into two schemes. The first scheme is a course with 2-classes (ground and non-ground) and the second scheme is a finer one with 8-classes (ground, vegetation, cars, trucks, powerlines, fences, poles, and buildings). The trained models are applied to three datasets to evaluate the result of the granularity on the accuracy of the DTM. The results show that the accuracy of the DTM based on the 2-class model outperforms the accuracy of the DTM obtained via the 8-classes model, as indicated by the used statistical measures such as RMSE, mean and standard deviation (STD). The semantic segmentation of the 8-classes model shows more misclassification, especially in complex urban areas, especially in distinguishing the ground points from non-ground objects. The study emphasizes the trade-off between class granularity and DTM accuracy, which shows that simpler classification schemes will lead to a better result of the DTM generation results.

1. Introduction

Digital Terrain Model (DTM) which provides an accurate representation of the bare ground surface is considered to be an essential component in different applications such as road design (AL-Areeq et al., 2023), forest monitoring (McCarley et al., 2020), urban planning (Dimitrov and Petrova-Antonova, 2021; Olivatto et al., 2023), and environmental and terrain analysis (AL-Areeq et al., 2023; Conforti et al., 2020; Wu et al., 2023). Recently, the main source for the DTM generation is either from LiDAR (Light Detection and Ranging) or photogrammetry. Despite the high accuracy of the LiDAR data, the process of filtering the ground point is still considered an important challenge (Mesbah et al., 2023). As a result, the accuracy of creating the DTM models mainly depends on point cloud classification or semantic segmentation, where the points are classified into different classes such as ground, buildings, and vegetation. For DTM generation, the ground points are considered to be important.

Traditional methods for the ground point filtering from LiDAR data have been very popular and widely used in the filtering ground point, such as Progressive TIN Densification (PTD) (Axelsson, 2000), Progressive Morphological Filter (PMF) (Zhang et al., 2003), Cloth Simulation Filtering (CSF) (Zhang et al., 2016), scale-irrelevant and terrain-adaptive approaches (Chen et al., 2021), and Bayesian approach (Sadeq, 2024), These algorithms are considered to be very successful in many cases, however, they struggle in complex terrain, slopped areas or regions with small objects near the ground, leading to inaccurate DTMs (Qin, Tan, Guan, et al., 2023). On the other hand, deep learning has become an active research area for point cloud filtering which aims to improve the results for point cloud classification, specifically in the challenging areas (Hu and Yuan, 2016; Qin, Tan, Ma, et al., 2023).

Various deep learning based semantic segmentations for the DTM generation have been developed. For instance, Rizaldy et al.(2018) applied a Fully Convolutional Network (FCN) on the converted point cloud into a single pixel thus can deal more robustly in the slopped terrain. Nurunnabi et al. (2021) developed a deep learning model based on the combination of local features instead of depending on the extracted features from raw data. Further to improve the accuracy, researchers have advanced in the field of deep learning. Winiwarter et al. (2019) enhanced the deep learning algorithm PointNet++ (Qi et al., 2017), an end-to-end algorithm which helps to infer from local neighbourhood features without the need for prior data. Jin et al., (2020) proposed an algorithm for filtering LiDAR point cloud which achieves better accuracy and performance compared to the well-known deep learning algorithms. Despite all these advancements in deep learning for the point cloud filtering algorithms, still none of these algorithms able to filter the ground points robustly to achieve 100% accuracy.

The process of LiDAR semantic segmentation entails assigning a predefined class to each point, these include features such as ground, buildings, vegetation, and electrical poles. Previous studies have explored the impact of various data strategies on data preparation to obtain better semantic segmentation results (Zou et al., 2021). For example, Wang and Yao (2022) utilized entropy regularization, sparse annotation, ensemble prediction constraints, and online soft pseudo-labelling to improve the result of the semantic segmentation. However, To the best of our knowledge, no studies investigated the effect of the number of classes on the accuracy of the DTM.

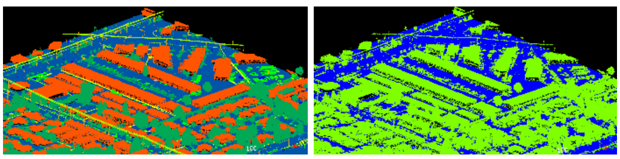
This study examines the impact of class granularity on the efficiency and accuracy of DTM generation by comparing deep learning models that are trained with different class schemes. Class granularity in the Semantic segmentation refers to less details of the classes in the scheme. The level of detail used in the point cloud semantic segmentation can influence the precision of the DTM. The coarser granularity will probably lead to simplifying the segmentation process and leading to better accuracy. On the other hand, the finer segmentation probably improves feature detection, meanwhile, it also may lead to introducing misclassification errors.

The structure of the study is as follows: the introduction and background studies are illustrated in this section, the used datasets are outlined in section 2, and the theoretical details of the deep learning and the used methodology are explained in section 3. The evaluation of the obtained result is illustrated in section 4, and the result and discussion are given in section 5. Finally, the conclusion is shown in section 6.

2. Datasets

In this study, the training data used for training the deep learning model is sourced from the DALES dataset (Varney et al., 2020), as shown in Figure 1. The datasets was classified into eight classes as indicated in Figure 1(a). These eight classes include ground, vegetation, cars, trucks, powerlines, fences, poles, and buildings.

To evaluate the effect of class granularity, the DALES dataset was modified into a 2-classes scheme, as shown in Figure 1(b). in the modified scheme, the two main classes are ground and non-ground. The non-ground points are obtained by combining all the non-ground classes from the original DALES dataset except the ground class.



(a) (b)

Figure 1 The DALES LiDAR datasets in the training of the RandLA-Net. (a) A sample of the point cloud is classified into 8-classes which include: ground, vegetation, cars, trucks, powerlines, fences, poles, and buildings. (b) A sample of the point cloud with two classes only, ground and non-ground.

For deep learning, the dataset has been classified into training and validation. Initially, the density of the dataset was very large which led to an effect on the computation efficiency. Therefore, it has been reduced by 75%, retaining only 25% of the original dataset. Furthermore, for the training only four tiles were used, and for validation only one tile was used. The tile size was measured to be 500 m by 500 m each, meanwhile, for the DTM generation and testing of the trained deep learning models, three tiles have been used as explained in section 3.3. Since the study is primarily for comparing class granularity schemes, therefore it is assumed that the minimizing point could not have affected the result.

3. Methodology

3.1 Point Cloud Semantic Segmentation

For assessing the impact of the granularity effect on the produced semantic segmentation, the RandLA-Net deep learning architectures have been selected. The RandLA-Net deep learning approach is based on 3D deep neural networks (DNN) that are used for encoding the LiDAR point cloud for the purpose of semantic segmentation. RandLA-Net developed by Hu et al. (2020), which is specifically designed for classifying large-scale point clouds. The RandLA-Net architecture includes a model that maintains the geometric details during the point sampling, furthermore, the module is based on a neural architecture which is specified to be efficient and lightweight and thus can be used for semantic labelling of large-scale point clouds. The main key advantage of the model is based on leveraging random sampling, which reduces memory consumption and computational cost, in contrast to the other segmentation frameworks in which their strategies are dependent on expensive sampling methods. RandLA-Net was selected for this study since is proven to be very efficient in point cloud semantic segmentation and compatible with ArcGIS-pro software. Moreover, it delivers a higher processing speed of processing compared to PointCNN which is shown to be slower and almost delivers the same result (Hu et al., 2020).

3.2 Training Deep Learning Models

The RandLA-Net model training process is applied in two stages. In the first stage, datasets consisting of 2-classes (ground and non-ground) result in a model known as 2-class model. In the second stage, the datasets that consist of 8-classes are used leading to obtaining an 8-class model. By applying the above two approaches in training the deep learning models, it is possible to conclude how the class granularity is affecting the accuracy of the generated DTM.

For the training, the datasets were divided into blocks with a size of 20 m assuming they are covering the object geometry within the study area. Each model was trained individually up to 25 epochs, and 100 % of iterations per epoch in order to ensure that all data is passed through the epoch. The learning rate was set to 0.005, and the batch size was set to 10 which means 10 patches were simultaneously processed within the model.

3.3 DTM generation

The obtained trained models (2-class and 8-class) were applied to test datasets, which were not used for the training or validation. These datasets were derived from LiDAR data that belong to the DALES dataset (Varney et al., 2020). The datasets have different types of man-made structures such as buildings, fences, poles and powerlines, in addition to the vegetation.

By applying the trained model, a semantic point cloud has been obtained for each study area as shown in Figure 2. The result of the 2-class models includes two classes, the first class includes bare ground. While the other class includes buildings, trees plants and other structures as shown in Figure 2 (a). In contrast, the 8-class model assigned a separate class to each type of object. For instance, the ground class, while buildings and other structures are each given different classes, which are considered 7 classes, as shown in Figure 2 (c).

To produce the DTM, only the ground class from each semantic segmentation was kept while other classes were eliminated. Later the remained point cloud will be interpolated for the DTM generation as shown in Figure 2 (b) and (d). This same procedure was applied to all datasets.

https://doi.org/10.5194/isprs-archives-XLVIII-M-6-2025-243-2025 | © Author(s) 2025. CC BY 4.0 License.

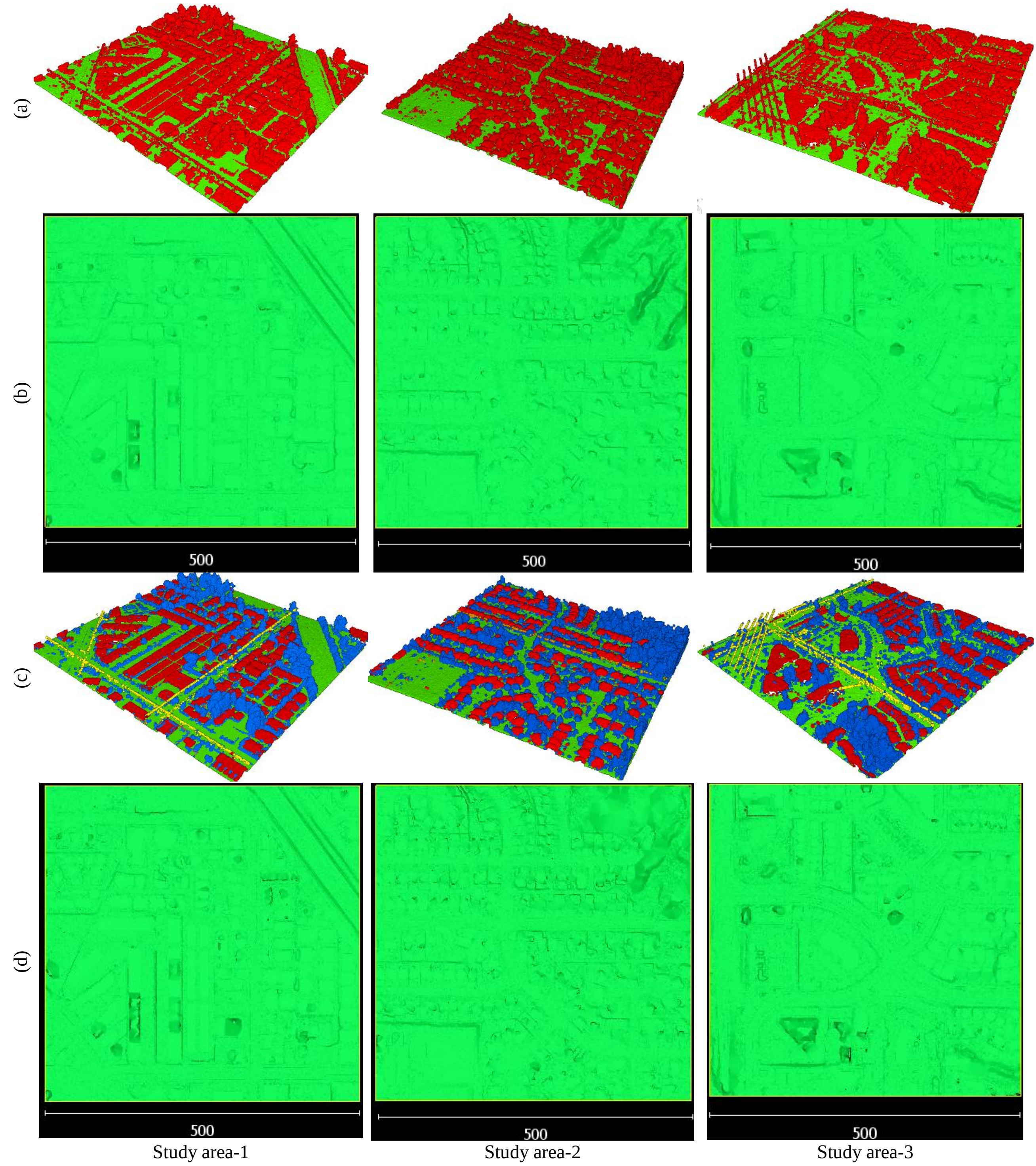


Figure 2 Semantic segmentation of the point cloud and the corresponding DTM processed with each model. (a) Segmented point cloud using the 2-class model. (b) Produced DTM based on using the ground class obtained from a 2-class model. (c) Segmented point cloud produced from an 8-class model. (d) Produced DTM based on using the ground class obtained from an 8-class model.

4. Evaluation

Several indices are available for the evaluation of the GF algorithm, such as the intersection over union for non-ground class (IoU1), intersection over union for ground class (IoU2), overall accuracy, precision, recall, F1-Score, and, Root Mean Square Error (RMSE) (Qin, Tan, Guan, et al., 2023). In this study, to quantify the error and assess the accuracy of the produced DTM the indices RMSE, mean and STD are used. These indices are mainly utilized for comparison purposes in order to assess which of the algorithms produces better results.

The DTMs generated by both models were compared with the true DTM, which was obtained by manually extracting the ground points. The true DTM was used as a reference and has been subtracted from the labelled DTM. For the subtraction process, the cloud compare software has been used (EDF R&D, 2011).

The difference or distance map, which is known as the C2M map, was obtained by subtracting the DTM generated from the semantic segmentation of the point cloud from the true DTM, as shown in Figure 3. The mean and STD were determined using the CloudCompare software, while the RMSE was calculated from the distance map (C2M map). For this purpose, the C2M map was exported to the txt file which includes vertical distance between the two DTMs. Later, by using these differences the RMSE is determined by developing a code with C++.

The results were calculated for the three study areas as shown in Table 1, including RMSE, mean and STD values that belonged to the DTM produced from the 2-classes model. The C2M approach is based on finding the changes between the point cloud and a true 3D mesh, this approach is considered to give a good result in the case of a flat surface such as comparing two DTMs as the case in this study (Lague et al., 2013).

In addition to the quantitative analysis, the qualitative analysis was also applied, by comparing and identifying the errors in the difference map. It is clear that some extra errors are clearly identified in the difference map that is obtained by the 8-class model, which has been identified by the red dashed circle in Figure 3. Further illustrations are given in Figure 4. Misclassification in the building was detected, which caused to deterioration of the accuracy of the produced DTMs.

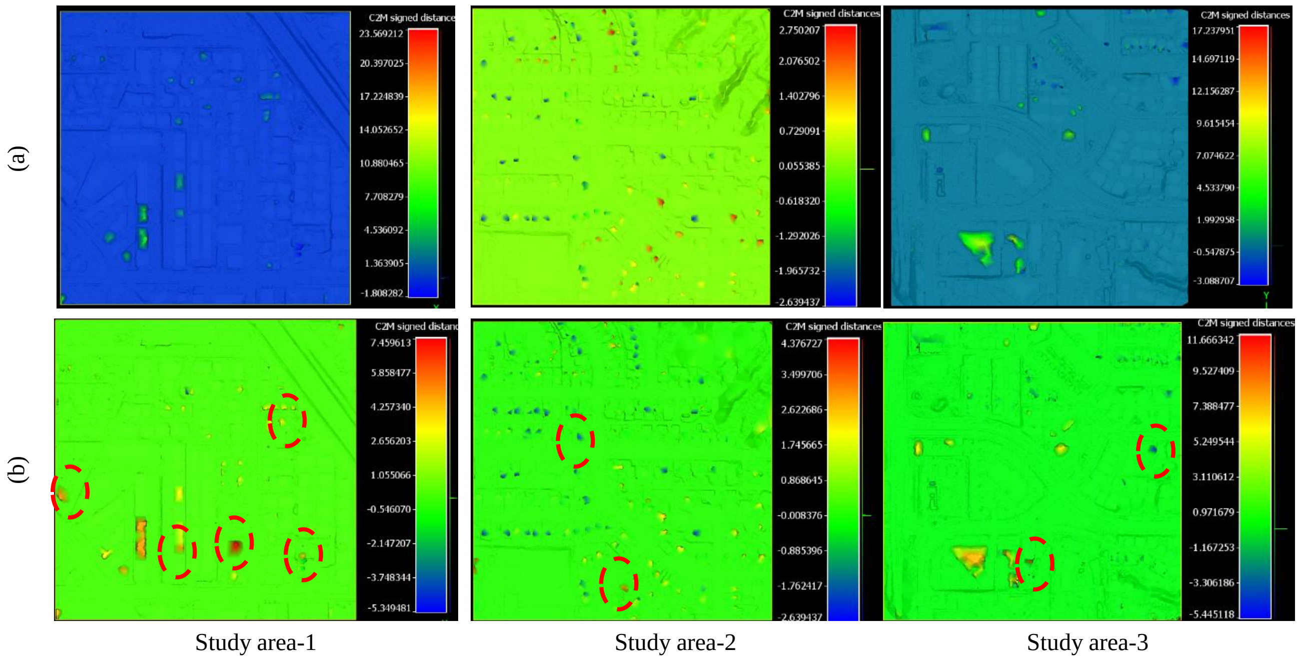


Figure 3 Difference map between the True DTM and the generated DTM. (a) The difference map was obtained by subtracting the True-DTM from 2-classes model DTM. (b) The difference map was obtained by subtracting the True-DTM from 8-classes model DTM. The red dashed circle identifies the locations of errors.

5. Result and Discussion

The main aim of this study is to examine the impact of class granularity on the accuracy of the DTM generated using a deep learning algorithm. Mainly, to investigate how changes in the level of the classes in semantic segmentation will affect the precision of ground point filtering from LiDAR data on the DTM generation. In particular, the performance of two deep learning models on DTM generation, that trained on datasets with two different class schemes is compared.

The finding indicates that the DTM obtained from ground filtering with 2-class model outperforms the DTM obtained from filtering the ground points using 8-class model. This is supported by the statistical analysis using the RMSE, mean and STD, as shown in Table 1, where is shown that the 2-class model gives the best result. Furthermore, it can be noticed that the number of points in the 2-class model is less than 8-class model which proves that some points have been misclassified and classified wrongly as a ground class. Furthermore, from the table, it can be noticed that the best accuracy is from the second study area which is shown to be 0.041760 m, compared to RMSE values of 0.081832 m, and 0.137628 in the first and third study areas, respectively. This discrepancy is because these study areas include large buildings which made the segmentation fail, mainly because the selected block size of 20 m during the model training was smaller than the actual size of the buildings.

The analysis shows that a simpler classification scheme (e.g. 2-class model) helps to reduce misclassification and streamline the segmentation process, particularly in problematic areas such as complex terrain or surface discontinuity areas. Reducing the level of detail (i.e. simplifying the categories), improves the model's ability to focus on the interest class which is the ground points, thus leading to more accurate DTM.

Conversely, the semantic segmentation with the model of 8-classes, which leads to classifying the point cloud into separate classes (e.g. ground, vegetation, cars, trucks, powerlines, fences, poles, and buildings) will lead to misclassification by hampering the model’s ability to classify the point cloud correctly. This finer granularity is problematic, especially in complex terrains or urban environments. The errors introduced by 8-class model are clearly shown in Figure 3 indicated by red dashed circles.

Based on the hypothesis in this paper, the study shows a trade-off between the level of class detail in semantic segmentation and misclassification. Although finer classification, such as 8-class model, theoretically gives better distinction of the features, it might lead to more errors in ground point classification which is considered to be critical in DTM generation. Misclassification of non-ground points as ground points or in reverse will lead to degradation of the DTM quality and reduce the accuracy. The 2-class model, by simplifying the segmentation into two classes only, ignores finer non-ground classes but performs better for identifying ground points.

	Study area-1 (The number of the points in the true DTM is 6,022,110)		Study area-2 (The number of the points in the true DTM is 5,470,158)		Study area-3 (The number of the points in the true DTM is 5,470,158)	
	2-class model	8- class model	2-class model	8- class model	2-class model	8- class mode
Number of points	**5,618,863**	5,767,074	**4,955,635**	5,100,182	**6,138,907**	6,306,851
Mean (m)	**0.001025**	0.004225	**0.000334**	0.001799	**0.002478**	0.007910
STD (m)	**0.081826**	0.134463	**0.041759**	0.065963	**0.137606**	0.231390
RMSE (m)	**0.081832**	0.134529	**0.041760**	0.065987	**0.137628**	0.231525

Table 1. Statistical analyses for the generated DTMs, for three study areas, using different deep learning models based on RANDLA-Net. The bold fonts refer to the best values.

By further analysis of the errors and their locations, it can be noticed that the misclassification of the buildings with ground points is clearly shown in Figure 4(b). It can be noticed that the 2-class model Figure 4 (a and c) clearly labelled the ground points as ground and the non-ground points, such as buildings and trees, without misclassifying them. However, in the 8-class models (Figure 4(b) and the zoomed-in portion(d)), buildings are misclassified as ground points which leads to introduced errors in the produced DTM. Further examination reveals that the buildings are also misclassified as trees, but since all the non-ground classes are eliminated, this misclassification will not affect the produced DTM.

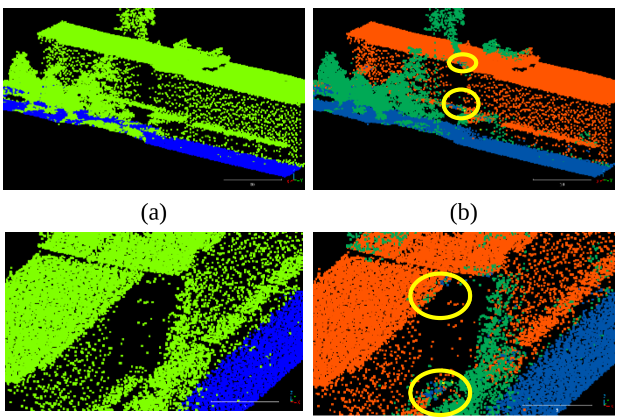


Figure 4 Comparison of semantic segmentation from 2-class and 8-class models. (a) Semantic segmentation with 2-class model, where non-ground buildings are correctly classified. (b) Semantic segmentation using 8-calss model, where buildings are misclassified as ground points, as marked with yellow circles. (c) and (d) illustrated a zoomed-in part of the highlighted area.

In the context of model selection, the RandLA-Net architecture has been demonstrated in order to efficiently process the semantic segmentation of the point cloud. It has been proven its efficiency in processing large-scale LiDAR point cloud segmentation. The selected lightweight deep learning architecture, which depends on random sampling, helps to reduce memory and computation costs. This technique proved to be performed accurately for the purpose of semantic segmentation. Moreover, its compatibility with the ArcGIS Pro makes it suitable for easy access by the user even those without high expertise in the programming language.

The implementation of the deep learning approach, as an alternative to the traditional filtering method is considered to be a remarkable advancement. Traditional methods such as CSF and PTD often fail in complex areas or with the object of small objects near the ground. However, deep learning models are more robust in such areas and deliver superior performance, offering more accurate ground filtering and generating precise DTM.

Although the study provided a remarkable insight into the impact of class granularity on the DTM generation, it has some limitations. For instance, only one dataset (e.g. DALES) was used, which might not represent all types of topography. Therefore, it is necessary to test different datasets representing different topography in another region or use more datasets in training the models.

When comparing the obtained result with the other studies, it is evident that the accuracy obtained by RandLA-Net using 2-class model outperforms the traditional algorithm. In terms of the accuracy of the DTM, various studies have assessed LiDAR point cloud segmentation algorithms. For example, Song and Jung, (2023) evaluated various lidar filtering algorithms for DTM generation, comprising LAStools, CSF, and PMF. The best DTM belonged to the LAStools, with an RMSE of 0.3 m, based on using Purdue University dataset, while Salleh et al. (2015), reported an accuracy of 0.379 m for the DTM low slopes terrain. Thus, it can be concluded that the obtained result in this study shows higher accuracy than the aforementioned studies.

6. Conclusion

This paper evaluates the impact of the class granularity in semantic segmentation on the accuracy of the obtained DTM. The study compares deep learning models based on the RandLA-Net architecture, which are trained using different class schemes. It has been found that, in terms of accuracy, the models generated via a scheme that consisted of 2-classes (e.g. ground and non-ground) provide better results than 8-classes scheme (e.g. ground, vegetation, cars, trucks, powerlines, fences, poles, and buildings). The accuracy has been measured by the RMSE, mean and STD values, all of which show that the model of 2-class scheme outperforms the model of 8-class scheme.

The analysis reveals that using the coarser class scheme will lead to improved results and higher DTM accuracy than the finer scheme, such as the 8-class approach. It was found that a 2-class scheme streamlines the segmentation process, by focusing only on the specific point cloud, thereby the misclassification is reduced, which is typically obtained from finer segmentation.

The findings highlighted a relationship between the segmentation details and the accuracy of the DTM. More detailed features lead to increased misclassification, thus reducing the accuracy, particularly in complex terrain. Moreover, the RandLA-Net architecture has proven its applicability for processing large

https://doi.org/10.5194/isprs-archives-XLVIII-M-6-2025-243-2025

LiDAR point cloud datasets, making it a promising option for DTM generation in large-scale areas. This suggests a strong candidate for the point cloud semantic segmentation task.

Acknowledgement

The author would like to thank the Esri company for providing the ArcGIS-pro license for the purpose of running the deep learning algorithm. Also, the author would like to express their sincere gratitude to Salahaddin University-Erbil for their financial support to achieve this research.

References

AL-Areeq, A.M., Sharif, H.O., Abba, S.I., Chowdhury, S., Al-Suwaiyan, M., Benaafi, M., Yassin, M.A., et al., 2023. Digital elevation model for flood hazards analysis in complex terrain: Case study from Jeddah, Saudi Arabia. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier B.V., 119, 103330.

Axelsson, P., 2000. DEM Generation from Laser Scanner Data Using Adaptive TIN Models. *International Archives of Photogrammetry and Remote Sensing*, Vol. 33 No. 4, pp. 110–117.

Chen, C., Chang, B., Li, Y., Shi, B., 2021. Filtering airborne LiDAR point clouds based on a scale-irrelevant and terrain-adaptive approach. *Measurement: Journal of the International Measurement Confederation*, Elsevier Ltd, 171, 108756.

Conforti, M., Mercuri, M., Borrelli, L., 2020. Morphological Changes Detection of a Large Earthflow Using Archived Images, LiDAR-Derived DTM, and UAV-Based Remote Sensing. *Remote Sensing*, 13(1), 120.

Dimitrov, H., Petrova-Antonova, D., 2021. 3D CITY MODEL AS A FIRST STEP TOWARDS DIGITAL TWIN OF SOFIA CITY. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B4-2(B4-2021), 23–30.

EDF R&D, T.P., 2011. CloudCompare (version 2.11). *[GPL Software]*., available at: http://www.danielgm.net/cc/ (accessed 20 October 2022).

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., et al., 2020. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 11105–11114.

Hu, X., Yuan, Y., 2016. Deep-Learning-Based Classification for DTM Extraction from ALS Point Cloud. *Remote Sensing*, 8(9), 730.

Jin, S., Su, Y., Zhao, X., Hu, T., Guo, Q., 2020. A Point-Based Fully Convolutional Neural Network for Airborne LiDAR Ground Point Filtering in Forested Environments. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 3958–3974.

Lague, D., Brodu, N., Leroux, J., 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS Journal of Photogrammetry and Remote Sensing*, International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS), 82, 10–26.

McCarley, T.R., Hudak, A.T., Sparks, A.M., Vaillant, N.M., Meddens, A.J.H., Trader, L., Mauro, F., et al., 2020. Estimating wildfire fuel consumption with multitemporal airborne laser scanning data and demonstrating linkage with MODIS-derived fire radiative energy. *Remote Sensing of Environment*, 251, 112114.

Mesbah, S., Madbouly, M.M., Ghanem, M., 2023. Error Assessment and Propagation in Digital Terrain Modeling: a Review. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 48(1/W2-2023), 611–617.

Nurunnabi, A., Teferle, F.N., Li, J., Lindenbergh, R.C., Hunegnaw, A., 2021. An efficient deep learning approach for ground point filtering in aerial laser scanning point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43(B1-2021), 31–38.

Olivatto, T.F., Inguaggiato, F.F., Stanganini, F.N., 2023. Urban mapping and impacts assessment in a Brazilian irregular settlement using UAV-based imaging. *Remote Sensing Applications: Society and Environment*, 29, 100911.

Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. PointNet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in Neural Information Processing Systems*, 5100–5109.

Qin, N., Tan, W., Guan, H., Wang, L., Ma, L., Tao, P., Fatholahi, S., et al., 2023. Towards intelligent ground filtering of large-scale topographic point clouds: A comprehensive survey. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier B.V., 125, 103566.

Qin, N., Tan, W., Ma, L., Zhang, D., Guan, H., Li, J., 2023. Deep learning for filtering the ground from ALS point clouds: A dataset, evaluations and issues. *ISPRS Journal of Photogrammetry and Remote Sensing*, 202, 246–261.

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 the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(2), 231–238.

Sadeq, H., 2024. A Bayesian framework for 3D point cloud filtering. *Journal of Spatial Science*, Taylor & Francis, 00(00), 1–24.

Salleh, M.R.M., Ismail, Z., Rahman, M.Z.A., 2015. ACCURACY ASSESSMENT OF LIDAR-DERIVED DIGITAL TERRAIN MODEL (DTM) WITH DIFFERENT SLOPE AND CANOPY COVER IN TROPICAL FOREST REGION. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-2/W2(2W2), 183–189.

Song, H., Jung, J., 2023. An Object-Based Ground Filtering of Airborne LiDAR Data for Large-Area DTM Generation. *Remote Sensing*, 15(16). doi.org/10.3390/rs15164105.

Varney, N., Asari, V.K., Graehling, Q., 2020. DALES: A large-scale aerial LiDAR data set for semantic segmentation. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 717–726.

Wang, P., Yao, W., 2022. A new weakly supervised approach for ALS point cloud semantic segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188, 237–254.

Winiwarter, L., Mandlburger, G., Schmohl, S., Pfeifer, N., 2019. Classification of ALS Point Clouds Using End-to-End Deep Learning. *PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, Springer International Publishing, 87(3), 75–90.

Wu, Y., Chen, Y., Chen, R., Chang, K., 2023. Applications of Multi-temporal DTMs in Mining Management and Environmental Analysis, V, 15511.

Zhang, K., Chen, S.C., Whitman, D., Shyu, M.L., Yan, J., Zhang, C., 2003. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(4) PART I, 872–882.

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., Yan, G., 2016. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sensing*, 8(6), 1–22.

Zou, Y., Weinacker, H., Koch, B., 2021. Towards urban scene semantic segmentation with deep learning from LiDAR point clouds: a case study in Baden-Württemberg, Germany. *Remote Sensing*, 13(16), 3220. doi.org/10.3390/rs13163220.