

Geometry-Aware Semantic Classification of LiDAR Point Clouds Using Image-Based Labels

Ozturk Ozan ^{1,*}, Seker Dursun Zafer²

¹ Faculty of Engineering and Architecture, Recep Tayyip Erdogan University, Rize, 53100, Türkiye, ozan.ozturk@erdogan.edu.tr

² Faculty of Civil Engineering, Istanbul Technical University, Istanbul, 34475, Türkiye, seker@itu.edu.tr

*Corresponding author

Abstract

The classification of LiDAR point clouds remains challenging due to the inherently irregular structure of the data and its lack of spectral information. In this study, a hybrid classification approach that transfers semantic information from labelled orthophoto images to the point cloud through spatial matching is proposed. To evaluate the geometric consistency of the transferred semantic information and to improve classification success, comprehensive feature analyses were performed on the point cloud. The vertical distribution of classes was examined using statistics such as average height and the standard deviation of height. The three-dimensional morphological structure of objects was characterized using linearity, planarity, and sphericity indices derived from eigenvalues of the local covariance matrix. The results showed that the tree class exhibited irregular structure with high vertical variability and surface roughness, together with low planarity and high sphericity. The results also showed that buildings and impervious surfaces are defined by dominant planar geometries with high planarity and low sphericity. Low vegetation and impervious surface classes showed overlapping geometric features in some tiles, whereas the tree class remained separable across tiles. These findings indicate that the integration of orthophoto-derived semantic information with point cloud geometric features establishes a robust and practical classification framework for LiDAR data.

Keywords: LiDAR point cloud classification, 2D–3D data fusion, semantic label transfer, urban point clouds, eigenvalue-based geometric features.

Introduction

LiDAR (Light Detection and Ranging) is a remote sensing method that enables the rapid, highly accurate acquisition of three-dimensional (3D) data of Earth's surface and objects on it. The acquired point cloud data serves as a primary source in applications such as urban modeling, environmental monitoring, and land use analysis. With the widespread integration of platforms such as Unmanned Aerial Vehicles (UAVs), the increasing availability of LiDAR data has made the effective processing and analysis of point clouds a critical research topic [1].

Classifying LiDAR point clouds by land-use categories is a significant challenge. Although LiDAR data is geometrically accurate, but it is disorder structure and does not contain semantic information related to object classes [2]. Different object forms can exhibit overlapping values in geometric descriptors, reducing class separability and complicating classification. Additionally, large-scale high-density point clouds increase the computational burden of point-wise analysis. For these reasons, there is a growing need for reliable classification approaches that maintain methodological simplicity while effectively distinguishing among different land-use categories.

Deep learning methods for point cloud classification have seen remarkable development in recent years. Pioneering studies such as PointNet [3] and the extended version PointNet++ [4], with its hierarchical feature learning capability, represented foundational contributions to the end-to-end learning on point clouds. Subsequently, several new approaches emerged from

these foundational architectures, including convolution-based methods for point clouds [5, 6], graph-based networks for modeling inter-point relationships [7, 8], and attention mechanisms and transformer-based models [9, 10]. However, these methods have practical limitations, such as high model complexity and reliance on large labeled datasets. Consequently, simpler, more interpretable, and more directly applicable alternatives remain in demand, particularly in operational applications [11, 12].

In this study, a spatial integration of 2D semantic images with a 3D LiDAR point cloud is proposed, enabling the transfer of image-based labels to the point cloud. The point cloud geometry is subsequently evaluated through data-driven local ground estimation and normalized height calculation. The transferred semantic information is then refined through multiple processing steps based on vertical profile analysis, spatial contextual refinement, and morphological constraints. Together, these steps constitute an end-to-end classification framework that leverages both image-derived semantic information and the three-dimensional structure of the point cloud.

Data

The ISPRS 2D Semantic Labeling Contest Potsdam dataset, which provides LiDAR point clouds, labeled orthophoto images, and corresponding land use label maps [13]. The dataset comprises 24 labeled tiles, each containing approximately 36 million LiDAR points and corresponding labelled orthophoto images at a spatial resolution of 6000×6000 pixels. 10 tiles were used for analysis in this study. Additionally, ground truth labels include six semantic classes. However, due to potential spatial misalignment between the car labels and the point cloud, the car and clutter or background classes were excluded from the analysis. Consequently, the study focused on four classes, which are impervious surfaces, building, low vegetation, and tree.

Method

The general workflow encompassing all processing steps from semantic label transfer to the final classified point cloud is illustrated in Figure 1.

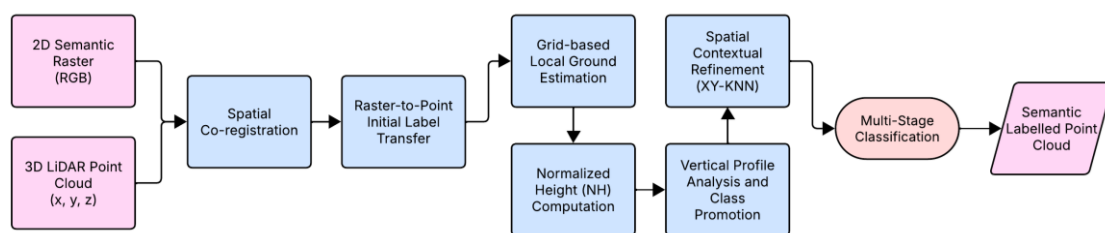


Fig. 1. Workflow of the proposed 2D–3D semantic labeling framework.

Spatial Co-registration and Raster-to-Point Initial Label Transfer

In the first stage, the two-dimensional image data and the three-dimensional LiDAR point cloud were spatially co-registered, enabling label transfer from the labelled images to the point cloud through spatial matching. For each LiDAR point, the corresponding image pixel was identified and its class label assigned, generating an initial semantic annotation of the point cloud before geometric analysis.

Grid-based Local Ground Estimation and Normalized Height Computation (NH)

Height normalization was applied to remove the influence of terrain elevation and to represent the relative vertical structure of objects within the point cloud. At this stage, the point cloud was split into regular grids in the horizontal plane, and the lower percentile of height values within each grid was adopted as the local ground reference. The normalized height value for

each point was then computed by subtracting this local ground reference from the absolute elevation of each point. Negative values, which may arise from inaccuracies in the grid-based ground reference, were set to zero.

Vertical Profile Analysis and Class Promotion

Grid-based vertical analysis was applied to correct labeling inconsistencies arising from the two-dimensional nature of orthophoto-derived labels, such as misclassified building facades and ground points under tree canopies. Points with a normalised height exceeding a defined threshold were identified as candidates for reclassification, regardless of their initial label. These points were reclassified into the most geometrically consistent class based on their proximity to the class-wise mean height profiles of the building and tree classes within the corresponding grid.

Spatial Contextual Refinement

To enhance the spatial consistency of classification and reduce local classification noise, a nearest-neighbour contextual refinement was applied. For each candidate point, only neighbouring points within a defined horizontal search radius were considered, and class labels were updated through majority voting. This approach, based on geometric neighbourhood relationships, helps reduce classification discrepancies, particularly along boundary regions between building roofs and adjacent vegetation.

Multi-Stage Classification

A multi-stage post-processing pipeline is applied to optimize classification accuracy. Points that retain their impervious surface label despite exhibiting high normalized height values are reclassified based on the spatial distribution of neighboring building and vegetation points. Classification noise occurring along transition zones between buildings and vegetation is suppressed using a median neighbourhood distance filtering approach. Finally, height-based morphological constraints are enforced to maintain geometric consistency and to prevent points with unrealistically low normalized heights from being assigned to the building or tree classes. This procedure generates the final semantically labelled point cloud.

Evaluation Metrics

Height-based statistics and local geometric features were used to quantitatively evaluate the geometric consistency of the classification. In this context, the mean normalized height per class, the standard deviation of height, and surface roughness were calculated. In addition, linearity, planarity, and sphericity indices derived from the eigenvalues of the local covariance matrix computed from k-nearest neighbours were used to characterize the three-dimensional morphological structure of each class. These indices serve as geometric features that capture the dominant dimensionality of the local point distribution, ranging from linear (1D) to planar (2D) and volumetric (3D) structures[14].

$$L_{\lambda} = \frac{\lambda_1 - \lambda_2}{\lambda_1}, \quad P_{\lambda} = \frac{\lambda_2 - \lambda_3}{\lambda_1}, \quad S_{\lambda} = \frac{\lambda_3}{\lambda_1} \quad (1)$$

Where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the ordered eigenvalues of the local covariance matrix, representing the three-dimensional structure tensor computed from the k-NN neighborhood of each point.

Results

Table 1 shows that the relative vertical distribution of the classes remains consistent across all tiles. The impervious and low vegetation classes are characterized by very low mean normalized heights together with low standard deviations, reflecting their ground-level nature. In contrast, the building and tree classes exhibit substantially higher mean normalized heights

and standard deviations, clearly distinguishing elevated objects from low-elevation classes. Since heights are normalized, these class-wise patterns remain comparable across tiles even where terrain conditions differ locally.

Table 1. Class-based statistical and geometric feature values obtained for different test areas (Tiles) on the ISPRS Potsdam dataset.

Tile	Class	Mean _z (m)	STD _z (m)	Roughness	Linearity	Planarity	Sphericity
02_10	Impervious	0.063	0.126	0.0229	0.2660	0.7046	0.0035
	Building	4.433	3.364	0.2662	0.3261	0.6075	0.0321
	Low Vegetation	0.084	0.157	0.0187	0.2545	0.7356	0.0011
	Tree	4.047	3.875	0.3054	0.3497	0.4976	0.0907
02_11	Impervious	0.073	0.137	0.0410	0.2688	0.6953	0.0068
	Building	4.493	4.846	0.1388	0.2944	0.6648	0.0095
	Low Vegetation	0.067	0.134	0.0220	0.2611	0.7278	0.0017
	Tree	3.886	3.258	0.4095	0.4099	0.3976	0.1404
02_12	Impervious	0.076	0.130	0.0490	0.2710	0.6877	0.0110
	Building	5.400	4.218	0.1544	0.2874	0.6694	0.0151
	Low Vegetation	0.049	0.086	0.0150	0.2531	0.7426	0.0007
	Tree	3.220	3.553	0.2657	0.3742	0.4686	0.0987
03_10	Impervious	0.096	0.168	0.0642	0.2821	0.6653	0.0148
	Building	6.581	5.353	0.2791	0.3228	0.6176	0.0259
	Low Vegetation	0.080	0.185	0.0218	0.2735	0.7110	0.0019
	Tree	4.799	4.439	0.5623	0.4142	0.3808	0.1528
03_11	Impervious	0.067	0.129	0.0371	0.2739	0.6938	0.0056
	Building	4.774	3.351	0.2498	0.3173	0.6379	0.0144
	Low Vegetation	0.061	0.124	0.0164	0.2632	0.7275	0.0011
	Tree	7.530	6.726	0.7525	0.4272	0.3732	0.1439
03_12	Impervious	0.090	0.165	0.0579	0.2785	0.6799	0.0112
	Building	5.682	4.840	0.4545	0.3466	0.5891	0.0244
	Low Veg	0.116	0.238	0.0283	0.2818	0.6967	0.0043
	Tree	5.562	4.786	0.6695	0.4185	0.3872	0.1409
04_10	Impervious	0.112	0.193	0.0708	0.2880	0.6520	0.0211
	Building	6.187	4.387	0.3993	0.3450	0.5874	0.0303
	Low Vegetation	0.235	0.480	0.0372	0.2793	0.6982	0.0039
	Tree	4.492	4.192	0.6140	0.4124	0.3870	0.1526
04_11	Impervious	0.093	0.167	0.0595	0.2806	0.6753	0.0131
	Building	6.427	4.897	0.6396	0.3713	0.5464	0.0356
	Low Vegetation	0.110	0.228	0.0286	0.2748	0.7052	0.0039
	Tree	4.533	4.031	0.5935	0.4198	0.3714	0.1582
04_12	Impervious	0.057	0.115	0.0261	0.2664	0.7122	0.0031
	Building	8.567	6.606	0.4963	0.3419	0.5969	0.0210
	Low Vegetation	0.083	0.184	0.0154	0.2585	0.7297	0.0014
	Tree	4.750	4.574	0.4049	0.3927	0.4323	0.1102
05_10	Impervious	0.074	0.142	0.0369	0.2687	0.7003	0.0056
	Building	6.122	4.692	0.3548	0.3293	0.6146	0.0209
	Low Vegetation	0.077	0.190	0.0161	0.2635	0.7240	0.0013
	Tree	4.844	4.507	0.5171	0.3995	0.4040	0.1453

The roughness metric constitutes a strong structural descriptor. Trees consistently demonstrate substantially greater roughness than impervious surfaces and low vegetation. In scene 03_11, for instance, tree roughness exceeds that of impervious surfaces and low vegetation by more than twenty-five times, highlighting its emphasized separability.

Among eigenvalue-based features, linearity shows considerable overlap across classes and therefore has a limited ability to distinguish classes. Although slightly higher values are observed for the tree class in some tiles, the overall distributions remain similar across classes. planarity and sphericity constitute the primary geometric features. planarity effectively differentiates volumetric vegetation structures from planar artificial and ground surfaces. Tree points consistently exhibit lower planarity values below 0.50, whereas impervious surfaces typically fall in the 0.65 to 0.71 range. Building surfaces also show relatively high planarity,

between 0.55 and 0.67, indicating partial overlap with other planar classes rather than a strict threshold above 0.70.

Sphericity further strengthens the separation of the Tree class. While impervious, low vegetation, and building classes remain near zero, generally below 0.03, the Tree class shows a distinct increase, approximately between 0.07 and 0.16, reflecting its three-dimensional volumetric structure. This morphological contrast enables reliable separation between tree and building classes, which may exhibit comparable height ranges but differ fundamentally in spatial structure. Therefore, the combined use of mean normalized height, height standard deviation, roughness, planarity, and sphericity provides a structurally consistent and tile-independent framework for class discrimination. Roughness and sphericity emerge as the most decisive parameters for identifying vegetation, whereas planarity most clearly characterizes planar surfaces such as impervious ground and buildings.

A visual comparison between the orthophoto images of the study area and the corresponding three-dimensional semantic classification results obtained using the proposed method is shown in Figure 2. The results demonstrate that building roofs and road surfaces are clearly identified, and that the tree canopy is accurately represented in terms of both its volumetric structure and spatial boundaries. The accurate representation of class boundaries, even in structurally complex areas, provides visual evidence of the effectiveness of the spatial contextual refinement and post-processing steps.

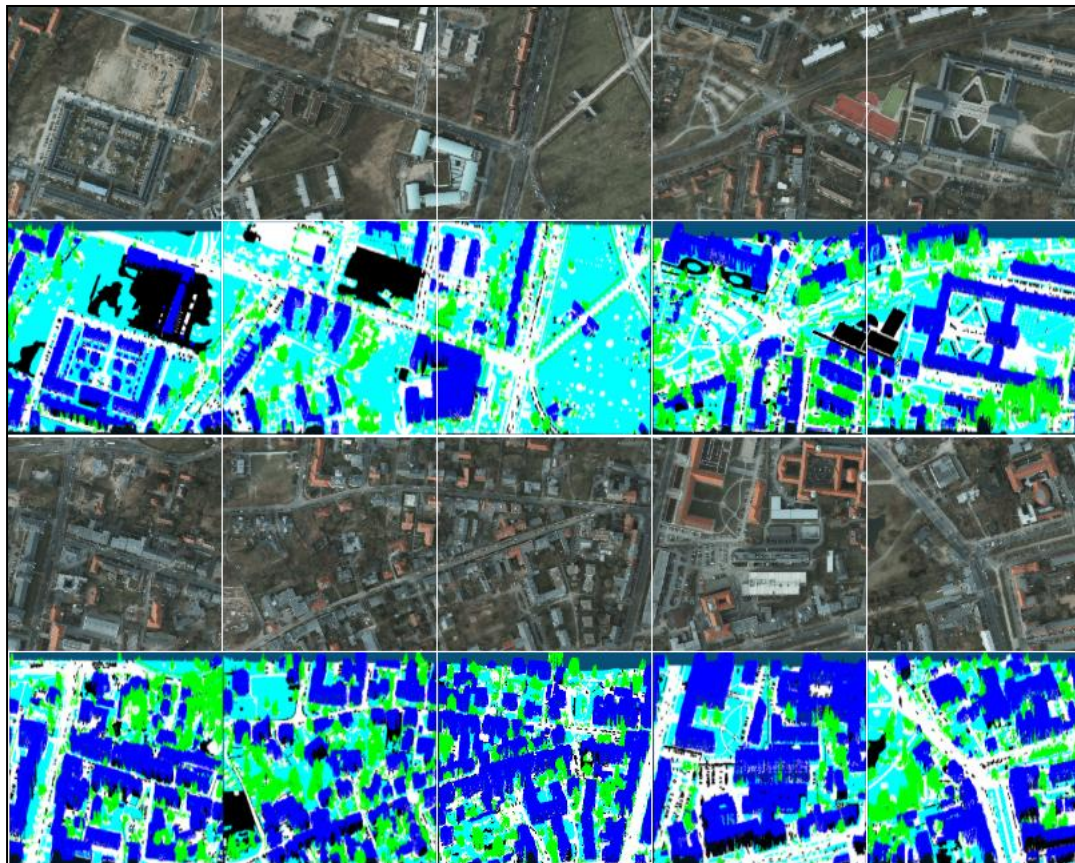


Fig. 2. 2D RGB orthophotos of the study area (top row) and the corresponding semantic point cloud classification results (bottom row). Class colors: buildings in blue, impervious surfaces in white, low vegetation in cyan, and trees in green.

Conclusion

Manually labeling LiDAR point clouds is a costly and labor-intensive process. Although AI-based methods have become increasingly prominent in this task, city-scale applications continue to face challenges associated with model complexity, dependence on large labelled datasets, and the processing of very large point clouds. The proposed method addresses this problem by integrating publicly available 2D semantic labels with geometric and morphological characteristics derived from LiDAR data. The results indicate that impervious surfaces, low vegetation, buildings, and trees can be effectively distinguished using these descriptors. Despite these strengths, the method can encounter limitations in complex areas and in mixed-class boundary regions. Beyond its role as a classification approach, the framework can also support the automatic generation of weakly supervised training data for subsequent deep learning applications. Overall, the proposed approach constitutes an interpretable and operationally applicable alternative for large area point cloud classification.

References

1. Zhang, Z., & Zhu, L. (2023). A Review on Unmanned Aerial Vehicle Remote Sensing: Platforms, Sensors, Data Processing Methods, and Applications. *Drones*, 7(6). <https://doi.org/10.3390/drones7060398>
2. Chen, G., Jin, Q., Zhang, X., & Zhang, H. (2021). Airborne Lidar Point Cloud Classification fusing Spectral Information. I 2021 3rd International Conference on Intelligent Control, Measurement and Signal Processing and Intelligent Oil Field (ICMSP) (s. 177–184). <https://doi.org/10.1109/ICMSP53480.2021.9513396>
3. Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *ArXiv*. <http://arxiv.org/abs/1612.00593>
4. Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *ArXiv*. <http://arxiv.org/abs/1706.02413>
5. Li, Y., Bu, R., Sun, M., Wu, W., Di, X., & Chen, B. (2018). PointCNN: Convolution On X - Transformed Points. *ArXiv*. <http://arxiv.org/abs/1801.07791>
6. Ma, Y., Zhang, Z., Yang, H. L., & Yang, Z. (2021). An adaptive adversarial domain adaptation approach for corn yield prediction. *Computers and Electronics in Agriculture*, 187, 106314. <https://doi.org/10.1016/j.compag.2021.106314>
7. Zhang, H., Wang, C., Yu, L., Tian, S., Ning, X., & Rodrigues, J. (2024). PointGT: A Method for Point-Cloud Classification and Segmentation Based on Local Geometric Transformation. *IEEE Transactions on Multimedia*, 26, 8052–8062. <https://doi.org/10.1109/TMM.2024.3374580>
8. Wu, J., Sun, M., Jiang, C., Liu, J., Smith, J., & Zhang, Q. (2024). Context-based local-global fusion network for 3D point cloud classification and segmentation. *Expert Systems with Applications*, 251, 124023. <https://doi.org/10.1016/j.eswa.2024.124023>
9. Zhou, W., Zhao, Y., Xiao, Y., Min, X., & Yi, J. (2024). TNPC: Transformer-based network for point cloud classification. *Expert Systems with Applications*, 239, 122438. <https://doi.org/10.1016/j.eswa.2023.122438>
10. Yang, H., Zhou, Q., Sun, H., Li, X., Liu, F., Lu, X., ... Yan, S. (2025). PointDGMamba: Domain Generalization of Point Cloud Classification via Generalized State Space Model. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(9), 9193–9201. <https://doi.org/10.1609/aaai.v39i9.32995>
11. Li, L., Shum, H. P. H., & Breckon, T. P. (2023). Less is More: Reducing Task and Model Complexity for 3D Point Cloud Semantic Segmentation. *ArXiv*. <http://arxiv.org/abs/2303.11203>

12. Hu, Q., Yang, B., Khalid, S., Xiao, W., Trigoni, N., & Markham, A. (2021). Towards Semantic Segmentation of Urban-Scale 3D Point Clouds: A Dataset, Benchmarks and Challenges. *ArXiv*. <http://arxiv.org/abs/2009.03137>
13. Franz Rottensteiner, Gunho Sohn, Markus Gerke, & Jan D. Wegner. (2014). ISPRS Semantic Labeling Contest. *International Society for Photogrammetry and Remote Sensing(ISPRS)*. <https://www.isprs.org/resources/datasets/benchmarks/UrbanSemLab/2d-sem-label-potsdam.aspx>
14. Weinmann, M., Jutzi, B., Hinz, S., & Mallet, C. (2015). Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, 286–304. <https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.01.016>