2025 European Conference on Computing in Construction

CIB W78 Conference on IT in Construction

Porto, Portugal
July 14-17, 2025

FINE-GRAINED FEATURE EXTRACTION FOR SEMANTIC SEGMENTATION OF POINT CLOUDS FOR CIVIL ENGINEERING STRUCTURES

Hristo Vassilev[1] and Jörg Blankenbach[1]
[1]Geodetic Institute and Chair for Computing in Civil Engineering & GIS, RWTH Aachen University

Abstract

Semantic point cloud segmentation is essential for creating geometric-semantic as-is models (GSMs) used in planning, construction, and digital twins. While conventional deep learning approaches often subsample the input—sacrificing fine-grained detail by focusing on large context windows—our proposed Fine-Grained Feature Extraction (FGE) method leverages local point cloud neighborhoods and full-waveform terrestrial laser scanning data to improve material classification. In our experiments, integrating color (RGB), online radiometric features (RiRiegl), raw full-waveform (FWF), and geometric features boosted the mean Intersection over Union (mIoU) from 2.1% (baseline using only point coordinates) to nearly 30%. Specifically, the sequential addition of these features yielded improvements of approximately 10.3%, 8.2%, 5.1%, and 5.9% mIoU respectively. These gains demonstrate that FGE offers a robust foundation for subsequent large-scale segmentation models that focus on long-range dependencies.

Introduction

As the volume of aging bridges in developed countries continues to grow, the need for robust and efficient as-is documentation of the built environment increased significantly (Haardt and Holst, 2008). Although laser scanning has become a common method of quickly capturing the state of existing assets with high geometrical accuracy (Tang et al., 2010), the cost and quality of the derived geometric models are still limited by the need for manual processing of the raw data. For this reason, research in the past years has focused on achieving more generalizable and efficient automation, specifically in the semantic segmentation of the captured point clouds, which is at the core of many downstream processes.

Especially data-driven methods have proven useful in cases where potentially noisy inputs have a complex and non-linear relationship with the desired outputs. Early methods focused on calibrating the measured reflected signal in order to achieve consistent measurements between campaigns and produce meaningful physical quantities such as the back-scatter coefficient (Wagner, 2010). In subsequent works Mallet (2011) supplemented calibration features with additional parameters produced by modelling the complete sequence of signals (colloquially known as full-waveform (FWF)) by mathematical models such as the generalized Gaussian. The combined feature set was evaluated for its relevance towards semantically segmenting point clouds captured with an aerial laser scanner (ALS) by utilizing a support vector machine (SVM). FWF-scanners as opposed to discrete laser scanners record the entire back-scattered signal as opposed to single intensity values, allowing for more refined subsequent analysis. The SVM approach to FWF-based segmentation was further refined by Lai et al. (2019), who showed that an ensemble of models and a wavelet kernel improved generalization.

The recent trend shows a departure from using handcrafted approaches to feature extraction and instead relies on the deployment of deep learning (DL) methods, which by observing a large corpora of data can learn to recognize reoccurring patterns and, on the basis of that, perform a subsequent segmentation of the points into the predefined semantic categories (Kellner et al., 2024). Originally successful in the field of image analysis DL methods have proven successful for processing point clouds after domain challenges such as order invariance and sparsity have been addressed (Qi et al., 2016). Shinohara et al. (2020) showed that this principle can be extended to incorporate full-waveform information, which was introduced as additional input channels along with the measured Cartesian coordinates into a PointNet-inspired DL-architecture. The method lead to increased performance in the semantic segmentation of ALS-based point clouds of road infrastructure, allowing better differentiation between road surfaces and vegetation.

FWF has seen successful use in the ALS domain as it allows a comprehensive view of the vertical profile of the return signal, enabling detailed analysis of areas with dense vegetation. At the same time, some more recent works explore for use cases of FWF data in related terrestrial applications. Pashaei et al. (2023) identified several challenges, inherent to FWF Terrestrial laser scanning (TLS), which are caused by the higher dynamic range of the applied sensors as well as their lower FWF sampling rate of 500 MHz instead of the 1 GHz, that is typical for ALS. Nevertheless, direct integration of FWF into a PointNet DL-Architecture showed a 7% increase of accuracy in the semantic segmentation of urban scenes captured by TLS.

In another study Janda et al. (2023) performed experiments in a controlled laboratory environment, which showed that FWF-data alone were enough to distinguish some materials such as Aluminium from Black Cloth with very high accuracy using a flash FWF laser scanner. This performance degraded significantly with more shallow incidence angles and objects of the same material but painted in different colors, showcasing the limitations of the FWF data. While DL methods have been showing steady gains in point cloud processing, due to hardware limitations, current algorithms often experience a trade-off between input resolution and maximum context size when extracting features at vastly different scales. Because especially in infrastructure assets the objects of interest tend to be more recognizable by spatially larger geometric patterns the best results are often observed when the resolution of the input point cloud is reduced, therefore enabling the use of larger context windows (Vassilev and Blankenbach, 2023). In this and the upcoming works, we hypothesize that a two-step process of fine-grained feature extraction followed by a large-context model can achieve gains in performance by levering otherwise discarded low-level information. To this end we develop a simple yet efficient neural network architecture which by leveraging derived as well as raw radiometric information (FWF) can process dense point clouds and extract feature-rich embeddings that are aligned with the task of material classification and can serve as inputs to various subsequent high-level segmentation tasks. Therefore this publication outlines the following contributions:

- A feature-rich dataset with multi-level annotations
- A DL architecture for embedding fine-grained material information for later use in large-scale models
- Analysis on the sensitivity of the proposed model towards different input features

Dataset

The first prerequisite for the joint training of a low-level feature extractor and a high-level segmentation model is the development of a dataset with known low-level and high-level semantics. This stems from the general goal of allowing higher flexibility in the choice of provided labels for supervising the respective model. To this end we develop a dataset consisting of 18 point clouds of which 5 contain building facades, while the rest contain bridges of various building materials (e.g. concrete, steel, natural stone). The dataset is split in a train (12 point clouds) and validation (8 point clouds), while aiming for a similar proportion of materials and semantic classes between the splits. To incorporate a multi-level labeling schema each point was annotated in four different categories as illustrated in figure 1. The first three categories contain semantic classes with increasing complexity, while the final level contains the material information of the target. The experiments presented within this work will mainly focus on material classification and will only utilize level 4, while the combination with other levels will be explored in subsequent works.

All point clouds are collected with the Riegl VZ-400 TLS and contain color information captured by an on-board camera (here refered to as RGB). The scanner is equipped with online waveform processing, which uses algorithms from Riegl to derive radiometric quantities (here refered to as RF-Riegl) for each point: 1) the raw peak amplitudes and 2) the derived reflectance values, which compensate for signal attenuation as a function of the distance to the scanner, 3) a 'deviation' value measuring the difference between the captured waveform and the expected waveform, 4) the number of total targets located on the path of a single laser beam and 5) the index of the current target within the beam. Next to the derived features we furthermore collect the raw full waveform in immediate proximity to the detected peaks (here refered to as FWF). The FWF data is a discrete time series, which captures the recorded back-scattering energy at a frequency of 500MHz. The regions containing fluctuations in the waveform may have a variable length and we therefore pad or crop them a fixed length of 32. Notably, given these considerations and assuming a 16-bit unsigned integer as the data type, the FWF data make up a significant portion, accounting for 47% of the total dataset volume. This is a major drawback, which has previously discouraged practitioners from offline processing of FWF information. Nevertheless, plotting all waveforms corresponding to two classes in our dataset (metal and mesh) we can make an initial observation that a smooth surface such as metal tends to have a distinct peak in intensity, whereas a meshed surface only partially reflects the incoming beam and therefore tends to lead to numerous secondary reflections close to a detected point (Figure 2). Therefore, it is a worthwhile investigation whether raw FWF information can contribute to material classification even in the case of terrestrial laser scanning.

Methodology

Towards the goal of enriching the feature representation of a sparse point cloud, we formulate the problem as follows: For every point $\mathbf{x}_0$ in the sparse point cloud, we gather the K_L number of nearest neighbors in the dense point cloud with their corresponding feature vectors $\mathbf{f} \in \mathbb{R}^F$ and waveforms $\mathbf{w} \in \mathbb{R}^W$, where W is the length of each FWF-sequence. Additionally, a smaller number of neighbors K_S is used to calculate the local covariance, from which, through eigen-decomposition, the normal vectors as well as the geometric properties of each point can be extracted (Kellner et al., 2024).

Normal Vectors

The covariance matrix $\mathbf{C}$ is computed for the K_S nearest neighbors of a query point $\mathbf{x}_0$ as:

$$\mathbf{C} = \frac{1}{K_S - 1} \sum_{i=1}^{K_S} (\mathbf{p}_i - \bar{\mathbf{p}})(\mathbf{p}_i - \bar{\mathbf{p}})^\top, \tag{1}$$

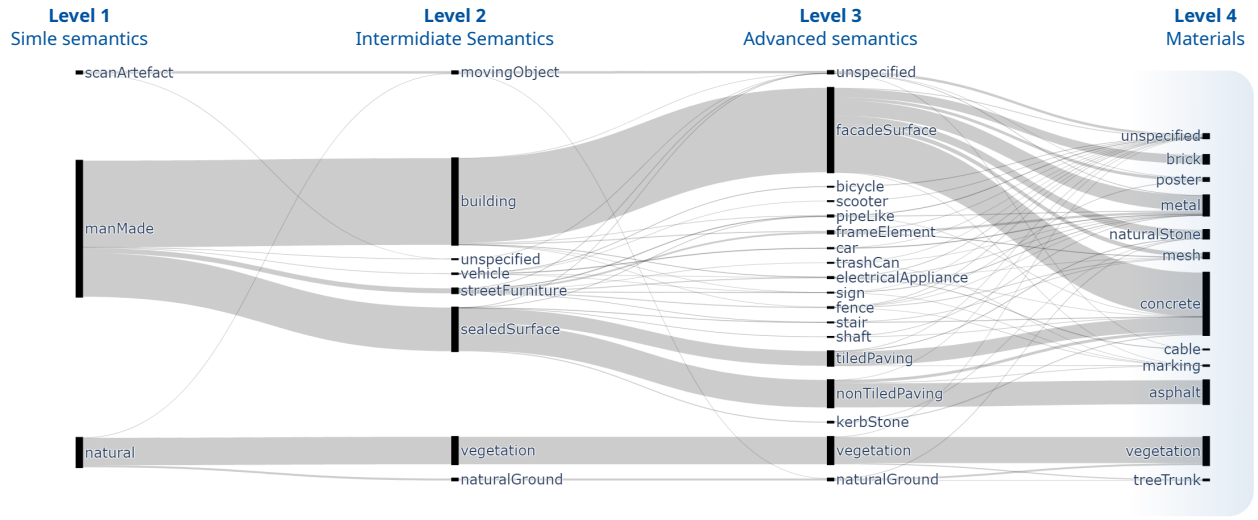

Figure 1: Datast structure. The sizes of the blocks as well as their connections visualize their relative size at each level.

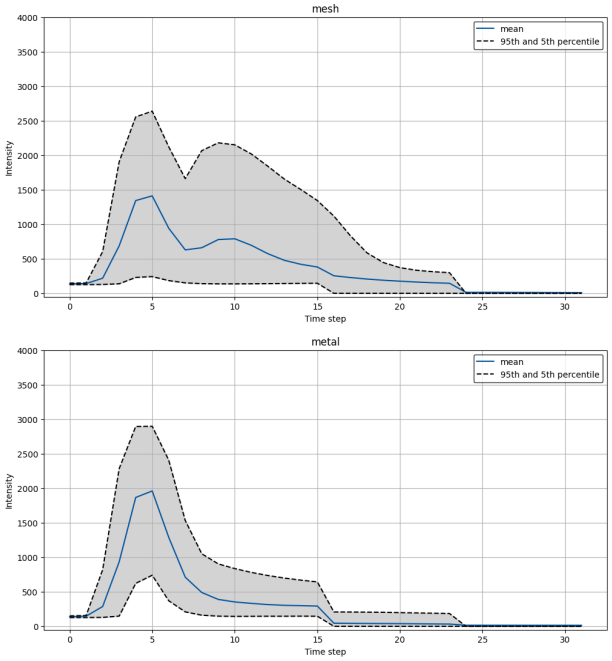

Figure 2: Illustration of the average as well as the 95th and 5th percentile waveform for meshed materials(top) and solid metal objects (bottom)

where $\mathbf{p}_i$ are the 3D coordinates of the neighboring points and $\bar{\mathbf{p}}$ is the mean of these neighbors. By performing eigen-decomposition of the covariance matrix, we obtain the eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. The normal vector $\mathbf{N}$ for the point $\mathbf{x}_0$ is given by the eigenvector $\mathbf{v}_3$ associated with the smallest eigenvalue λ_3, as it represents the direction of least variance in the neighborhood. The orientation of the normal vector $\mathbf{N}$ is flipped if needed to point toward the scanner origin, ensuring consistency.

Incidence Angles

The incidence angle α_i for a point $\mathbf{P}_i$ is calculated as:

$$\alpha_i = \cos^{-1}\left(\frac{\mathbf{P}_i \cdot \mathbf{N}}{\|\mathbf{P}_i\|\|\mathbf{N}\|}\right), \tag{2}$$

where $\mathbf{P}_i$ is the vector from the scanner to the point, and $\mathbf{N}$ is the normal vector at the point.
This formulation ensures that the incidence angle captures the relationship between the surface orientation and the direction of the incoming laser beam.

Geometric Features

The eigenvalues of the covariance matrix are further used to compute geometric features that describe the local shape properties of the neighborhood. These include:

$$\text{Linearity} = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \varepsilon} \tag{3}$$

$$\text{Planarity} = \frac{\lambda_2 - \lambda_3}{\lambda_1 + \varepsilon} \tag{4}$$

$$\text{Sphericity} = \frac{\lambda_3}{\lambda_1 + \varepsilon} \tag{5}$$

$$\text{Omnivariance} = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3} \tag{6}$$

$$\text{Anisotropy} = \frac{\lambda_1 - \lambda_3}{\lambda_1 + \varepsilon} \tag{7}$$

$$\text{Eigenentropy} = -\sum_{i=1}^{3} \frac{\lambda_i}{\sum_{j=1}^{3} \lambda_j} \log\left(\frac{\lambda_i}{\sum_{j=1}^{3} \lambda_j} + \varepsilon\right) \tag{8}$$

Here, ε is a small positive constant to prevent division by zero. These features capture the local structure of the neighborhood, distinguishing between linear, planar, and volumetric configurations.

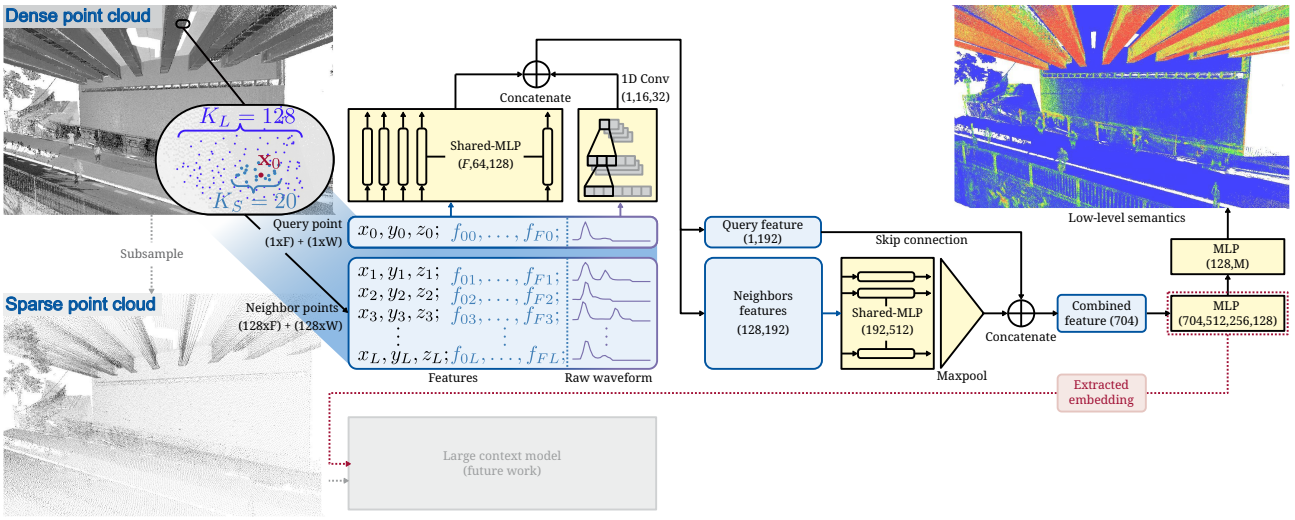

Figure 3: Architecture of the proposed model

Together, these build the feature $\mathbf{X} \in \mathbb{R}^{K_L \times (F+3)}$ and waveform matrices $\mathbf{W} \in \mathbb{R}^{K_L \times W}$. Since the points and their features are an unordered set, we use the permutation invariant combination of a shared multi-layer perceptron (MLP) and a max-pooling layer as first suggested by Qi et al. (2016). Rather than extending this mechanism to the FWF data, we deviate from prior works by leveraging the inherent structure of the time series and convolving it with 3 consecutive convolutional layers, each with a 1D-kernel size of 3, followed by max pooling with stride 2. The resulting FWF features are concatenated with the extracted point-wise features (marked by $\oplus$ in figure 3) and fed to further shared-MLP layers. Afterwards, a global feature is derived by taking the maximum value in each column of the neighborhood feature matrix, resulting in a single 704-dimensional vector describing the entire vector. To enhance gradient flow and preserve local information, skip connections are employed between the processed query point features and the aggregated neighbor features, ensuring that the relevance of the query point is not lost during the embedding process. A final MLP-block then reduces the dimensionality of the aggregated feature, yielding a 128-dimensional embedding that is cached for subsequent processing in downstream models (see Figure 3). The FGE is well-suited for such tasks as it only captures local information and is applicable at high resolution (e.g. 1cm). Evaluating these embeddings in larger contextual settings is a topic for further feature research.

Although the final goal is to extract features in the subsampled resolution, during training we instead sample random points from the full-resolution cloud. Rather than using a uniform probability distribution we sample each point with a probability, that is inversely proportional to its frequency in the dataset in order to negate the effects of class imbalance. This ensures that under-represented classes are given equal importance during training, mitigating biases that may arise from imbalanced class distributions. In specific, this is achieved by deriving the class-weights as:

$$w_k = \frac{\frac{1}{n_k+1}}{\sum_{j=1}^{N} \frac{1}{n_j+1}} \tag{9}$$

where:

- n_k is the frequency of class k,
- N is the total number of classes,

During the random subsampling of the dataset, the probability of selecting a point p_i is derived as the average of the weights of all the class labels associated with that point. For a point p_i with class labels $l_1, l_2, \ldots, l_m$, the sampling probability $P(p_i)$ is calculated as:

$$P(p_i) = \frac{1}{m} \sum_{k=1}^{m} w_{l_k} \tag{10}$$

where:

- m is the number of label levels associated with the point p_i,
- w_{l_k} is the weight of the k-th class label.

To further increase the effective size of the training data and to stimulate an invariance towards irrelevant effects we perform established data augmentation methods during training such as gaussian perturbation of the features and coordinates of the inputs and random rotations about the z-axis. To reduce any effects of overfitting we introduce a dropout with a 10% probability on the input features and a 30% on all shared-MLP layers in the network. Batch normalization is applied after each convolutional layer to stabilize training and improve convergence.

Experiments

Classifying materials in point clouds requires a different approach than identifying large-scale structures. While geometric patterns suffice for large objects, materials often

share geometry, making small textural details and radiometric information critical. To quantify feature relevance, we trained the our proposed fine grarined feature extractor (FGE) model with all features and used average integrated gradients for sensitivity analysis (Figure 4). While using the raw gradients is a common method for sensitivity analysis in linear models, they can miss important relationships in more complex models such as artificial neural networks. Instead, approximating the integral of the gradients along straight a path between a zero-vector and the input feature vector has been axiomatically proven to provide a more robust measure of the true contribution of each feature (Sundararajan et al., 2017).

While the sensitivity analysis highlights feature importance in a complete set it still lacks insights into the performance under a constrained feature set. For this reason we furthermore perform various combinations of feature sets and present them as an ablation study in tables 1 and 2.Therefore, the experiments aim to dissect the relative importance and potential redundancy among the input components. To ensure statistical significance and eliminate any variance introduced through the stochastic training process we repeat each experiment a total of 4 times and report the mean and standard deviation of the calculated metrics.

Results

Sensitivity analysis

The result of the sensitivity analysis shows that on average the model chose to focus primarily on laser-derived radiometric features such as the raw waveform and reflectance values being closely followed by the different color channels captured by the RGB camera (figure 4). With the exception of eigenentropy most geometric features showed to be on average less relevant. Notably, the normalized intensity value (Reflectance) has a much larger significance than the raw intensity value (Amplitude).

Ablation study

Ablation experiments (Tables 1, 2) assess contributions of features and design choices. Exp. #0 serves as a baseline by using only the Cartesian (XYZ) coordinates, while exp. #1 adds color information (RGB) to quantify its effect on classification accuracy. Subsequent experiments progressively incorporate derived radiometric features (RF-Riegl), waveform data (FWF), and local geometric properties such as normal vectors and incidence angles to assess their impact. Additionally, exp #7 investigates the necessity of using neighboring points by training the model without neighborhood information, and exp #8 explores the role of radiometric data by jointly excluding RF-Riegl and FWF features.

Results from these experiments reveal several key findings. The two largest contributions are attributed to the sequential addition of color information and RF-Riegl features by accounting for a change in mIoU of approximately +10.25% and +8.22% respectively. The addition of geometric features led to an increase of 5.91%, while the 1D convolution mechanism of processing the raw FWF contributed +5.1% mIoU. Furthermore, introducing the scan-target relationship information of the incidence angles and the distance to the laser scanner lead to additional gains of +3.1%. Notably, the initial gains observed from including FWF processing in #5 could not be validated if by removing FWF from the complete feature set in #9.

Neighborhood information proved to be critical, as evidenced by exp. #7, where removing neighboring points caused performance to drop to baseline levels (0.55% mIoU). This indicates that the local context provided by neighboring points is essential for effective feature extraction. Finally, exp. #8 validated the importance of radiometric data (RF-Riegl and FWF) by showing a significant drop of 14.38% mIoU when both were excluded, leaving only geometric, RGB and positional information. Overall the findings demonstrate the complementary roles of color, geometry, radiometric, and target-scanner features in material classification based on small-scaled neighborhoods.

Training data volume

The impact of training data volume on model performance was evaluated by gradually increasing the number of point clouds used for training while keeping the validation dataset fixed. As illustrated in Figure 5, the validation mIoU demonstrates a consistent upward trend as more training data is introduced. However, the progression is not strictly linear, with noticeable fluctuations at intermediate stages, likely caused by the variability in the scenes and materials included in the dataset. For smaller training volumes (1-5 point clouds), the performance remains relatively low and stable, indicating insufficient representation of diverse material and geometric configurations. A sharp improvement is observed as the training data size increases beyond six point clouds, suggesting that the model benefits from encountering a broader variety of examples that capture the high variability of real-world materials and structures. Despite the observed improvement, the saturation point has not yet been reached, as the mIoU continues to rise even at the largest dataset size (12 point clouds).

Qualitative results

Results from the FGE model are visualized in figures 6 to 12 on a chosen hard example. The individual output probabilities show potential use as embeddings for a more complex model but underscore the difficulty in differentiating very similar materials such as asphalt and concrete.

Discussion and Result Analysis

The experiments revealed that incorporating a rich feature set enables a simple local MLP network to achieve moderate results in material classification, boosting its performance to nearly 30% mIoU. While the inclusion of raw FWF data benefits material classification—contributing a 5.1% increase in mIoU—it is not essential, as its role can be partly replaced by other available features (e.g., geometric or online radiometric features). Nevertheless, a

Table 1: Overview of the ablation study. XYZ: Point coordinates, RGB: color channels, RF-Riegl: online radiometric features, Geom.: geometric features, Norm: surface normals, I.Ang.: angle of incidence, Dist.: distance from scanner to target, FWF: offline full-waveform processing, Neighb.: whether neighbors are processed by the network

Exp.	XYZ	RGB	RF-Riegl	Geom.	Norm.	I.Ang.	Dist.	FWF	Neighb.	mIoU		mAcc	
#0	✓								✓	2.12	±2.49	6.87	±3.86
#1	✓	✓							✓	12.37	±0.75	22.66	±0.81
#2	✓	✓	✓						✓	20.59	±0.13	33.07	±1.22
#3	✓	✓	✓			✓	✓	✓	✓	23.69	±0.46	35.96	±1.83
#4	✓	✓				✓	✓	✓	✓	20.22	±0.77	32.26	±0.62
#5	✓	✓						✓	✓	17.47	±1.12	30.17	±1.61
#6	✓	✓	✓	✓	✓	✓	✓	✓	✓	29.60	±0.76	41.52	±1.27
#7	✓	✓	✓	✓	✓	✓	✓	✓		0.55	±0.33	4.80	±0.65
#8	✓	✓		✓	✓				✓	15.57	±0.74	25.19	±1.06
#9	✓	✓	✓	✓	✓	✓	✓		✓	**29.95**	±1.05	**42.90**	±2.90
#10	✓	✓	✓	✓	✓				✓	27.06	±0.33	40.70	±1.19

Table 2: Class-wise IoU for the experiments defined in table 1. Unsp.: Unspecified, Asph.: Asphalt, Conc.: Concrete, Mark.: Road markings, N.Stn.: Natural stone, Post.: Posters and billboards, Veg.: Vegetation

Exp.	Unsp.	Asph.	Brick	Cable	Conc.	Mark.	Mesh	Metal	N.Stn.	Post.	Tree	Veg.
#0	0.00	10.63	0.00	0.04	5.65	0.00	0.00	0.00	0.00	2.21	0.15	6.77
#1	1.39	39.99	0.00	1.52	33.72	2.22	6.08	0.69	0.00	10.24	2.94	49.60
#2	7.92	22.73	0.00	4.37	43.61	2.52	62.67	9.51	0.00	24.80	6.75	62.20
#3	4.84	34.98	0.01	3.10	45.94	15.30	57.83	35.68	0.00	23.90	6.22	56.52
#4	7.55	44.82	**0.08**	6.70	42.72	8.40	17.73	34.32	0.00	20.19	5.15	54.98
#5	6.57	34.47	0.01	3.53	40.41	4.42	25.69	19.30	0.00	19.36	5.07	50.78
#6	9.11	48.13	0.06	**25.91**	**48.11**	18.19	62.96	**40.62**	0.00	23.57	9.97	68.62
#7	0.00	0.93	0.00	0.00	3.89	0.00	0.00	0.00	0.00	1.74	0.00	0.02
#8	4.94	**51.30**	0.00	5.16	32.88	2.12	1.86	3.00	0.00	17.19	2.39	65.97
#9	**9.31**	44.48	0.00	24.57	47.58	18.28	61.81	42.52	0.00	**30.36**	**12.53**	67.93
#10	8.26	43.72	0.00	13.39	45.30	**23.10**	**64.59**	25.29	0.00	21.58	10.85	**68.67**

model with a complete feature set still relied heavily on the raw waveform as indicated by high its integrated gradient. This suggests that the FWF-based embeddings may potentially capture more intricate features, that are not necessary reflected by the chosen performance metrics (accuracy and mIoU) as they only reflect a discrete overlap with the ground truth. While the proposed dataset of 18 point clouds offers substantial volume and annotation detail varying the trained volume indicates that further gains in performance could be achieved by incorporating additional training samples, particularly those capturing underrepresented materials or complex geometric structures.

Conclusion

This work presented a flexible and efficient model for extracting fine-grained features from small and dense neighborhoods in 3D point clouds. By combining and compressing geometric and radiometric information, the proposed approach demonstrates promising results as a feature embedding method for subsequent use in larger deep learning models which tend to perform better when operating on moderate point cloud density and larger context windows. The results underline the importance of incorporating complementary information sources, with the greatest contribution given by laser-derived radiometric features and color data, followed by derived geometric features and scanner-target relationships such as angle of incidence. Despite these advances, several areas for improvement remain. Firstly, the performance of the FWF component can be optimized by designing an extractor tailored to the signal's characteristics. Previous research suggests that processing only a small subset of samples (e.g., 9 samples around the peak) could yield similar benefits while significantly reducing computational overhead. Future work should explore whether a more selective approach could better leverage the unique properties of FWF data. Secondly, the current model treats all neighbors equally in aggregation. Future improvements could include attention mechanisms or weighting schemes to adjust neighbor importance and enhance performance. An-

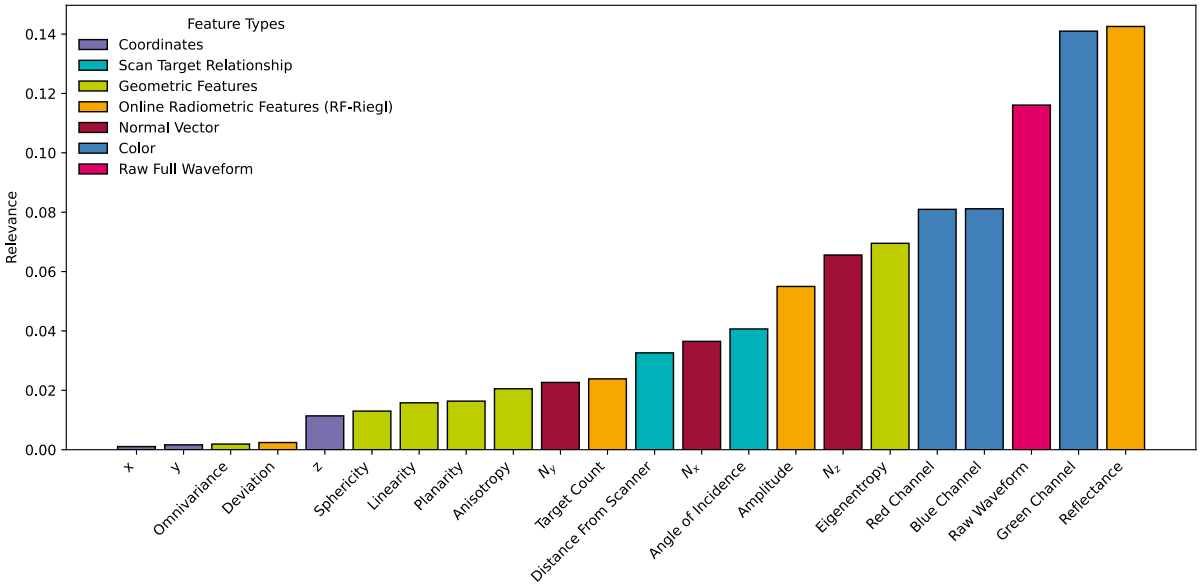

Figure 4: Feature relevance computed as the average absolute integrated gradient of the predicted class' score with respect to each input feature.

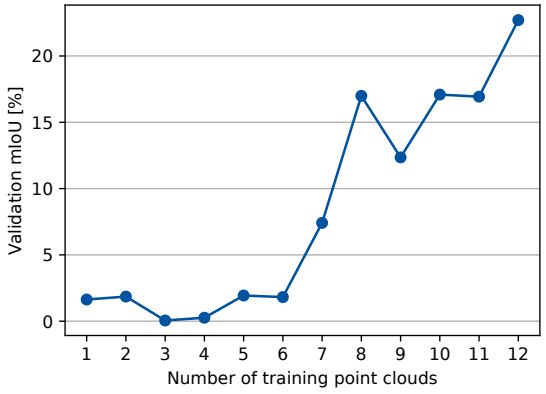

Figure 5: Performance as a result of the training data volume given a fixed validation dataset.

Figure 6: Color (RGB)

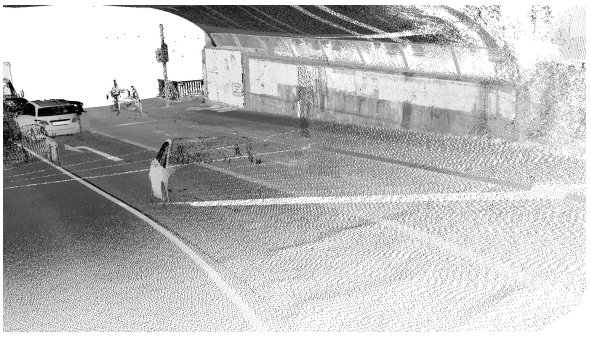

Figure 7: Intensity (Reflectance)

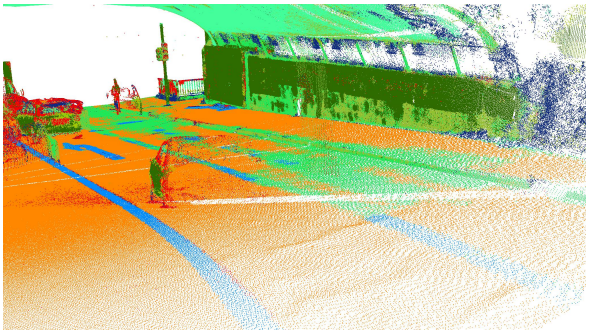

Figure 8: Predictions

other key challenge is handling class imbalance. Lastly, the effect of the number and density of neighbors (K_L and K_S) on model performance remains an open question. Future research could focus on systematically evaluating the impact of these parameters, with the goal of balancing computational efficiency and model accuracy. Additionally, experiments on larger and more diverse datasets could help identify edge cases and further validate the model's generalizability. The overarching goal of this study was not to create a perfect segmentation model but to design a robust and versatile embedding mechanism that could serve as a building block for hierarchical deep learning workflows. While the mIoU values remain modest, they are a promising foundation for future improvements. By

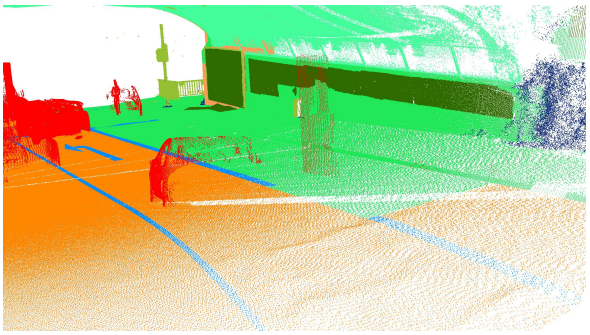

Figure 9: Ground truth

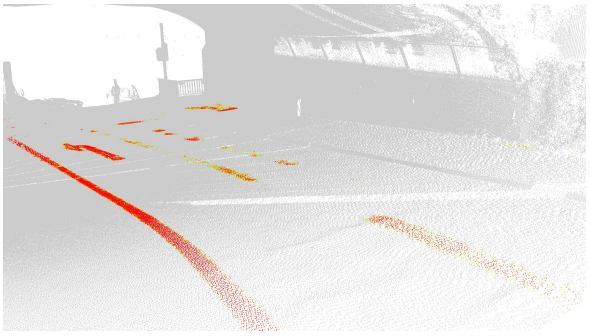

Figure 10: $p(\mathbf{x}_0 = marking)$

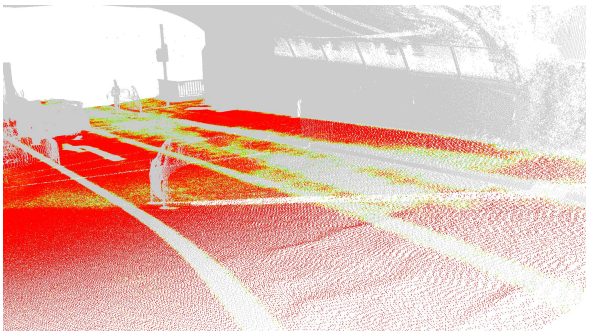

Figure 11: $p(\mathbf{x}_0 = asphalt)$

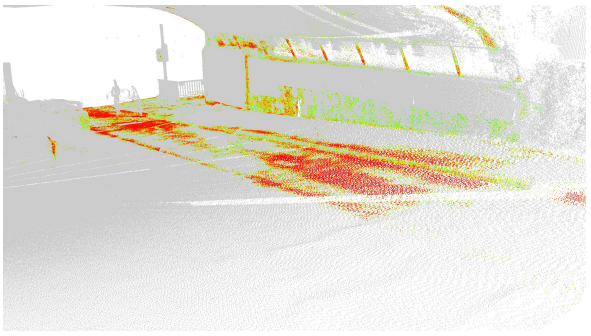

Figure 12: $p(\mathbf{x}_0 = concrete)$

iteratively addressing the identified context window limitations, the proposed approach has the potential to advance the state of point cloud semantic segmentation.

Acknowledgements

This research has received funding by (DFG, German Research Foundation) within the framework of "SPP 2388: Hundert plus - Verlängerung der Lebensdauer komplexer Baustrukturen durch intelligente Digitalisierung" with project number 501834640.

References

Haardt, P. and Holst, R. (2008). The German approach to bridge management: From reactive to predictive management procedures. In Tenth International Conference on Bridge and Structure Management, pages 3–15.

Janda, A., Merriaux, P., Olivier, P., and Kelly, J. (2023). Living in a material world: Learning material properties from full-waveform flash lidar data for semantic segmentation.

Kellner, M., Vassilev, H., Busch, A., Blaskow, R., Ferrandon Cervantes, M., Poku-Agyemang, K. N., Schmitt, A., Weisbrich, S., Maas, H.-G., Neitzel, F., Reiterer, A., and Blankenbach, J. (2024). Scan2bim – a review on the automated creation of semantic-aware geometric as-is models of bridges. avn – allgemeine vermessungsnachrichten, (3):159–181.

Lai, X., Yuan, Y., Li, Y., and Wang, M. (2019). Full-waveform LiDAR point clouds classification based on wavelet support vector machine and ensemble learning. Sensors, 19.

Mallet, C. (2011). Relevance assessment of full-waveform lidar data for urban area classification. ISPRS Journal of Photogrammetry and Remote Sensing, 66.

Pashaei, M., Starek, M. J., Glennie, C. L., and Berryhill, J. (2023). Classification of terrestrial lidar data directly from digitized echo waveforms. IEEE Trans. Geosci. Remote Sensing, 61:1–12.

Qi, C. R., Su, H., Mo, K., and Guibas, L. J. (2016). PointNet: Deep learning on point sets for 3d classification and segmentation.

Shinohara, T., Xiu, H., and Matsuoka, M. (2020). FWNet: Semantic segmentation for full-waveform LiDAR data using deep learning. Sensors, 20.

Sundararajan, M., Taly, A., and Yan, Q. (2017). Axiomatic attribution for deep networks.

Tang, P., Huber, D., Akinci, B., Lipman, R., and Lytle, A. (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. 19(7):829–843.

Vassilev, H. and Blankenbach, J. (2023). 22. internationale geodätische woche obergurgl 2023. In Weinold, T., editor, 22. Internationale Geodätische Woche Obergurgl 2023, pages 90–101, Berlin; Offenbach. Wichmann. Buchbeitrag, Beitrag zu einem Tagungsband.

Wagner, W. (2010). Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: Basic physical concepts. ISPRS Journal of Photogrammetry and Remote Sensing, 65.