Semantic segmentation of point clouds with the 3D medial axis transform

Giulia Ceccarelli[1], Weixiao Gao[1],*, Ravi Peters[2]

[1]Faculty of Architecture and the Built Environment, Delft University of Technology, The Netherlands -
giulia.ceccarelli94@gmail.com, w.gao-1@tudelft.nl
[2]3DGI, Zoetermeer, The Netherlands - ravi.peters@3dgi.nl

Keywords: 3D point cloud, Semantic Segmentation, Medial Axis Transform, Deep Learning, Geometric Features

Abstract

Semantic segmentation of 3D point clouds is pivotal for urban modeling and autonomous systems, yet challenges like irregular data structure and complex geometry hinder accurate segmentation. This study explores integrating the 3D Medial Axis Transform (MAT)—a topological skeleton encoding shape geometry via maximally inscribed balls—into deep learning frameworks to enhance semantic reasoning. We propose a feature fusion approach embedding MAT-derived attributes (radii, separation angles, medial bisectors) into point-based (PointNet++) and graph-based (Superpoint Graph) networks, enabling explicit geometric context for local points and superpoint relationships. Experiments on diverse datasets (3DOM, SynthCity, SHREC) demonstrate that MAT-enhanced features, particularly radii and separation angles, improve mean intersection over union (mIoU) by 5.8–12.4% compared to baseline RGB-only models, especially for classes like grass and shrubs where appearance features are ambiguous. However, MAT-guided geometric partitioning requires careful regularization to avoid over-segmentation, and graph convolutions benefit most from mean MAT attributes for global structure modeling. This work establishes MAT as a valuable geometric prior for point cloud segmentation, highlighting its potential to bridge topological structure and data-driven learning.

1. Introduction

Semantic segmentation of 3D point clouds, which aims to assign each point to a predefined class, has emerged as a critical task in various fields, including urban planning, autonomous driving, and cultural heritage preservation (Qi et al., 2016; Landrieu and Simonovsky, 2017). As point clouds are unordered, irregular, and invariant under geometric transformations, deep learning methods for this task must address challenges such as permutation invariance, local structure capture, and scale adaptivity. Early approaches like PointNet (Qi et al., 2016) leverage on symmetric functions (e.g., max-pooling) to handle unordered points, while hierarchical architectures like PointNet++ (Qi et al., 2017) recursively aggregate features from nested local neighborhoods. Graph-based methods, such as Superpoint Graph (Landrieu and Simonovsky, 2017), partition point clouds into semantically homogeneous superpoints and model adjacency relationships via graph convolutions. Despite these advancements, they often overlook explicit structural representations that could enhance semantic reasoning, particularly in complex urban environments with occlusions and density variations.

The 3D Medial Axis Transform (MAT), a skeleton representation encoding a shape's topological and geometric properties via maximally inscribed balls, has shown promise in shape analysis and organization (Peters, 2018). Each medial atom in the MAT captures local geometry through properties like radius and separation angle, offering a compact descriptor for spatial context. Despite its proven value as a shape descriptor in computational geometry and applications in urban mesh analysis (Gao et al., 2021, 2025) as well as structural modeling of airborne LiDAR point clouds (Peters, 2018), its role in point cloud semantic segmentation—particularly for mobile-based point clouds with varying densities and occlusions—remains

unclear. Specifically, it is unclear whether MAT's inherent structural information (e.g., medial radii) can effectively regularize learning in irregular point distributions or complement appearance features like RGB and intensity.

This work presents the first systematic study integrating 3D MAT with deep learning for point cloud semantic segmentation. Our key innovations include: (1) A feature fusion framework embedding MAT-derived attributes (radii, separation angles, medial bisectors) into point-based and graph-based networks; (2) Quantitative analysis of MAT's complementary benefits across different sensor modalities (MLS, dense image matching); (3) Empirical validation showing MAT-enhanced models achieve 5.8-12.4% IoU improvements on real-world mobile LiDAR datasets. The methodology preserves MAT's theoretical advantages while maintaining computational efficiency through parameterized feature extraction, demonstrating generalizability across synthetic and real-world urban scenes.

2. Related work

2.1 Deep Learning for Point Cloud Understanding

Deep learning in point clouds faces challenges due to their unordered, irregular, and transformation-invariant nature. Point clouds represent complex urban scenes with issues like occlusion and irregular point distribution, making semantic segmentation—classifying each point into K classes—critical yet demanding. Early methods like PointNet (Qi et al., 2016) use symmetric functions (e.g., max-pooling) to handle unordered points, while hierarchical architectures such as PointNet++ (Qi et al., 2017) recursively aggregate features from local neighborhoods. Graph-based approaches like Superpoint Graph (Landrieu and Simonovsky, 2017) partition point clouds into semantically homogeneous superpoints and model adjacency via graph convolutions. Projection networks and point convolution networks Thomas et al. (2019) also contribute to

---

* Corresponding author

the field. Despite these advancements, exploiting geometric and topological descriptors to enhance segmentation accuracy remains underexplored. This study aims to address this by integrating the 3D MAT into existing frameworks.

2.2 3D MAT and Its Applications

The 3D MAT is a skeletal representation that is dual to an object's boundary and encodes both topological and geometric properties. It is defined via maximally inscribed medial balls, on both sides of the object boundary, as illustrated in Fig. 1(b). The skeletal structure of the MAT is defined by the centers of
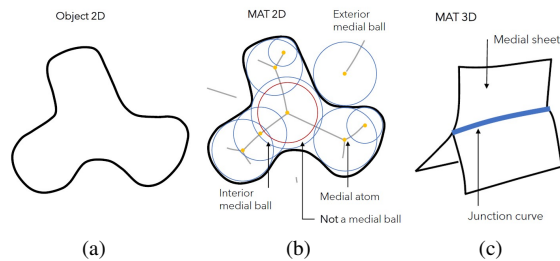
Figure 1. Visualizations of MAT in 2D and 3D. (a) depicts the 2D boundary of an object, (b) illustrates the 2D MAT, and (c) shows the medial sheets of the 3D MAT.

the medial balls. For 2D objects, this skeleton consists of points and curves, while for 3D objects, it can also contain sheets. Each medial atom, defined by the center of a medial ball and the points where the medial ball touches the object boundary, captures local geometry (e.g., thickness, curvature) through properties like the medial radii, medial bisectors, and separation angles (see also Fig. 4). The MAT of a point cloud can be approximated using the improved ball-shrinking algorithm from Peters (2018). This yields a point cloud approximation called the unstructured MAT, consisting of medial atoms without explicit skeletal topology. The structured MAT segments the atoms into medial clusters or sheets and describes the topology of the sheets through an adjacency graph. The MAT has been applied in shape analysis and modeling (Broersen et al., 2017; Gao et al., 2021, 2025), but its role in point cloud semantic segmentation—especially for mobile-based point clouds—remains unclear. This research explores integrating MAT properties into deep learning frameworks to improve segmentation accuracy, addressing the gap in leveraging topological descriptors for semantic understanding.

3. Methodology

To address the limitations of existing methods in handling complex urban point clouds with occlusions and density variations, this study proposes a three-pronged approach integrating Medial Axis Transform (MAT) geometric priors with deep learning architectures. The framework systematically combines MAT-derived features at multiple representation levels through: 1) Point-wise feature augmentation in PointNet++, 2) MAT-guided geometric partitioning in Superpoint Graph, and 3) MAT-attribute enhanced graph convolutions. This methodology bridges the gap between explicit structural modeling and data-driven learning through parametric MAT feature extraction and adaptive fusion mechanisms.

3.1 Pre-processing

The preprocessing stage focuses on preparing point clouds to enable accurate and efficient 3D MAT computation, which is critical for subsequent feature integration. Key steps include normal vector computation and orientation, subsampling and outlier removal, and data partitioning.
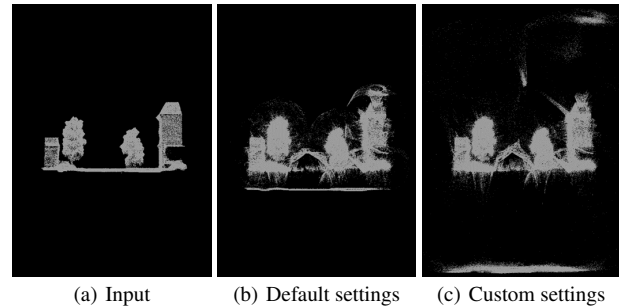
Figure 2. MAT customized preprocessing

- Normal vector's computation and orientation is a crucial pre-processing step as it enables the accurate construction of the 3D MAT and the correct separation between interior and exterior sheets. This step was carried out in CloudCompare (CC, 2016), using the program's default parameters: plane local surface model, automatic radius of the neighbors, and orientation with minimum spanning tree (knn=6). Then, for each dataset, the computed normal vectors were oriented heuristically to obtain the best possible fit.

- Subsampling is the process of reducing the number of points in a point cloud. This step was carried out in CloudCompare (CC, 2016) with the purpose of reducing the weight of the datasets while preserving fine structures. The spatial subsampling method was used; this involves selecting the minimum distance between points in the output point cloud.

- Outliers removal is a way to clean the input dataset by eliminating those elements of the point cloud that do not respect a given condition. An example are points that lie at abnormal distances from the others.

- Cropping the input point cloud into smaller input files was performed to enable a better comparison between the training and test sets and the validation set. This step is important as the files used in the deep learning algorithm and those used for evaluation should have similar characteristics, as the number of points or represented objects.

- Classes' imbalance weighting consists of assigning a weight to each class of a data-set based on the number of times it is present in the data-set. This can be done through a bin count of the points in each class, followed by an inverse weighting method. This makes sure that classes that appear less are not penalized in the deep learning phase.

- Class reduction was needed for strongly imbalanced data-sets. This phase consisted of the identification of the classes most present in the dataset, together with the classes that could be merged and those that could be eliminated.

These steps facilitate robust MAT computation by ensuring point clouds have homogeneous density, minimal noise, and correctly oriented normals, which are essential for defining medial atoms (interior/exterior radii, separation angles) and structured medial sheets. Custom parameters for MAT computation (e.g., the initial radius is 40 m, and the denoising planar and preservation thresholds are 0.6 and 0.5, respectively) were optimized to avoid spurious connections between distant objects, as validated visually using Geoflow Peters et al. (2022) (see Fig. 2).

3.2 Point-wise feature augmentation

To integrate MAT into the point-based deep learning framework PointNet++ (Qi et al., 2017), we leverage its hierarchical structure to incorporate MAT attributes as extra features. First, the algorithm's set abstraction layers—sampling, grouping, and feature extraction—were analyzed to identify optimal points for feature enrichment. MAT properties (interior/exterior coordinates, radii, separation angles) were mapped to the input points and appended to default point features (XYZ, RGB) to encode long-range geometric relationships (see Fig. 3). Three core experiments were conducted:

- Coordinate integration: interior and exterior MAT coordinates were added to represent spatial relationships between surface points and their medial atoms.

- Radius/Separation angle integration: numeric attributes (radii, separation angles) were normalized and combined with default features to describe surface curvature and local geometry.

- Hybrid features: a combination of coordinates, radii, and angles was tested to evaluate cumulative effects on segmentation accuracy.

The Pytorch implementation of PointNet++ was modified to include these features, with hyperparameters (batch size ranges from 1 to 32) tuned to balance computational efficiency and model performance. By treating MAT features as additional dimensions in the input tensor, the algorithm learned enhanced geometric representations.
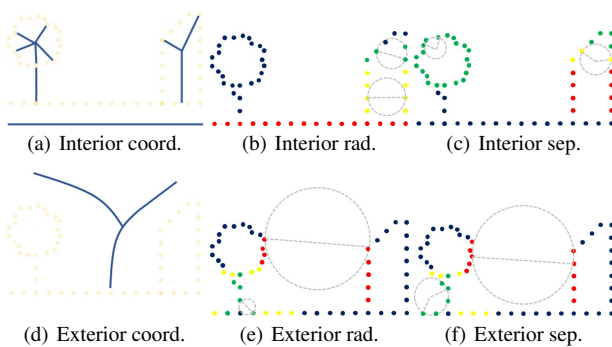

Figure 3. 3D medial axis transform (MAT) as a point feature. Here, 'coord.' represents coordinates (a 3D feature vector), "rad." denotes radius (a scalar feature value), and 'sep.' stands for separation angle (also a scalar feature value). Colors in (b), (c), (e), (f) indicate scalar feature values from low to high, and the interior and exterior shrinking balls are represented by gray dashed lines.

3.3 MAT-guided geometric partitioning

For the graph-based algorithm Superpoint Graph (SPG) (Landrieu and Simonovsky, 2017), the 3D MAT was used to refine the cut-pursuit (Landrieu and Obozinski, 2017) partitioning algorithm, which decomposes point clouds into homogeneous superpoints by minimizing a cost function balancing geometric feature consistency and adjacency weights:

$$\arg \min_{g \in \mathbb{R}^{dg}} \sum_{i \in C} \|g_i - f_i\|^2 + \mu \sum_{(i,j) \in E_{nn}} w_{i,j}[g_i - g_j \neq 0]$$

Here, f_i denotes geometric features (linearity, planarity, etc.), $w_{i,j}$ are edge weights inversely proportional to point distances, and μ controls partition coarseness. To enhance this process, MAT properties (interior/exterior radii, separation angles, medial bisectors) were integrated into the geometric feature set (linearity, planarity, scattering, verticality) to better capture structural relationships (see Fig. 4).
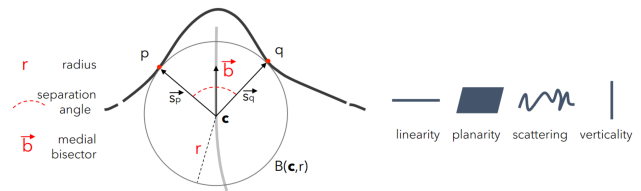

Figure 4. Comparison of MAT properties (left) and eigenvalue-based geometric descriptors (right).

Key modifications included:

- Feature enrichment: radii and angles were normalized and appended to default features, providing quantitative measures of surface curvature and medial sheet orientation.

- Edge weight adjustment: points belonging to the same medial sheet (derived from structured MAT segmentation) had their adjacency weights strengthened, promoting coherent superpoint formation.

- Segmentation refinement: medial bisectors (unit vectors along medial sheets) were used to enforce geometric consistency during partitioning, aligning superpoints with underlying MAT structures.

These adjustments aimed to make the partition more consistent with the hierarchical organization of the MAT, thereby improving the semantic coherence of superpoints of the algorithm.

3.4 MAT-enhanced graph convolutions

In the final integration step, MAT properties were used to enhance node and edge features in the Superpoint Graph, enabling more informative graph convolutions. For each superpoint (node), mean/max/min values of interior/exterior radii and separation angles were computed to characterize its geometric structure. For superedges (edges between nodes), differences in these values between adjacent superpoints were used to encode relational geometry (see Fig. 5). The workflow involved:

- Node feature embeddings: radii and angles were aggregated (mean, max, min) for each superpoint, supplementing default attributes (centroid, volume, point count).

- Edge feature embeddings: delta values of MAT properties between connected superpoints were added to edge features, enhancing the graph's ability to model inter-object relationships.

This approach capitalized on the MAT's hierarchical structure to improve both local (node) and global (edge) feature representations, particularly for complex urban scenes with diverse geometric patterns.
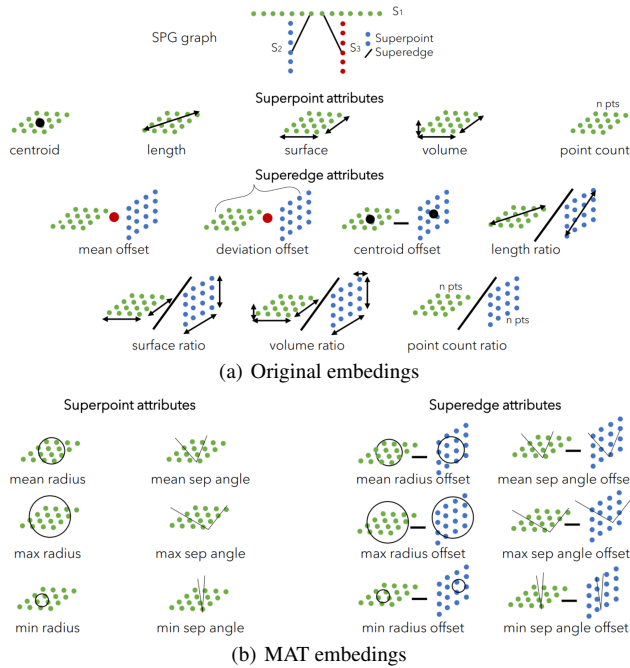

(a) Original embedings

(b) MAT embedings

Figure 5. Feature embeddings for superpoint graph

4. Evaluation

4.1 Datasets

This research utilizes three primary datasets to validate the proposed methodology, including one internal and two open-source datasets, summarized in Table 1.

	Platforms	Tiles	Points	Classes
SHREC	MLS	80	3M	5
3DOM	DIM	1	28M	6
SynthCity	MLS	9	15M to 50M	9

Table 1. Datasets specification

The SHREC dataset, a subset of CycloMedia's internal dataset, was utilized in the 3D Object Retrieval 2020 competition[1]. Comprising 80 point clouds, it is divided into a training set of 60 point clouds and a test set of 20 point clouds. These point clouds are manually labeled into five semantic classes (building, car, ground, pole, vegetation).

The 3DOM dataset, generated through dense image matching over a controlled urban artifact (Özdemir et al., 2019), features a reference point cloud of 28 million points, with training/test

---

[1] https://workshop.cgv.tugraz.at/3dor2020/

subsets (2 million points each) labeled into six classes (ground, grass, shrub, tree, facade, roof). Notable for its homogeneous point density and complete object representation, it was used to test MAT integration under ideal geometric conditions, with a manually labeled validation set derived from the full cloud.

SynthCity, a synthetic MLS dataset simulating a Velodyne scanner, includes nine areas (8 training, 1 test) with 15–52 million points total, segmented into nine semantic classes (building, car, ground, pole, road, street furniture, tree, pavement). Its stable point density and complete object geometry, regardless of material, make it suitable for evaluating MAT completeness. The dataset was subsampled to 0.01m point spacing in Cloud-Compare, reducing size by 80% while preserving fine structural details.

4.2 Results of point-wise feature

Table 2 presents the point-wise feature performance comparison in mean intersection over union (mIoU) across three datasets—3DOM, SynthCity, and SHREC—using different configurations. The default configuration employs XYZ and RGB, while additional configurations integrate Medial Axis Transform (MAT) attributes: MAT-C (interior and exterior MAT coordinates), MAT-I (interior MAT coordinates), MAT-RS (interior and exterior radii with separation angles), MAT-SP (interior spoke vectors), and MAT-BIS (bisector angles). Notably, MAT-C and MAT-I are tailored for 3DOM and SynthCity, whereas MAT-SP and MAT-BIS are exclusive to SHREC.

	RGB	MAT-C(SP)	MAT-I(BIS)	MAT-RS
3DOM	64.4%	44.2%	48.1%	**77.9%**
SynthCity	73.0%	42.5%	49.1%	**78.7%**
SHREC	34.8%	42.7%	38.7%	**43.6%**

Table 2. Comparison of point - wise feature performance in mIoU(%) among datasets. Default configurations utilize XYZ and RGB. Additional configurations include MAT - C (Interior and Exterior MAT Coordinates), MAT - I (Interior MAT Coordinates), MAT - RS (Interior and Exterior Radii with Separation Angles), MAT - SP (Interior Spoke Vectors), and MAT - BIS (Bisector Angles). MAT - C and MAT - I are applicable to 3DOM and SynthCity, whereas MAT - SP and MAT - BIS are specific to SHREC.

For 3DOM and SynthCity, the MAT-RS configuration outperforms others, achieving the highest mIoU values of 77.9% and 78.7%, respectively. This suggests that explicit geometric information from radii and separation angles enhances feature discriminability. In contrast, MAT-C and MAT-I yield lower mIoU (44.2%–49.1%) for these datasets, likely due to ambiguous coordinate relations that hinder deep learning generalization. For SHREC, MAT-RS again leads with 43.6% mIoU, slightly surpassing MAT-SP (42.7%) and outperforming the default RGB (34.8%), indicating that even with class imbalance and density variation, structural MAT features improve segmentation.

Figure 6 visualizes point-wise classification results and scalar features, where colors in subfigure (b) denote semantic classes (e.g., ground, roof, facade), and subfigures (c)–(f) show scalar feature values from low to high via a shared legend. Similarly, Figure 7 depicts scalar features, highlighting how MAT-derived attributes capture geometric details critical for distinguishing classes like vegetation and man-made structures.
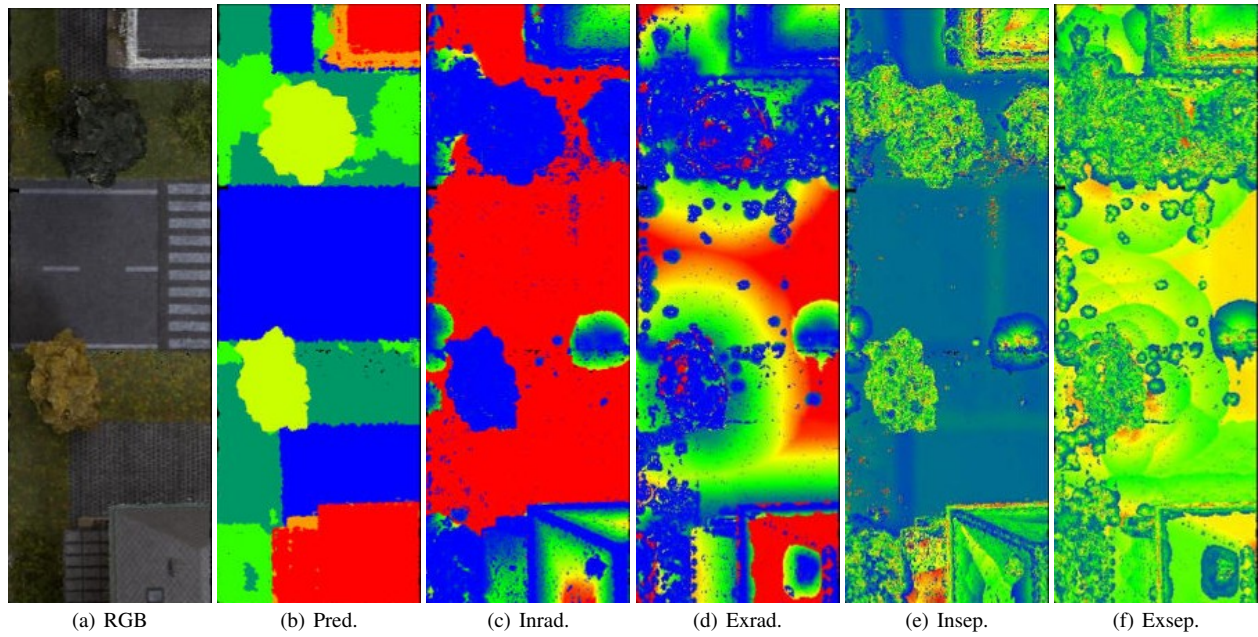
| (a) RGB | (b) Pred. | (c) Inrad. | (d) Exrad. | (e) Insep. | (f) Exsep. |

Figure 6. 3DOM point-wise classification results and scalar feature visualization. 'Pred.' is predicted classes; 'Inrad.', 'Exrad.' are interior and exterior radii; 'Insep.', 'Exsep.' are interior and exterior separation angles. In (b), colors represent ■ ground, ■ roof, ■ facade, ■ grass, ■ tree, and ■ shrub. From (c) to (f), colors ▬▬▬ indicate scalar feature values from low to high.
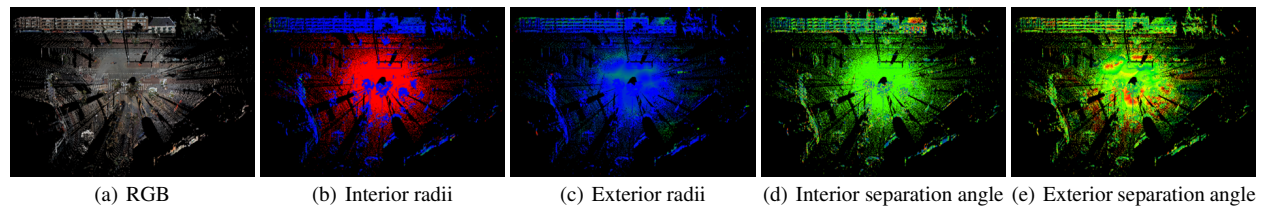

| (a) RGB | (b) Interior radii | (c) Exterior radii | (d) Interior separation angle | (e) Exterior separation angle |

Figure 7. SHREC scalar feature visualization. From (b) to (e), colors ▬▬▬ indicate scalar feature values from low to high.

4.3 Results of geometric partitioning

Table 3 compares geometric partitioning outcomes in terms of segment count (Noseg.) and semantic segmentation performance (mIoU) for 3DOM and SynthCity datasets using default, MAT, bisector, and edge weight configurations. In 3DOM, the MAT configuration increases segments by over double (10221 vs. default 4433) but reduces mIoU to 32.1%, indicating excessive over-segmentation that disrupts homogeneous structures. The bisector configuration (5375 segments, 42.0% mIoU) balances segment count and performance, while adjusting edge weights (3517 segments, 37.7% mIoU) reduces over-segmentation and improves structural integrity.

	Dataset	Default	MAT	Bisector	Edge weight
Noseg.	3DOM	4433	10221	5375	3517
	SynthCity	6578	7707	15577	7294
mIoU	3DOM	**46.1%**	32.1%	42.0%	37.7%
	SynthCity	42.6%	**46.3%**	23.9%	35.3%

Table 3. Comparison of geometric partition by number of segments (Noseg.) and semantic segmentation (mIoU)

For SynthCity, the MAT configuration moderately increases segments (7707 vs. default 6578) and boosts mIoU to 46.3%, suggesting better adaptation to synthetic scenes with uniform

density. Conversely, the bisector configuration (15577 segments, 23.9% mIoU) leads to severe over-segmentation, highlighting sensitivity to noise-free synthetic data. Edge weight adjustment (7294 segments, 35.3% mIoU) balances segmentation but underperforms the MAT configuration.

| (a) Default | (b) MAT |

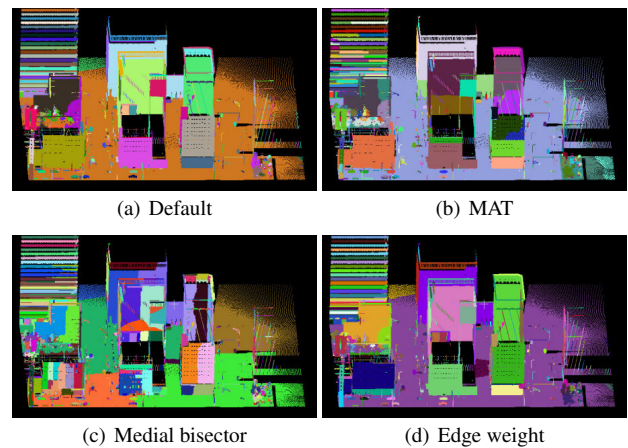| (c) Medial bisector | (d) Edge weight |

Figure 8. SynthCity geometric partition comparison.

Figure 8 visually contrasts partitions, showing that the default configuration preserves large homogeneous regions, while MAT introduces finer segmentation suitable for structured

scenes. Excessive segments from bisector or unoptimized configurations lead to fragmented partitions, degrading semantic accuracy. These results emphasize the trade-off between segment granularity and segmentation quality, with MAT attributes improving performance when appropriately regularized.

4.4 Results of graph convolutions

Table 4 evaluates 3DOM semantic segmentation using graph convolutions with different MAT-derived graph attributes: mean radii and separation angles (MeRS), min-max radii and separation angles (MiaRS), and their combinations with default attributes (DeMeRS, DeMiaRS). The default configuration achieves the highest mIoU (46.1%), while MeRS yields the best overall accuracy (OA, 74.04%), indicating that mean geometric attributes enhance global structure recognition. MiaRS performs poorly (26.2% mIoU), likely due to noise from extreme values, and DeMeRS/DeMiaRS show intermediate results, suggesting that direct integration of MAT attributes without optimization introduces redundancy or ambiguity. The overall accuracy of approximately 75% is reasonable for the 3DOM dataset, as the ground class (which is large and relatively easy to segment) dominates the scene. However, the mean IoU of 46.1% indicates that performance on minority classes (e.g., shrub, grass) remains challenging due to their geometric complexity and similarity.

	Default	MeRS.	MiaRS.	DeMeRS.	DeMiaxRS.
OA	72.64%	**74.04%**	72.77%	70.12%	73.64%
mIoU	**46.1%**	38.1%	26.2%	40.1%	36.8%

Table 4. 3DOM semantic segmentation performance measured by OA and mIoU for different graph attributes. 'MeRS.': mean radii and separation angles; 'MiaRS.': min-max radii and separation angles; 'DeMeRS.': default + mean radii and separation angles; 'DeMiaRS.': default + min-max radii and separation angles.

Figure 10 visualizes graph attributes with MAT embedding, illustrating how mean radii and separation angles (MeRS) highlight consistent structural features, while min-max values (MiaRS) emphasize local variations. The default configuration (subfigure (b)) shows balanced class discrimination, whereas MeRS (subfigure (c)) sharpens boundaries between classes like ground and vegetation. Poor performance of MiaRS (subfigure (d)) aligns with table results, as extreme values disrupt the smoothness of geometric representations critical for graph-based learning.

4.5 Results summary

In summary, MAT-derived point-wise features (especially radii and separation angles) consistently improve segmentation across datasets, while geometric partitioning requires careful regularization to avoid over-segmentation. Graph convolutions benefit from mean MAT attributes for global structure modeling, but extreme values or unoptimized combinations hinder performance. These findings underscore the value of explicit geometric information in 3D point cloud analysis, contingent on proper integration and regularization for different tasks.

4.6 Limitations of MAT

While the integration of MAT-derived features enhances semantic segmentation in controlled environments, our experi-

ments reveal inherent limitations in noisy, occluded, or geometrically irregular settings. These limitations stem from MAT's sensitivity to point cloud quality and structural complexity:

Noise Sensitivity: MAT computation relies on stable normal vectors and continuous surfaces. In noisy environments (e.g., mobile LiDAR with irregular sampling), noise severely degrades MAT's utility, causing medial balls to overshrink and collapse radii and separation angles into near-identical values. This eliminates the geometric discriminability of MAT-RS features, reducing their performance to baseline RGB levels. Such sensitivity arises because MAT atoms depend on maximally inscribed balls, which are unstable under point perturbations.

Occlusion and Sparsity: Thin or sparsely sampled structures (e.g., poles, wires) yields unstable MAT representations. In the SHREC dataset, the "pole" class consistently achieved 0% IoU across all MAT-enhanced experiments due to insufficient point density for reliable medial ball computation. Similarly, facade decorations in the 3DOM dataset generated fragmented medial sheets, exacerbating over-segmentation during geometric partitioning. This fragmentation is intrinsic to MAT's requirement of closed surfaces; partial scans break medial continuity.

Geometric Irregularities: Exterior MAT properties (radii and separation angles) introduced ambiguity in complex urban scenes. When used in Superpoint Graph partitioning, exterior features incorrectly linked distant objects, reducing performance compared to interior-only features. This occurs since exterior MAT captures inter-object relationships rather than intrinsic geometry, making it unsuitable for defining homogeneous superpoints.

Future Mitigations: To address these limitations, we propose: 1) *Adaptive MAT Filtering*: Developing noise-robust MAT extraction that adjusts denoising thresholds (planar/preservation) based on local point density and curvature to stabilize medial atoms in sparse regions. 2) *Learning-Based MAT Extraction*: Replacing heuristic MAT computation with neural estimators (e.g., graph networks) trained on noisy-clean point cloud pairs to recover stable skeletons under occlusion. 3) *Topology-Aware Partitioning*: Aligning superpoint graphs with MAT sheet adjacency (rather than spatial proximity) to leverage inherent structural hierarchies in irregular data.

5. Conclusion

This research presents the first systematic integration of 3D Medial Axis Transform (MAT) properties into deep learning for point cloud semantic segmentation, demonstrating that explicit geometric attributes like radii and separation angles can significantly enhance segmentation performance. By embedding MAT features into PointNet++ and Superpoint Graph frameworks, we show consistent improvements in mIoU across synthetic and real-world datasets, particularly for classes challenging to distinguish using RGB alone (e.g., grass, shrubs). MAT-derived point-wise features enrich local geometric context, while graph-based integrations highlight the importance of mean structural attributes for global scene understanding. These findings underscore the value of topological descriptors in complementing appearance features, offering a new paradigm for leveraging shape semantics in deep learning.

However, the study also reveals limitations: MAT-guided geometric partitioning often leads to excessive over-segmentation

(a) Linearity (b) Planrity (c) Verticality (d) Scattering

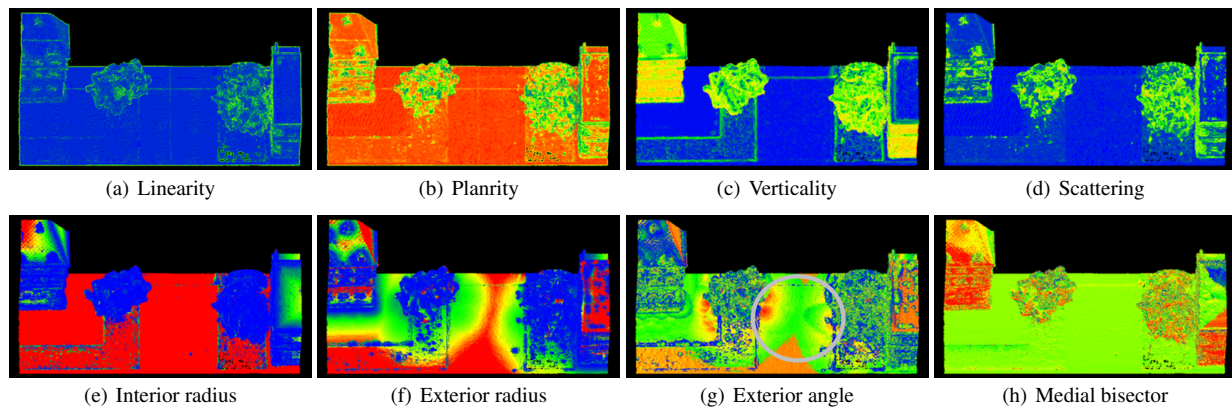(e) Interior radius (f) Exterior radius (g) Exterior angle (h) Medial bisector

Figure 9. 3DOM scalar feature visualization. The top row shows the eigenvalue-based features, and the bottom row shows the MAT-based features. Colors indicate scalar feature values from low to high.
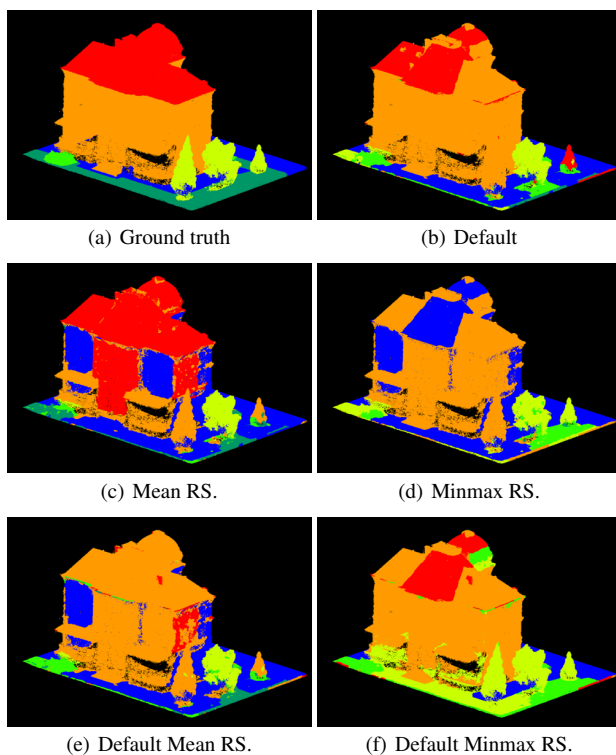

(a) Ground truth (b) Default

(c) Mean RS. (d) Minmax RS.

(e) Default Mean RS. (f) Default Minmax RS.

Figure 10. 3DOM graph attributes with MAT embedding. 'Mean RS.': mean radii and separation angles; 'Minmax RS.': min-max radii and separation angles; 'Default Mean RS.': default + mean radii and separation angles; 'Default Minmax RS.': default + min-max radii and separation angles.

in complex scenes, necessitating adaptive regularization strategies. Additionally, direct integration of extreme MAT values (e.g., min/max radii) into graph convolutions introduces noise, emphasizing the need for feature optimization. The discrepancy between MAT's theoretical structure and the practical adjacency of Superpoint Graph nodes also limits the effectiveness of edge feature enrichment, highlighting structural mismatches between topological skeletons and graph-based representations.

Future work should focus on developing automated MAT parameter selection tailored to dataset characteristics, exploring MAT's utility in aerial LiDAR datasets, and using feature attribution to dissect how MAT properties influence deep learning layers. Additionally, aligning Superpoint Graph topology with MAT's medial sheet adjacency could unlock full potential in graph-based segmentation, bridging structural modeling and semantic reasoning for more robust urban scene analysis.

Acknowledgements

References

Broersen, T., Peters, R., Ledoux, H., 2017. Automatic identification of watercourses in flat and engineered landscapes by computing the skeleton of a LiDAR point cloud. *Computers & Geosciences*, 106, 171-180.

Gao, W., Nan, L., Boom, B., Ledoux, H., 2021. SUM: A Benchmark Dataset of Semantic Urban Meshes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 179, 108-120.

Gao, W., Nan, L., Ledoux, H., 2025. SUM Parts: Benchmarking part-level semantic segmentation of urban meshes. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE/CVF.

Girardeau-Montaut, D., 2016. CloudCompare (version 2.13.2). GPL software. Retrieved from `https://www.cloudcompare.org/`.

Landrieu, L., Obozinski, G., 2017. Cut pursuit: Fast algorithms to learn piecewise constant functions on general weighted graphs. *SIAM Journal on Imaging Sciences*, 10(4), 1724–1766.

Landrieu, L., Simonovsky, M., 2017. Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs. *CoRR*, abs/1711.09869. http://arxiv.org/abs/1711.09869.

Peters, R., 2018. Geographical point cloud modelling with the 3D medial axis transform. PhD thesis, Delft University of Technology. ISBN: 978-94-6186-899-2.

Peters, R., Dukai, B., Vitalis, S., van Liempt, J., Stoter, J., 2022. Automated 3D Reconstruction of LoD2 and LoD1 Models for All 10 Million Buildings of the Netherlands. *Photogrammetric Engineering and Remote Sensing*, 88(3), 165–170.

Qi, C. R., Su, H., Mo, K., Guibas, L. J., 2016. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *CoRR*, abs/1612.00593. http://arxiv.org/abs/1612.00593.

Qi, C. R., Yi, L., Su, H., Guibas, L. J., 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *CoRR*, abs/1706.02413. http://arxiv.org/abs/1706.02413.

Thomas, H., Qi, C. R., Deschaud, J., Marcotegui, B., Goulette, F., Guibas, L. J., 2019. KPConv: Flexible and Deformable Convolution for Point Clouds. *CoRR*, abs/1904.08889. http://arxiv.org/abs/1904.08889.

Özdemir, E., Toschi, I., Remondino, F., 2019. A multi-purpose benchmark for photogrammetric urban 3D reconstruction in a controlled environment. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-1/W2, 53–60. https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-1-W2/53/2019/.