applied sciences

MDPI

Article

Attention-Guided Semantic Segmentation and Scan-to-Model Geometric Reconstruction of Underground Tunnels from Mobile Laser Scanning

Yingjia Huang, Jiang Ye *, Xiaohui Li and Jingliang Du

College of Earth and Planetary Sciences, Chengdu University of Technology, Chengdu 610059, China; huangyingjia@stu.cdut.edu.cn (Y.H.); lixiaohui@stu.cdut.edu.cn (X.L.); dujingliang@stu.cdut.edu.cn (J.D.)
* Correspondence: yejiang@cdut.edu.cn

Abstract

Mobile Laser Scanning (MLS) integrated with Simultaneous Localization and Mapping (SLAM) has emerged as a key technology for digitizing GNSS-denied environments, such as underground mines. However, the automated interpretation of unstructured, high-density point clouds into semantic engineering models remains challenging due to extreme geometric anisotropy in point distributions and severe class imbalance inherent to narrow tunnel environments. To address these issues, this study proposes a highly automated scan-to-model framework for precise semantic segmentation and vectorized two-dimensional (2D) profile reconstruction. First, an enhanced hierarchical deep learning network tailored for point clouds is introduced. The architecture incorporates a context-aware sampling strategy with an expanded receptive field of up to 10 m to preserve axial continuity, coupled with a spatial–geometric dual-attention mechanism to refine boundary delineation. In addition, a composite Focal–Dice loss function is employed to alleviate the dominance of wall points during network training. Experimental validation on a field-collected dataset comprising 16 mine tunnels demonstrates that the proposed model achieves a mean Intersection over Union (mIoU) of 85.15% (±0.29%) and an Overall Accuracy (OA) of 95.13% (±0.13%). Building on this semantic foundation, a robust geometric modeling pipeline is established using curvature-guided filtering and density-adaptive B-spline fitting. The reconstructed profiles accurately recover the geometric mean surface of the tunnel wall, yielding an overall filtered Root Mean Square Error (RMSE) of 4.96 ± 0.48 cm. The proposed framework provides an efficient end-to-end solution for deformation analysis and digital twinning of underground mining infrastructure.

Keywords: Mobile Laser Scanning; point cloud semantic segmentation; geometric modeling; attention mechanism; mine tunnel

Check for updates

Academic Editor: Ludovit Kovanic

Received: 28 January 2026
Revised: 10 March 2026
Accepted: 19 March 2026
Published: 21 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

As underground space development accelerates, the demand for high-precision, efficient, and automated geometric mapping of complex environments—particularly narrow and cluttered mine tunnels—has become increasingly urgent [1,2]. Accurate tunnel geometry is a fundamental prerequisite for critical engineering tasks such as deformation monitoring, structural safety assessment, and digital mine management. However, traditional surveying techniques, including total station measurements, are predominantly manual, time-consuming, and poorly suited to the harsh, confined, and dynamic conditions

of underground environments. These limitations lead to low operational efficiency, high labor costs, and inconsistent measurement quality, making such approaches incompatible with the demands of modern mining operations [3].

In recent years, handheld Simultaneous Localization and Mapping (SLAM) devices have emerged as a practical alternative for underground mapping. By integrating LiDAR-based ranging with real-time localization, SLAM systems enable rapid and flexible acquisition of dense three-dimensional point clouds with minimal setup requirements and strong robustness to environmental disturbances [4–7]. Ongoing advancements in sensor hardware and localization algorithms have further expanded their applicability in underground engineering and geospatial measurement tasks [8–10]. Nevertheless, while SLAM significantly improves data acquisition efficiency, it does not directly provide the structured geometric representations required for downstream engineering analysis.

In practical tunnel engineering, cross-sectional geometry serves as a critical basis for deformation assessment and safety evaluation [11]. Reliable comparison of tunnel cross-sections over time requires point cloud data that is both geometrically accurate and structurally consistent. However, raw SLAM point clouds often contain substantial non-structural elements—such as mine carts, rail tracks, ventilation ducts, and cables—that severely interfere with geometric interpretation. Conventional post-processing workflows typically rely on manual slicing, heuristic filtering, and expert-driven interpretation, which not only limits automation but also compromises scalability and reproducibility.

Previous studies have attempted to address these issues using geometric filtering techniques. Nie et al. [12] combined statistical and median filtering to suppress noise; Zhao et al. [13] employed Delaunay triangulation to remove non-wall points; and several approaches developed for subway tunnel analysis have demonstrated effectiveness in short-distance or low-curvature scenarios [14,15]. However, real-world mine tunnels are often characterized by long spatial extents, large curvature variations, and non-uniform cross-sections. Under such conditions, traditional filtering-based methods suffer from sensitivity to parameter selection and increasing computational cost, resulting in degraded performance in high-curvature or long-range segments [16].

To reduce reliance on handcrafted rules, point cloud semantic segmentation has been increasingly adopted as a means of isolating structurally meaningful components. By classifying points into semantically distinct categories based on geometric and contextual cues, semantic segmentation enables automatic identification of tunnel walls and other structural elements [17]. This paradigm has been successfully applied in various engineering domains, including tunnel maintenance [18], sewer inspection [19], and infrastructure monitoring. Representative models include classic architectures such as PointNet++ [20], PointCNN [21], and RandLA-Net [22], as well as recent state-of-the-art attention-based models like Point Transformer V3 (PTv3) [23].

Among these methods, PointNet++ has demonstrated strong generalization capability in diverse scenarios such as urban as-built modeling [24], coal yard safety inspection [25], and transmission line extraction [26,27]. Recent studies have further enhanced its performance through attention mechanisms, including GateNet [28], SO-PointNet++ [29], and coordinate attention modules [30]. Similarly, recent advancements in remote sensing have explored graph attention and transformer-based networks for large-scale point cloud segmentation [31]. More recently, specialized frameworks such as EDA-TCNet have introduced dual-attention mechanisms to refine segmentation in tunnel construction, while other task-specific approaches have focused on the precise extraction of tunnel face excavation areas using B-spline interpolation [32,33]. However, directly applying these models to entire mine tunnel environments remains challenging. Unlike the aforementioned studies that often prioritize localized excavation zones or relatively regular tunnel segments,

real-world mine tunnels exhibit extreme geometric anisotropy—extending hundreds of meters longitudinally while spanning only a few meters transversely—which renders conventional spherical neighborhood sampling ineffective for capturing long-range structural continuity. Moreover, uneven point density and severe class imbalance, where tunnel walls dominate the scene while interfering objects occupy only a small proportion, further degrade segmentation robustness in these vast, cluttered environments.

More importantly, segmentation alone does not address the full engineering problem. Tunnel deformation analysis and geometric assessment ultimately require vectorized, continuous, and repeatable cross-sectional representations rather than discrete semantic labels. Existing workflows typically treat segmentation and geometric modeling as separate stages, often involving manual intervention that disrupts spatial continuity and limits end-to-end automation [34].

To bridge this gap, this study proposes a highly automated, engineering-oriented framework that integrates SLAM-based data acquisition, tunnel-adaptive semantic segmentation, and robust geometric modeling into a unified workflow. Instead of emphasizing algorithmic novelty, the proposed approach focuses on methodological robustness, automation, and practical applicability in real mine environments.

The main contributions of this work are summarized as follows:

1. A highly automated, tunnel-adaptive processing framework. We propose an end-to-end methodology that integrates SLAM-based acquisition, deep semantic segmentation, and vectorized modeling to address the challenges of geometric anisotropy and clutter in narrow mines. By replacing manual filtering with intelligent perception, the framework achieves an Overall Accuracy (OA) of 95.13% ($\pm$0.13%) on real-world datasets, providing a unified solution for digitizing underground environments.

2. An enhanced geometry-aware segmentation network. To overcome long-range contextual loss and severe class imbalance, we improve the PointNet++ architecture by introducing a hierarchical sampling strategy (up to 10 m), a spatial–geometric dual-attention (GASE) module, and a composite Focal–Dice loss. These algorithmic innovations yielding a mean Intersection over Union (mIoU) of 85.15% ($\pm$0.29%) (a 6.93% improvement over the baseline) and an F_1-Score of 86.48% ($\pm$0.28%) for small-sample interference objects.

3. An engineering-oriented geometric reconstruction pipeline. We establish a robust application workflow for transforming discrete wall points into continuous vectorized profiles using curvature-guided sampling and density-adaptive B-spline fitting. The resulting cross-sections accurately recover the tunnel's geometric mean surface with an overall filtered Root Mean Square Error (RMSE) of 4.96 $\pm$ 0.48 cm, satisfying the sub-decimeter precision requirements for deformation monitoring and safety assessment.

2. Data and Methodology

In this study, we develop an end-to-end framework for highly automated tunnel geometry extraction from laser SLAM point clouds, with a particular focus on practical applicability in narrow and cluttered mine environments. The proposed methodology integrates tunnel-adaptive semantic segmentation with automated geometric modeling to transform raw point cloud data into vectorized tunnel cross-sections suitable for engineering analysis.

As illustrated in Figure 1, the overall workflow consists of three major stages: (1) data acquisition and preprocessing, (2) semantic segmentation of tunnel inner walls, and (3) automated cross-sectional profile reconstruction. The framework is designed to minimize manual intervention throughout the entire pipeline. Specifically, the preprocessing stage includes a preliminary manual cropping step to remove far-field sparse point clouds at the extremities of the acquired data. These sparse regions are inherent instrumental artifacts

resulting from the maximum effective range limits and laser divergence of the Mobile Laser Scanning (MLS) equipment. Removing them ensures that the subsequent algorithms operate within a reliable geometric density, thereby guaranteeing the boundary quality of the final reconstructed models. Aside from defining this effective Region of Interest, from handheld SLAM-based data collection to parametric fitting of tunnel cross-sections, the proposed highly automated system enables high-throughput and standardized geometry extraction, significantly reducing reliance on skilled labor and improving reproducibility in engineering practice.

To support this workflow, a specialized semantic segmentation dataset for mine tunnels was constructed using laser SLAM-acquired point cloud data. Each point cloud was annotated with two semantic categories: inner wall, representing the structural tunnel surface, and interfering objects, including rails, ventilation ducts, cables, and other non-structural elements commonly present in underground mine environments. An enhanced PointNet++ model is employed to perform semantic segmentation and accurately extract the inner wall regions, which serve as the geometric basis for subsequent profile reconstruction.

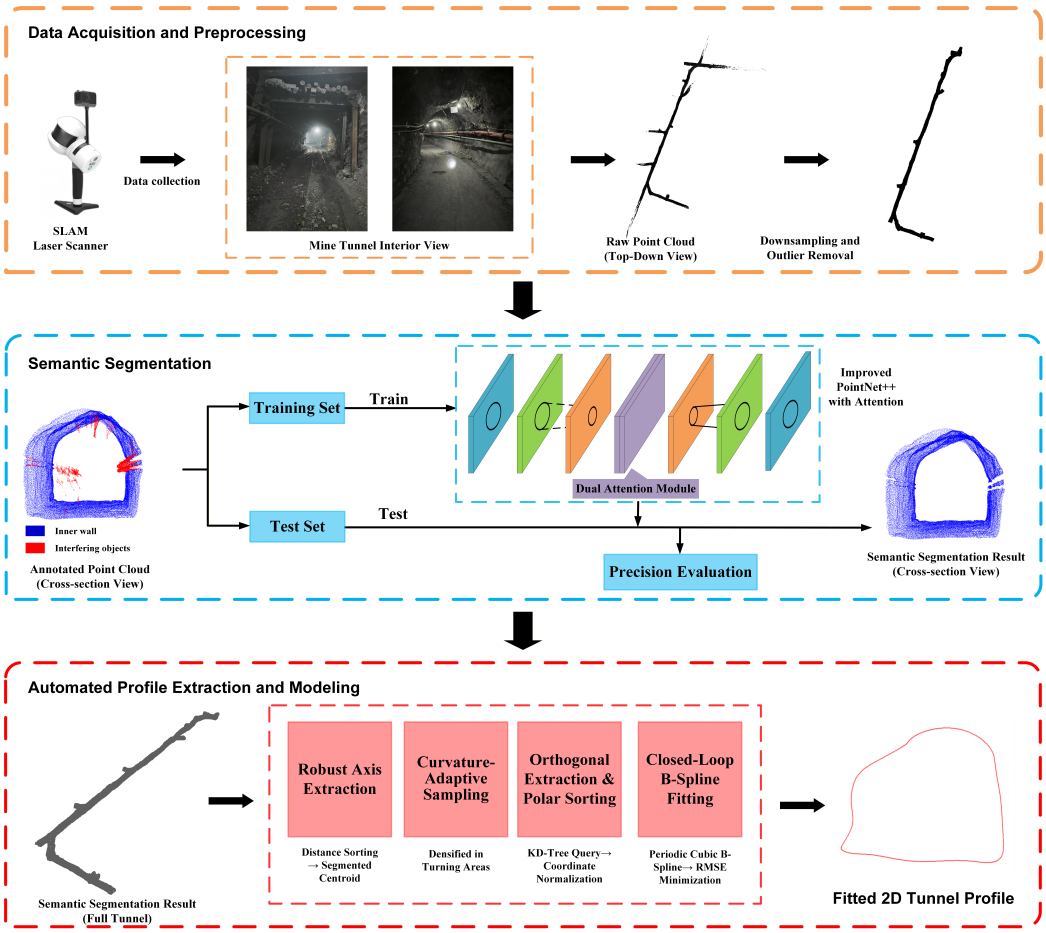

Figure 1. Workflow of point cloud semantic segmentation and profile extraction for narrow mine tunnels.

2.1. Data Acquisition and Preprocessing

The experimental data were collected from an underground tunnel project in a metal mine located in the Xizang region. A handheld 3D laser SLAM scanner was used to acquire high-density point cloud data under real operating conditions. Handheld SLAM systems are particularly suitable for underground mine environments due to their flexibility, rapid deployment, and ability to operate without external positioning infrastructure.

The raw point clouds are characterized by high-density nature, which pose challenges in terms of computational efficiency and storage requirements [35]. To address these issues, spatial downsampling was performed using CloudCompare software [36], with a minimum point spacing of 0.03 m. This downsampling strategy effectively reduces point density while preserving the essential geometric characteristics of tunnel structures, providing a balance between computational efficiency and geometric fidelity.

The quality of SLAM-acquired point clouds is influenced by multiple factors, including sensor accuracy, surface reflectivity, edge effects, and environmental conditions such as temperature and humidity [37]. These factors may introduce outliers and isolated noise points that adversely affect subsequent segmentation and geometric modeling. To ensure dataset reliability, manual inspection and removal of obvious outliers were conducted during preprocessing. Although this step involves limited human intervention, it reflects common engineering practice and significantly improves data stability for model training and evaluation.

For experimental validation, four middle sections of the mine tunnel network were selected, comprising a total of 16 individual tunnels, each with a length ranging from several hundred meters. The complete dataset contains approximately 80 million points. The tunnels were divided into training and testing subsets using a 3:1 split, with 12 tunnels used for model training and 4 tunnels reserved for testing. This split ensures sufficient scene diversity for training while reserving independent tunnel sections for unbiased performance evaluation. All point clouds were manually annotated into two semantic classes—inner wall and interfering objects—to establish ground truth labels for supervised learning. Representative examples of the labeled data are shown in Figure 2.

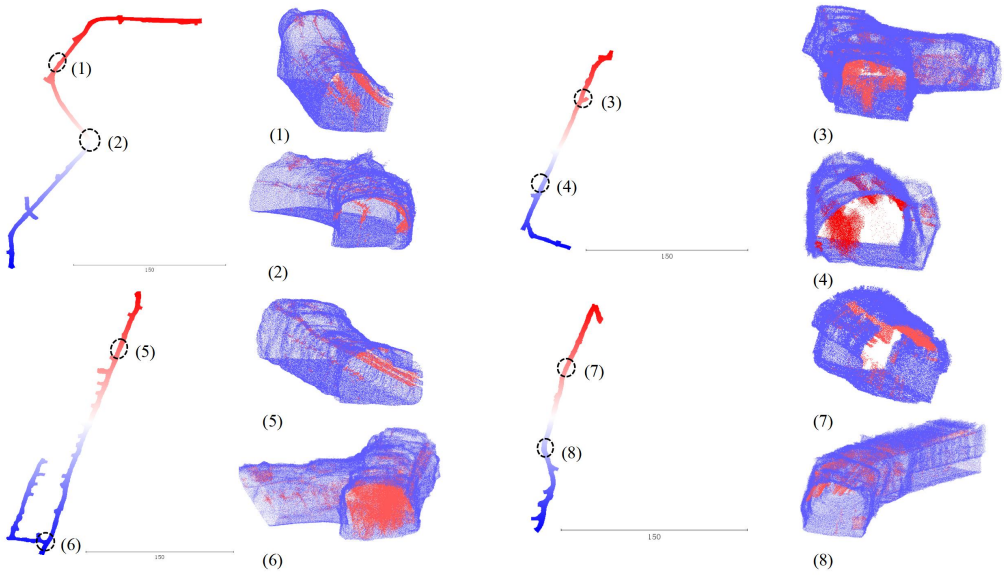

Figure 2. Examples of manually annotated tunnel point clouds. For each sample, the left image shows the top view of the tunnel, while the right image presents a three-dimensional view of the region highlighted by the black circle. The numbers (1)–(8) denote different local regions, where blue points represent the inner wall and red points represent interfering objects.

2.2. Improved PointNet++ Network Architecture for Tunnel Environments

Unlike conventional indoor point cloud datasets, mine tunnel data exhibit extreme geometric anisotropy—stretching hundreds of meters in length but only a few meters in width. Such elongated spatial configurations lead to long-range contextual dependencies that fixed-scale PointNet++ structures cannot capture effectively. To address these challenges, this study proposes a task-oriented enhancement of PointNet++ tailored for

tunnel environments. Specifically, the design focuses on enhancing geometric continuity perception, boundary sensitivity, and robustness under severe class imbalance—all of which are critical requirements in underground engineering applications. The proposed enhancements include: (1) a geometry-aware hierarchical sampling strategy expands the receptive field to preserve both local detail and axial continuity; (2) a dual-attention mechanism integrates spatial and geometry-aware channel attention to emphasize structural boundaries; and (3) a composite Focal–Dice loss mitigates class imbalance and improves boundary consistency [38,39].

2.2.1. Hierarchical Expansion of Sampling Radius

In the traditional PointNet++ framework, the Set Abstraction (SA) module employs a maximum sampling radius of 0.8 m, which is insufficient to capture the continuous and elongated spatial structures characteristic of mine tunnels, leading to the loss of long-range contextual information. To better align the receptive field with the physical geometry of tunnels, the sampling radii of the four SA layers (SA1–SA4) are exponentially expanded to 1.0 m, 2.0 m, 4.0 m, and 10.0 m, respectively. Each layer corresponds to a distinct spatial scale within the tunnel environment: SA1 captures fine-grained local details such as surface roughness; SA2 and SA3 correspond to typical tunnel cross-sectional widths of approximately 4.0–8.0 m, facilitating mesoscale structural perception; and SA4 models the global axial alignment and overall geometric continuity. This hierarchical sampling design extends the network's spatial perception while reflecting the anisotropic geometry and layered continuity of tunnels. As a result, it forms a task-driven, multi-scale feature extraction mechanism that effectively balances local precision with global contextual consistency. These radii were empirically determined based on the statistical distribution of tunnel widths and axial continuity observed in the dataset, ensuring that each SA layer approximately corresponds to a meaningful physical scale. For tunnels with significantly larger cross-sectional dimensions, the sampling radii can be proportionally adjusted following the same hierarchical scaling principle, without modifying the overall network structure.

2.2.2. Dual Attention Modules for Spatial and Geometric Features

Although attention mechanisms have been extensively studied in the field of 3D vision [40,41], their application to laser SLAM point cloud segmentation tasks in narrow mine tunnels remains in the exploratory phase. The distinct anisotropic characteristics of tunnel environments—namely, strong axial elongation, limited cross-sectional width, complex curvature variations, and occluded regions—pose significant challenges. As a result, traditional attention mechanisms struggle to effectively model long-range spatial dependencies and capture local structural variations. To address these challenges, we have redesigned a dual-attention module specifically tailored for feature extraction in tunnel point clouds. This module simultaneously captures spatially distributed features and incorporates geometric priors, thereby improving the network's ability to represent features accurately and recognize structural patterns. The module consists of two complementary branches. The Lightweight Spatial Attention Branch enhances the network's sensitivity to local boundaries and regions with geometric discontinuities. On the other hand, the Geometry-Aware Squeeze-and-Excitation Branch (GASE) introduces channel-level feature recalibration by incorporating the statistical properties of point cloud coordinates, such as variance and mean. This approach allows the recalibration of channels to be specifically adapted to the geometry of the tunnel environment. Together, these two branches enable the module to highlight critical regions locally while maintaining global geometric consistency.

(1) Lightweight Spatial Attention Module, which emphasizes local boundaries, curvature transitions, and occluded areas. The spatial attention branch is designed to enhance feature responses in local regions with strong geometric significance, including structural boundaries, curvature transition zones, and partially occluded areas commonly observed in tunnel environments. Instead of adopting computationally expensive global attention mechanisms, this branch employs a lightweight design based on channel compression and point-wise convolution, enabling efficient learning of spatial importance while maintaining compatibility with the hierarchical structure of the Set Abstraction layers.

As illustrated in Figure 3, let the input feature tensor be denoted as $X \in \mathbb{R}^{B\times C\times K\times S}$, where B denotes the batch size, C is the number of feature channels, K represents the number of localized points, and S corresponds to the number of sampled points. To reduce computational complexity and introduce nonlinearity, the module first applies a 1×1 convolution to compress the feature channels from C to $C/8$, followed by a ReLU activation function. Subsequently, a second 1×1 convolution generates a single-channel spatial attention map $A_s \in \mathbb{R}^{B\times 1\times K\times S}$. This attention map is normalized using a sigmoid activation function to produce spatial attention weights in the range $[0, 1]$. Finally, the learned attention weights are applied to the original feature tensor via element-wise multiplication, $X' = X \odot A_s$, yielding a refined feature representation X'. Here, A_s is broadcast along the channel dimension to ensure consistency with the input feature map. This mechanism enhances the network's sensitivity to informative spatial regions and improves the discrimination of geometrically critical features, particularly in boundary and transition regions, while introducing minimal computational overhead.

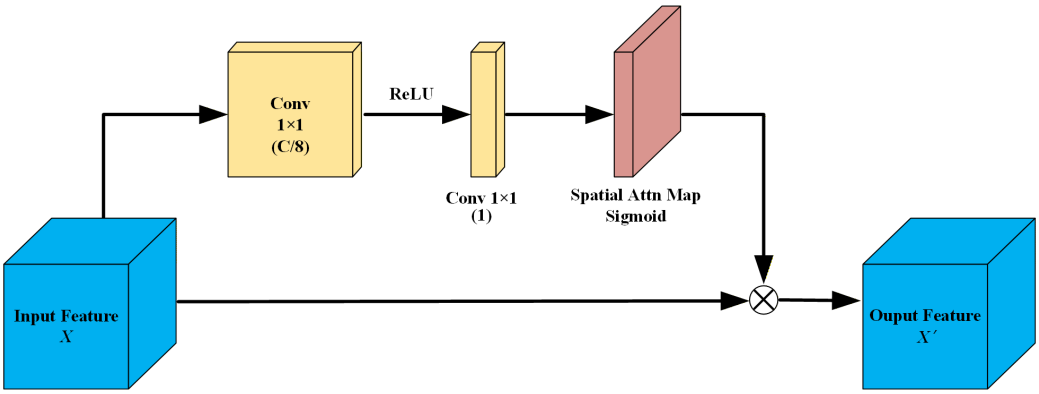

Figure 3. Architecture of the spatial attention module embedded in the Set Abstraction layer.

(2) Geometry-Aware Channel Attention, which recalibrates channel responses using statistical descriptors of point coordinates to enable geometry-aware feature enhancement.

To enhance geometric feature representation in tunnel point cloud data, we introduce a Geometry-Aware Squeeze-and-Excitation (GASE) module prior to the feature propagation layer. This module extends the conventional SE mechanism by explicitly incorporating geometric priors derived from point coordinates, thereby adapting channel-wise feature responses to the anisotropic and structured characteristics of tunnel environments, as illustrated in Figure 4.

Unlike the standard SE module [42], which models global channel dependencies solely based on aggregated feature activations (thus remaining fundamentally blind to the explicit 3D spatial distribution of the data), the proposed GASE module explicitly incorporates the macro-geometric shape of the point cloud. Furthermore, compared to conventional geometry-enhanced attention approaches that typically rely on the simple early concatenation of raw coordinates or normal vectors—which often struggle to capture global anisotropic characteristics—our GASE module adopts a statistical approach. It

extracts global statistical descriptors (specifically, the means and variances of the spatial coordinates) to directly guide channel recalibration. This design allows the network to selectively emphasize geometry-sensitive channels based on the tunnel's inherent anisotropy. For example, in ceiling regions, features correlated with the vertical (Z-axis) structure are strengthened, whereas in corner or side-wall regions, features associated with lateral continuity are adaptively enhanced. Such geometry-aware modulation enables more discriminative and spatially adaptive feature encoding for complex tunnel structures.

(1) Standard SE path

Given the input feature tensor $X \in \mathbb{R}^{B\times C\times K\times S}$, global average pooling is applied to obtain the channel-wise descriptor vector $s \in \mathbb{R}^{B\times C}$, defined as:

$$s_c = \frac{1}{K\times S}\sum_{k=1}^{K}\sum_{s=1}^{S} X_c(k,s) \tag{1}$$

The descriptor s is then passed through two fully connected layers with a reduction ratio r, where a ReLU activation follows the first (descending) layer and a Sigmoid activation follows the second (ascending) layer to generate the standard channel attention weights:

$$z = \sigma(W_2\,\mathrm{ReLU}(W_1 s)) \in \mathbb{R}^{B\times C} \tag{2}$$

where $W_1 \in \mathbb{R}^{\frac{C}{r}\times C}$, $W_2 \in \mathbb{R}^{C\times\frac{C}{r}}$, $\sigma(\cdot)$ is the Sigmoid function.

(2) Geometric sensing path

To explicitly capture geometric characteristics, the coordinate channels are extracted from the input features:

$$X_{geo} = X_{(:,1:3,:,:)} \in \mathbb{R}^{B\times 3\times K\times S} \tag{3}$$

Based on X_{geo} global geometric statistics are computed, including the means and variances of the three coordinate dimensions:

$$\begin{cases} \mu_d = \dfrac{1}{K\times S}\displaystyle\sum_{k=1}^{K}\sum_{s=1}^{S} X_{geo}(d,k,s) \\ \sigma_d^2 = \dfrac{1}{K\times S}\displaystyle\sum_{k=1}^{K}\sum_{s=1}^{S}\left(X_{geo}(d,k,s)-\mu_d\right)^2 \end{cases} \quad d\in\{x,y,z\} \tag{4}$$

These six scalar statistics $[\mu_x,\mu_y,\mu_z,\sigma_x^2,\sigma_y^2,\sigma_z^2]$ are concatenated to form a geometric descriptor vector, which is mapped through two fully connected layers to generate geometry-aware channel weights:

$$W = \sigma\left(W_{geo2}\,\mathrm{ReLU}(W_{geo1} g)\right) \in \mathbb{R}^{B\times C} \tag{5}$$

where g denotes the geometric descriptor vector, $W_{geo1} \in \mathbb{R}^{\frac{C}{r}\times 6}$ and $W_{geo2} \in \mathbb{R}^{C\times\frac{C}{r}}$.

(3) Dual-path fusion

Finally, the standard SE weights (z) and the geometry-aware weights (W) are fused via element-wise multiplication and applied to the input features:

$$\tilde{X} = X \odot (z \odot W) \tag{6}$$

where $\odot$ denotes channel-wise multiplication, z and W are broadcast along the spatial dimensions. This dual-path fusion allows the network to jointly exploit semantic feature dependencies and geometric priors, resulting in geometry-consistent feature recalibration.

As detailed in the subsequent complexity analysis, this dual-path fusion introduces an additional 0.217 M parameters and 12.59 ms of latency, maintaining the overall efficiency of the network.

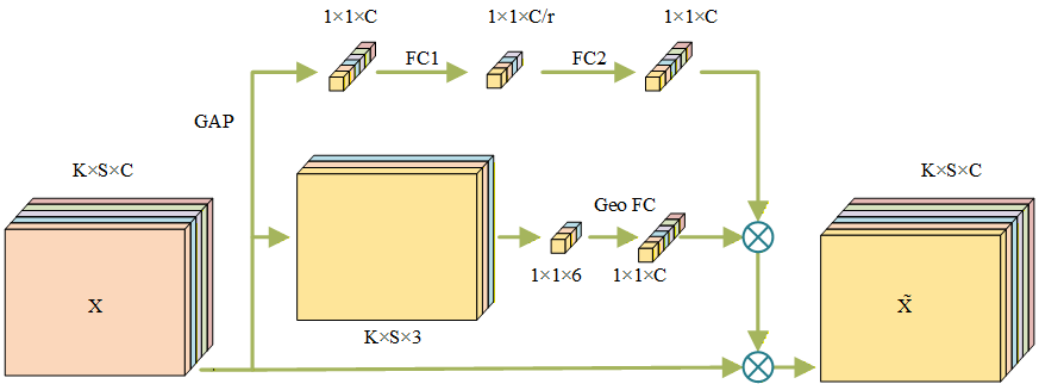

Figure 4. GASE module for channel-wise attention.

By jointly expanding the receptive field of the Set Abstraction layers and embedding a lightweight spatial–geometric dual attention mechanism, the proposed framework explicitly adapts network perception to the elongated, anisotropic, and structurally constrained geometry of underground tunnels. The large-radius hierarchical sampling strategy enables effective modeling of long-range axial continuity, while the spatial attention module emphasizes local geometric discontinuities and boundary regions. Meanwhile, the geometry-aware channel attention further incorporates coordinate-derived statistical priors to guide feature recalibration in a structure-consistent manner.

Together, these designs allow the network to preserve axial structural coherence and enhance the discriminative representation of cross-sectional geometry and boundary details, even under conditions of occlusion, uneven point density, and geometric degradation commonly encountered in narrow mine tunnels. The overall architecture of the proposed tunnel-oriented semantic segmentation framework, and the interaction between its key components, is summarized in Figure 5.

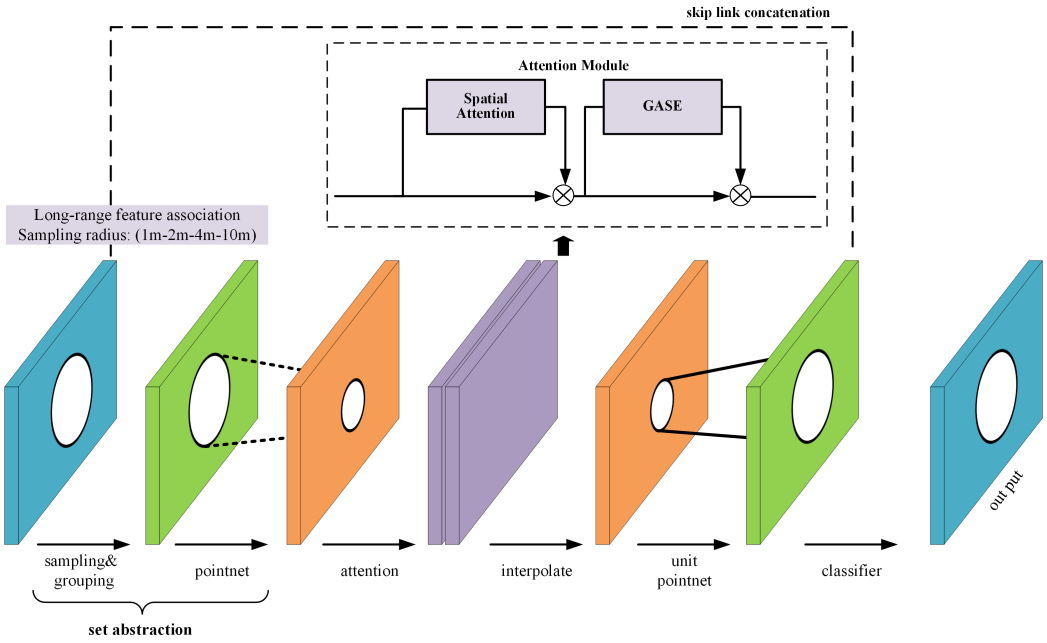

Figure 5. Architecture of the tunnel-oriented semantic segmentation framework featuring hierarchical large-radius sampling and spatial–geometric dual attention.

2.2.3. Composite Loss Function: Focal Loss + Dice Loss

Due to the pronounced class imbalance in mine tunnel point cloud data—where inner wall points typically account for more than 90% of all samples—the standard cross-entropy loss employed in the original PointNet++ architecture tends to bias the network toward the dominant class. This bias often leads to insufficient learning of minority classes, such as small-scale interference objects, and results in degraded boundary delineation. To alleviate this issue, a composite loss function that combines Focal Loss and Dice Loss is adopted in this study, jointly addressing class imbalance and structural consistency in semantic segmentation.

Focal Loss (FL) is specifically designed to mitigate class imbalance by dynamically down-weighting easy samples and focusing training on hard-to-classify points. Its formulation is given as:

$$FL(p_t) = -\alpha_t(1-p_t)^{\gamma}\log(p_t) \tag{7}$$

where p_t denotes the prediction probability of the ground-truth class, α_t is a class-balancing factor, and γ is a focusing parameter that controls the relative emphasis on hard samples. In this study, γ is set to 2.0 and α_t is set to 0.25, which effectively increases the contribution of minority-class points and prevents the network from being dominated by the majority wall class.

Dice Loss (DL) is widely used in semantic segmentation to directly optimize region overlap and shape consistency. Unlike point-wise classification losses, Dice Loss evaluates the similarity between predicted and ground-truth regions and is particularly effective in preserving structural integrity. It is defined as:

$$Dice = \frac{2\sum_i p_i y_i + \epsilon}{\sum_i p_i + \sum_i y_i + \epsilon}, \qquad \begin{cases} Dice = \dfrac{2\cdot|A\cap B|}{|A|+|B|} \\ DL = 1 - Dice \end{cases} \tag{8}$$

where p_i denotes the predicted probability, y_i denotes the corresponding ground-truth label, and ϵ (e.g., 1×10^{-6}) a small constant for numerical stability. Here $|A|$ and $|B|$ represent the predicted and ground-truth point sets, respectively. In practice, a soft Dice formulation is employed, in which $|A\cap B|$ corresponds to the summed probabilities of correctly predicted points, enabling stable gradient propagation during training.

To fully leverage the complementary strengths of the two loss terms, the final training objective is defined as a weighted combination:

$$Loss = \lambda_{fl}FL + \lambda_{dl}DL \tag{9}$$

where λ_{fl} and λ_{dl} control the relative contributions of Focal Loss and Dice Loss. Through empirical validation, both weights are set to 1.0, ensuring a balanced optimization between point-wise classification accuracy and region-level structural consistency.

This composite loss design is particularly well suited for narrow mine tunnel environments, where small interference objects are spatially sparse yet geometrically critical, and where accurate preservation of wall continuity is essential for subsequent geometric modeling and deformation analysis.

2.3. Automated Tunnel Profile Extraction and Modeling

To address the engineering imperatives of digital mine management, we developed an automated framework to transform the segmented inner-wall point clouds into vectorized geometric cross-sections. This pipeline addresses challenges such as large-scale coordinate precision, varying tunnel curvatures, and sensor noise. As illustrated in Figure 6, The

workflow consists of four consecutive stages: data preprocessing, robust central axis extraction, curvature-adaptive sampling, and parametric profile reconstruction.

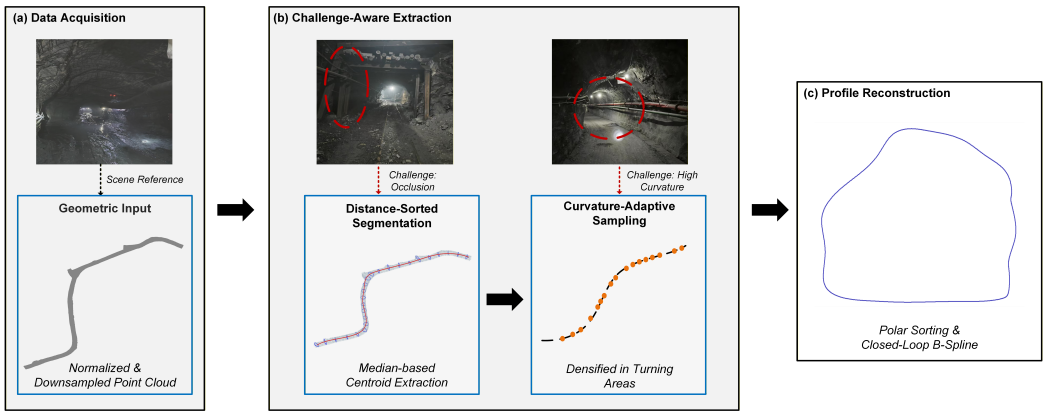

Figure 6. Workflow of the proposed automated tunnel profile extraction pipeline. The framework transforms raw SLAM point clouds into vectorized cross-sections through three sequential steps: geometric input preprocessing, curvature-guided adaptive sampling for robust axis extraction, and parametric B-spline modeling.

2.3.1. Data Preprocessing and Coordinate Normalization

Raw point cloud data acquired from laser SLAM often operates in large geospatial coordinate systems (e.g., UTM), which can induce floating-point truncation errors during complex geometric computations. To ensure numerical stability, we first normalize the point cloud P by translating its geometric centroid to the origin of a local coordinate system: $P' = P - \bar{P}$. Subsequently, a Statistical Outlier Removal (SOR) filter is applied to eliminate discrete noise, followed by a voxel grid down-sampling (voxel size set to 0.05 m) to homogenize the point density for efficient processing.

2.3.2. Distance-Sorted Central Axis Extraction

Extracting a stable central axis from curved tunnel point clouds is challenging due to the absence of a dominant global orientation. To address this issue, a distance-sorted central axis extraction strategy is adopted, which does not rely on principal direction estimation.

First, the geometric centroid of the entire point cloud is computed. The point with the maximum Euclidean distance from this centroid is selected as the starting point P_{start}. For each point $\mathbf{p}_i$, the Euclidean distance $d_i = \|\mathbf{p}_i - P_{\text{start}}\|_2$ is calculated to establish a global ordering along the tunnel's longitudinal extension. The point cloud is sorted according to d_i and uniformly divided into $N = 100$ consecutive bins with equal point counts. For each bin $\mathcal{B}_k$, the geometric median is computed to represent the local tunnel center: $\mathbf{c}_k = \text{median}(\mathcal{B}_k)$. Compared with the arithmetic mean, the median-based estimation effectively suppresses the influence of outliers such as hanging cables and irregular attachments.

To remove redundant or unstable center points, consecutive centers with inter-point distances smaller than 0.1 m are filtered out. The remaining ordered centers are then smoothed using a cubic B-spline interpolation. The smoothing factor is set proportional to the number of center points ($s = 0.5 \times |\{c_k\}|$), ensuring a balance between noise suppression and global shape preservation. The resulting spline represents a continuous and differentiable three-dimensional central axis.

2.3.3. Curvature-Guided Adaptive Cross-Section Sampling

Based on the extracted centerline, adaptive cross-sections are generated according to local geometric complexity. The centerline spline is uniformly sampled in parameter space, and the curvature $\kappa(u)$ is computed using first- and second-order derivatives. The sampling

interval is adaptively determined as $\Delta l = 1.0 + 4\exp(-5\kappa)$, where κ is the curvature magnitude of the centerline, normalized to ensure numerical stability. Here, regions with high curvature are sampled more densely, while flatter regions are sampled more sparsely. For each sampling location, the tangent direction of the centerline is used as the normal vector of the cross-sectional plane, forming an orthogonal slicing configuration.

2.3.4. Cross-Section Point Extraction

For each cross-section, candidate points are selected via a radius-based neighborhood query with a radius of 15.0 m. Points whose orthogonal distance to the slicing plane is smaller than half of the predefined thickness (0.3 m) are retained. This operation extracts a thin slab of points representing the local tunnel boundary.

2.3.5. Periodic B-Spline Profile Fitting

The extracted cross-sectional points are projected onto a local two-dimensional coordinate system defined on the slicing plane. The projected points are sorted by polar angle to ensure a consistent boundary order, and the first point is appended to the end to enforce closure.

A periodic cubic B-spline is fitted to the ordered boundary points using a small smoothing parameter proportional to the number of points ($s = 0.0001 \times N$). This configuration allows the fitted curve to closely adhere to the measured boundary while maintaining continuity and smoothness. The fitted 2D spline is subsequently mapped back to 3D space, yielding a smooth and closed tunnel cross-sectional profile.

3. Experimental Evaluation and Engineering Applications

The experimental environment is configured with an Intel® Core™ i5-14600KF CPU, an NVIDIA GeForce RTX 4070 Super GPU with 12 GB of video memory, and 32 GB of RAM. The implementation is based on the PyTorch (version 2.9.1+cu128) deep learning framework. The model is trained for 100 epochs with a batch size of 16 and an initial learning rate of 0.001. Specifically, the Adam optimizer is employed for parameter optimization with a weight decay of 1×10^{-4}. The learning rate is governed by a step decay strategy, decreasing by a factor of 0.7 every 10 epochs, with a minimum threshold set at 1×10^{-5}. For a fair comparison, the baseline PointNet++ network is trained under identical hyperparameter settings.

3.1. Evaluation Metrics

The accuracy and effectiveness of point cloud semantic segmentation are evaluated using several widely adopted performance metrics, including Intersection over Union (IoU), Mean Intersection over Union (mIoU), Overall Accuracy (OA), Precision, Recall, and the F_1-score. These metrics jointly assess both class-level performance and overall classification accuracy. The corresponding mathematical definitions are summarized in Table 1.

In these formulations, TP_i, FP_i, and FN_i denote the numbers of true positives, false positives, and false negatives for class i, respectively. The total number of semantic categories is denoted by C, and the total number of points by N. As defined in the standard practice, the $mIoU$ is computed as the arithmetic mean of IoU across all C classes, providing a balanced evaluation of accuracy and robustness, especially for imbalanced 3D point cloud data.

Table 1. Evaluation metrics used for assessing point cloud semantic segmentation performance.

Evaluation Indicators	Formulas
IoU_i	$\frac{TP_i}{TP_i + FP_i + FN_i}$
$mIoU$	$\frac{1}{C}\sum_{i=1}^{C} IoU_i$
OA	$\frac{\sum_{i=1}^{C} TP_i}{N}$
$Precision_i$	$\frac{TP_i}{TP_i + FP_i}$
$Recall_i$	$\frac{TP_i}{TP_i + FN_i}$
F_1-$Score_i$	$\frac{2 \cdot Precision_i \cdot Recall_i}{Precision_i + Recall_i}$

3.2. Experimental Results

To rigorously validate the stability of the proposed modules and mitigate performance variance caused by network initialization, our model was trained and evaluated across three independent runs using different random seeds (0, 42, and 2025). Baseline networks were evaluated under the same hardware environment and controlled training settings to ensure fair comparison. Consequently, the performance metrics reported for our method are presented as mean ± standard deviation.

Table 2 summarizes the comprehensive quantitative performance of the improved model alongside other state-of-the-art networks on the mine tunnel test set. The results indicate that our framework achieves an mIoU of 85.15% (±0.29%) and an OA of 95.13% (±0.13%). Notably, the IoU for the inner wall class reaches 94.13% (±0.16%) with a corresponding F_1-Score of 96.98% (±0.09%), which demonstrates the model's strong capability in capturing the structured geometric features of narrow mine tunnels. This high precision in wall extraction provides a reliable data foundation for subsequent vectorized profile reconstruction and engineering deformation analysis.

Table 2. Quantitative comparison of semantic segmentation performance between the proposed model and other state-of-the-art networks in tunnel scenarios.

Method	OA (%)	mIoU (%)	F_1 (%)	IoU-Inner Wall (%)	F_1-Inner Wall (%)	IoU-Interfering Objects (%)	F_1-Interfering Objects (%)
PointNet++	92.17 ± 0.08	78.22 ± 0.30	87.22 ± 0.22	90.80 ± 0.08	95.18 ± 0.04	65.65 ± 0.57	79.26 ± 0.40
RandLA-Net	91.46 ± 0.13	77.30 ± 0.17	86.60 ± 0.09	90.34 ± 0.11	94.95 ± 0.01	64.26 ± 0.24	78.24 ± 0.18
PointCNN	91.87 ± 0.19	77.50 ± 0.17	86.77 ± 0.10	90.06 ± 0.17	94.75 ± 0.10	64.94 ± 0.16	78.75 ± 0.12
PTv3	94.66 ± 0.08	**86.42 ± 0.22**	**92.57 ± 0.13**	93.35 ± 0.09	96.56 ± 0.05	**79.48 ± 0.34**	**88.57 ± 0.21**
Ours	**95.13 ± 0.13**	85.15 ± 0.29	91.73 ± 0.18	**94.13 ± 0.16**	**96.98 ± 0.09**	76.18 ± 0.44	86.48 ± 0.28

Note: Bold values indicate the best performance in each column.

In contrast, the segmentation performance for the interference class yields a relatively lower IoU of 76.18% (±0.44%) and a corresponding F_1-Score of 86.48% (±0.28%). This limitation is primarily attributed to the extreme class imbalance, where interference objects constitute less than 10% of the training samples, and the uneven point density of laser SLAM data that introduces ambiguity in local feature representation. Nevertheless, the model demonstrates strong predictive robustness and generalization ability in complex underground environments. These experimental outcomes confirm that the proposed attention-guided framework effectively mitigates the challenges of severe geometric anisotropy and class imbalance in narrow mine tunnels.

The visualization comparison results are presented in Figure 7. Across four sets of representative samples, a comparison between the original tunnel point cloud (left column)

and the segmentation results produced by the proposed model (right column) demonstrates that the inner wall is accurately segmented, with the resulting structure exhibiting well-preserved geometric continuity. Notably, the model performs effectively in high-curvature areas and regions with complex topological variations, achieving precise delineation of structural boundaries.

However, minor mis-segmentation artifacts are observed in the transitional regions between interference objects and the inner wall, as specifically highlighted by the red dashed circles in Figure 7b. These localized errors are primarily attributed to the ambiguity in feature representation near class boundaries, where point cloud characteristics of adjacent categories exhibit high similarity, challenging the model's discriminative capability.

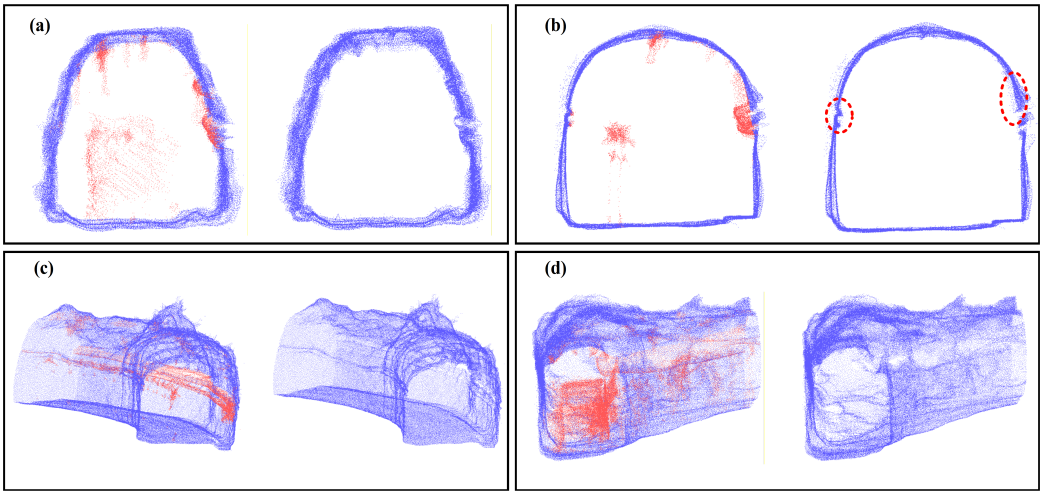

Figure 7. Comparison of cross-sectional profiles before and after segmentation: (**a**,**b**) show regular tunnel sections, (**c**,**d**) correspond to curved tunnel segments. In each group, the left sub-image shows the raw point cloud, and the right shows the segmentation result. The red dashed circles in (**b**) highlight minor mis-segmentation artifacts in the transitional regions.

3.3. Comparison with Baseline Networks

To comprehensively evaluate the segmentation performance of the proposed framework, comparative experiments were conducted against representative point cloud segmentation networks, including classic lightweight models (PointNet++, RandLA-Net, PointCNN) and the recent large-scale model (PTv3). The comprehensive quantitative results are summarized in Table 2.

PointNet++ employs a hierarchical multi-scale feature learning strategy, which provides strong capability in modeling local geometric structures. On the tunnel dataset, it achieved an F_1-Score of 87.22% (±0.22%), an mIoU of 78.22% (±0.30%), an IoU (inner wall) of 90.80% (±0.08%), and an overall accuracy (OA) of 92.17% (±0.08%). However, its fixed-radius sampling strategy limits the effective perception of long-range spatial dependencies, making it difficult to preserve continuous structural features along the elongated tunnel axis. As a result, segmentation discontinuities are observed in complex tunnel sections.

RandLA-Net improves computational efficiency through random sampling and lightweight local feature aggregation. It achieved an F_1-Score of 86.60% (±0.09%), an mIoU of 77.30% (±0.17%), an IoU (inner wall) of 90.34% (±0.11%), and an OA of 91.46% (±0.13%). Although effective in large-scale open scenes, its random sampling mechanism tends to lose fine-grained geometric information in narrow and confined tunnel environments, leading to reduced segmentation accuracy compared with PointNet++.

PointCNN applies adaptive convolution kernels directly on point clouds and achieved an F_1-Score of 86.77% (±0.10%), an mIoU of 77.50% (±0.17%), an IoU (inner wall) of 90.06% (±0.17%), and an OA of 91.87% (±0.19%). While its performance exceeds that of RandLA-

Net, the model does not fully exploit its feature modeling advantages in elongated tunnel geometries, resulting in limited improvement in structural continuity.

As a recent transformer-based model, PTv3 demonstrates strong global feature modeling capability, achieving an mIoU of 86.42% (±0.22%) and an interfering objects IoU of 79.48% (±0.34%). However, due to its higher memory consumption, spatial downsampling (grid size = 0.15 m) was required to fit within GPU memory constraints, resulting in a relatively sparse output representation.

In contrast, the proposed tunnel-oriented semantic segmentation framework achieves the highest Overall Accuracy (95.13%) and the best IoU for the inner wall category (94.13%). Compared with PointNet++, the proposed method improves the mIoU, IoU (inner wall), and OA by 6.93%, 3.33%, and 2.96%, respectively. While PTv3 obtains a higher overall mIoU, our model excels in extracting the precise tunnel inner wall, which is the most critical foundation for downstream engineering tasks such as vectorized profile reconstruction and deformation analysis. These results suggest that aligning network perception with tunnel-scale geometric characteristics—through extended receptive fields and spatial–geometric dual attention—effectively enhances both long-range structural continuity and boundary discrimination. Consequently, the proposed framework demonstrates stronger robustness and generalization capability in complex underground tunnel environments.

To further illustrate the model's robustness in different geometric environments, visual comparative analyses of the extracted inner walls were conducted across curved and straight tunnel segments, as depicted in Figure 8. The top row presents the curved segments, while the bottom row displays the straight sections. In the Ground Truth, the inner wall is colored blue, and interfering objects are marked in red. For the model predictions, only the points predicted as the inner wall (blue) are visualized, and regions outlined by red dashed circles highlight significant extraction artifacts.

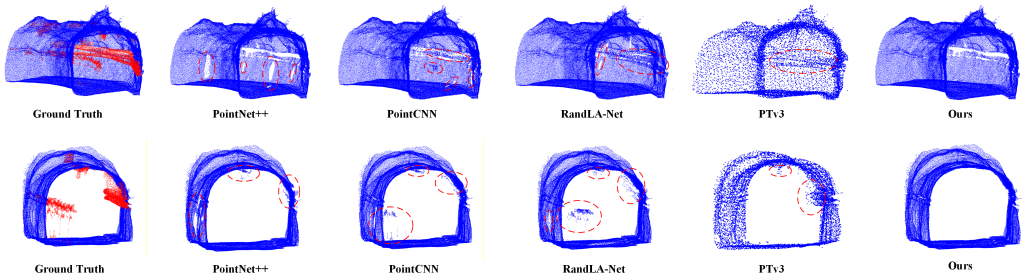

Figure 8. Visual comparison of the extracted inner tunnel walls in different geometric environments. The top row shows the curved segments, and the bottom row shows the straight sections. Red dashed circles highlight artifacts caused by misclassified interfering objects. Regions outlined by red dashed circles highlight significant extraction artifacts. Note that PTv3 is visualized at a sparse grid size of 0.15m due to the hardware memory constraints mentioned in Section 3.

As observed in both geometric scenarios, classical baseline models such as PointNet++, PointCNN, and RandLA-Net frequently misclassify complex interfering objects (e.g., hanging pipes) as the tunnel wall. This results in severe geometric artifacts and noisy bumps on the reconstructed surface. Notably, although the recent SOTA model PTv3 possesses a large parameter scale, it still exhibits evident missegmentation, erroneously classifying part of the interfering objects as the inner wall. Furthermore, as a consequence of the aggressive spatial downsampling mentioned previously, the output of PTv3 is inherently sparse, leading to a substantial loss of fine-grained structural details in both straight and curved sections.

In comparison, our proposed model accurately delineates these complex boundaries regardless of the tunnel geometry and without the need for extreme downsampling. It successfully filters out challenging interfering objects—yielding a smooth, artifact-free

surface in both curved and straight segments—while accurately preserving the original density and continuous geometric topology of the tunnel walls. Overall, the improved model exhibits superior segmentation quality, enhanced structural recognition accuracy, and highly effective engineering feasibility across diverse tunnel environments.

3.4. Ablation Study

To comprehensively evaluate the effectiveness of the proposed network enhancements and to rigorously reduce the potential influence of random initialization, a series of ablation experiments were conducted. Point cloud networks can be sensitive to initialization; therefore, evaluating models based on a single run may introduce randomness that obscures the true contribution of architectural changes. To mitigate this concern, all models in this ablation study were independently trained and evaluated across three distinct random seeds (0, 42, and 2025). The results are reported as the mean ± standard deviation.

The improved architecture incorporates three key components: an extended sampling radius, a spatial–geometric dual attention mechanism, and a combinatorial loss function. These components were evaluated individually against the baseline model (PointNet++). Table 3 presents the comparative results, where "−" indicates the absence of a module and "√" denotes its inclusion.

Table 3. Performance comparison of different module combinations in the ablation study (Mean ± Std across 3 random seeds).

Extended Sampling Radius	Dual-Attention Mechanism	Combinatorial Loss Function	OA (%)	mIoU (%)	F_1 (%)	IoU-Inner Wall (%)	F_1-Inner Wall (%)	IoU-Interfering Objects (%)	F_1-Interfering Objects (%)
−	−	−	92.17 ± 0.08	78.22 ± 0.30	87.22 ± 0.22	90.80 ± 0.08	95.18 ± 0.04	65.65 ± 0.57	79.26 ± 0.40
√	−	−	94.86 ± 0.03	84.76 ± 0.06	91.48 ± 0.04	93.88 ± 0.04	96.84 ± 0.02	75.63 ± 0.10	86.12 ± 0.06
−	√	−	93.48 ± 0.57	81.48 ± 1.13	89.40 ± 0.73	92.26 ± 0.70	95.97 ± 0.38	70.70 ± 1.57	82.83 ± 1.08
−	−	√	92.45 ± 0.12	78.76 ± 0.50	87.58 ± 0.35	91.12 ± 0.11	95.36 ± 0.06	66.40 ± 0.87	79.81 ± 0.61
√	√	√	**95.13 ± 0.13**	**85.15 ± 0.29**	**91.73 ± 0.18**	**94.13 ± 0.16**	**96.98 ± 0.09**	**76.18 ± 0.44**	**86.48 ± 0.28**

Note: Bold values indicate the best performance in each column.

The multi-run results indicate that each enhanced module consistently improves segmentation performance across different random seeds, with relatively small standard deviations suggesting stable behavior. Specifically, extending the sampling radius alone provides a substantial improvement, raising the mIoU by 6.54 percentage points (from 78.22% to 84.76%) and increasing the OA to 94.86%, while exhibiting a low standard deviation (±0.06% for mIoU), indicating stable receptive field expansion. Similarly, integrating the dual attention mechanism independently leads to a 3.26 percentage point improvement in mIoU and a 5.05 percentage point increase in the IoU for interfering objects (from 65.65% to 70.70%), reflecting its effectiveness in modeling complex spatial dependencies.

Although the combinatorial loss function results in a more modest gain—improving the IoU for interfering objects by 0.75 percentage points—its relatively stable variance suggests that it contributes to better handling of class imbalance without introducing additional instability. When all three modules are applied jointly, the proposed full model achieves the best overall performance among the evaluated configurations. It reaches an mIoU of 85.15% (±0.29%) and an OA of 95.13% (±0.13%). The combined configuration yields a 10.53 percentage point absolute improvement in the segmentation of interfering objects compared to the baseline. These findings collectively demonstrate the effectiveness of the proposed framework in improving segmentation accuracy and robustness across different initializations.

3.5. Computational Complexity Analysis

To quantitatively evaluate the computational overhead of the proposed modules and provide transparency regarding model efficiency, a comprehensive complexity analysis was conducted. As detailed in Table 4, the baseline model operates with 1.882 M parameters and an inference latency of 320.24 ms per batch. Notably, the integration of the spatial–geometric dual-attention mechanism is highly efficient: it introduces only 0.217 M additional parameters and adds a mere 12.59 ms to the inference latency (a marginal 3.9% increase), maintaining a lightweight computational profile. Similarly, the combinatorial loss function primarily targets training-phase optimization and imposes virtually no additional memory or latency burden during inference.

Table 4. Quantitative analysis of computational complexity, peak memory footprint, and inference speed for different module combinations. Tests were conducted on a single NVIDIA RTX 4070 Super GPU with a batch size of 1.

Extended Sampling Radius	Combinatorial Loss Function	Dual-Attention Mechanism	Params (M)	Peak GPU Memory (MB)	Inference Latency (ms)	FPS (Hz)
–	–	–	1.882	161.88	320.24	3.12
√	–	–	1.882	312.58	582.84	1.72
–	√	–	1.915	162.01	316.17	3.16
–	–	√	2.099	163.74	332.83	3.00
√	√	√	2.132	314.57	642.49	1.56

Note: √ and – indicate the inclusion and exclusion of the corresponding module, respectively.

However, the experimental data reveals that the inference latency is highly sensitive to the receptive field expansion. Specifically, extending the sampling radius significantly increases the point density and processing volume within each local neighborhood during the grouping phase (e.g., Ball Query), which is computationally intensive and inherently consumes more Graphics Processing Unit memory. Consequently, the independent application of the extended sampling radius drives the latency up to 582.84 ms and increases the peak memory footprint to 312.58 MB. The full model achieves its optimal segmentation performance with a total of 2.132 M parameters, a peak memory footprint of 314.57 MB, and an inference speed of 1.56 Frames Per Second (642.49 ms). While the architectural enhancements introduce an acceptable computational trade-off compared to the simplistic baseline, the overall memory requirements remain exceptionally low and well within the limits of modern edge-computing hardware. Given the substantial improvements in boundary precision and the effective mitigation of severe class imbalance, this trade-off is highly justified for near-real-time deployment in mine tunnel engineering.

3.6. Application Results and Evaluation

To evaluate the practical effectiveness of the proposed framework, it was applied to real-world point cloud data collected from underground mine tunnels using a handheld LiDAR SLAM system. The evaluation focuses on geometric integrity, reconstruction accuracy, and engineering robustness.

3.6.1. Reconstruction Performance and Local Fidelity

Figure 9 illustrates the reconstruction results on a representative tunnel segment. As shown in the 3D visualization (Left panel), the extracted tunnel centerline (red) exhibits high continuity, and the reconstructed cross-sectional profiles (blue) maintain consistent alignment with the wall surfaces even in regions affected by partial occlusions. To further examine the geometric fidelity at a local scale, a detailed error analysis was conducted on

the curved segment indicated by the red dashed circle. The circumferential RMSE heatmap (Right panel) reveals that the fitted profile closely approximates the geometric mean surface of the tunnel wall, achieving a local RMSE of 3.64 cm. This demonstrates that the proposed framework maintains reliable centimeter-level accuracy even in geometrically complex regions with increased curvature and potential SLAM-induced trajectory drift.

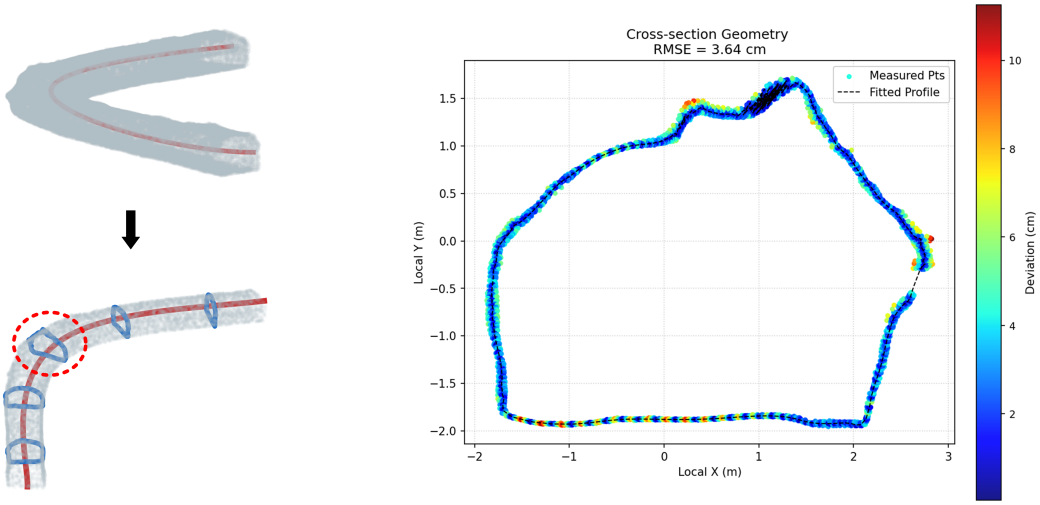

Figure 9. Application results of the proposed framework on real mine tunnel data: (**Left**) 3D view of the reconstructed centerline and profiles. (**Right**) Local geometric accuracy heatmap of the selected curved section, yielding an RMSE of 3.64 cm.

3.6.2. Cross-Sectional Reconstruction Accuracy

To quantitatively assess the geometric fidelity of the reconstructed tunnel profiles, the Root Mean Square Error (RMSE) was computed between the fitted B-spline curves and the corresponding segmented inner-wall point clouds.

All distance calculations were performed within the local 2D sectional plane. To ensure complete data transparency and evaluate the influence of surface noise, we analyzed the performance across a continuous tunnel segment measuring 246.88 meters, from which 246 cross-sections were extracted. The raw average RMSE for the entire segment was 5.05 cm. By applying a 5 cm distance threshold to differentiate the structural backbone from local surface roughness, approximately 29.98% of the points were identified as geometric outliers; however, the resulting overall Filtered RMSE remained remarkably stable at 4.96 ± 0.48 cm.

The minimal difference of only 0.09 cm (0.9 mm) between the raw and filtered RMSE metrics provides definitive evidence of the algorithm's robustness. This confirms that the fitting process successfully converges on the structural centroid of the tunnel without being biased by high-frequency surface noise or structural interference. As visualized in the boxplot in Figure 10, the filtered distribution (Blue) exhibits a more concentrated error profile compared to the raw data (Red), effectively mitigating extreme fitting errors near the 7 cm range. These results demonstrate that the pipeline consistently maintains sub-decimeter precision, satisfying the stringent requirements for practical engineering deformation analysis.

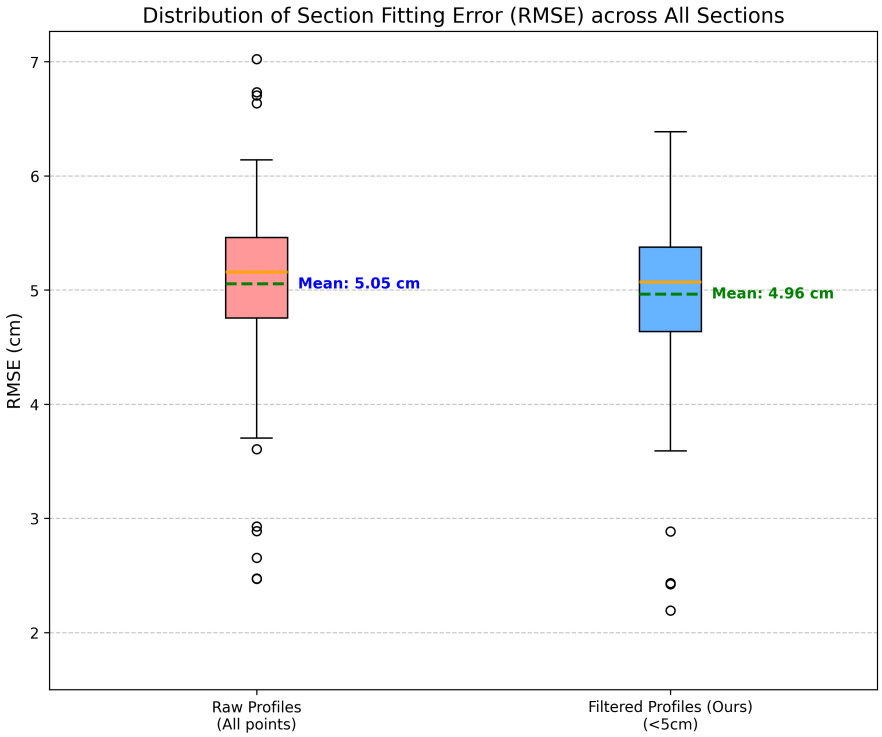

Figure 10. Statistical distribution of the RMSE across 246 reconstructed tunnel cross-sections. The minimal variance and the close alignment (0.09 cm difference) between raw and filtered performance confirm the longitudinal stability and robustness of the modeling pipeline.

4. Conclusions

This study addresses the challenges of geometric anisotropy and severe clutter in narrow mine tunnels by proposing a highly automated framework that integrates handheld laser SLAM with deep learning. By enhancing the PointNet++ architecture with a hierarchical sampling-radius expansion and a spatial–geometric dual-attention mechanism, the proposed method effectively preserves long-range axial continuity and resolves class imbalance issues. Experimental results on real-world datasets demonstrate that the model achieves an mIoU of 85.15% ($\pm$0.29%) and an Overall Accuracy of 95.13% ($\pm$0.13%), significantly outperforming baseline methods in extracting structural inner-wall points.

Building on accurate segmentation, the developed geometric modeling pipeline—utilizing curvature-adaptive sampling and density-adaptive B-spline fitting—successfully transforms discrete point clouds into continuous vectorized cross-sections. The reconstructed profiles achieve a robust overall Filtered RMSE of 4.96 $\pm$ 0.48 cm, satisfying the sub-decimeter accuracy requirements for deformation monitoring. This end-to-end workflow not only replaces labor-intensive manual filtering but also provides a high-precision data foundation for the full-lifecycle digital management and stability assessment of mining infrastructure.

Despite the promising results, a limitation of this study is that the current dataset is derived from 16 tunnels within a single mining project. While the proposed framework effectively captures the common geometric features of typical drill-and-blast tunnels, its direct generalization to tunnels with entirely different support structures (e.g., dense steel arches) or varying excavation profiles (e.g., circular TBM tunnels) may require target-domain fine-tuning. Therefore, evaluating and enhancing the model's generalization capability across diverse mining sites and geological conditions remains an important direction. Additionally, while the extraction of structural inner walls achieved high accuracy, the segmentation performance for interfering objects (mean IoU of 76.18 $\pm$ 0.44%)

indicates room for improvement. The extreme geometric heterogeneity, sparse point distribution, and severe occlusion of elements such as thin cables and irregular ventilation ducts pose significant challenges. To address this bottleneck, future research will explore multi-modal fusion—incorporating LiDAR intensity data or RGB imagery for material cues—and instance-level data augmentation (e.g., 3D Copy-Paste techniques) to enrich the geometric diversity of minority classes. Furthermore, future work will focus on extending the framework to detailed multi-class semantic parsing (e.g., pipelines and support structures) and fusing underground SLAM data with surface InSAR observations to construct a comprehensive "air-space-ground" digital twin for intelligent mining operations.

Author Contributions: Conceptualization, Y.H. and J.Y.; methodology, Y.H.; software, Y.H.; validation, Y.H., X.L. and J.D.; formal analysis, Y.H.; investigation, J.Y.; resources, J.Y.; data curation, X.L.; writing—original draft preparation, Y.H.; writing—review and editing, J.Y.; visualization, J.D.; supervision, J.Y.; project administration, J.Y.; funding acquisition, J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Sichuan Science and Technology Program, grant number 2019YJ0504.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to mine security regulations.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sun, S.; Chen, C.; Yang, B.; Xu, Y.; Zhao, L.; He, Y.; Jin, A.; Li, L. ALM-LED: Autonomous LiDAR mapping in underground space with Luojia explorer anti-collision drone. *ISPRS J. Photogramm. Remote Sens.* **2025**, *230*, 346–373. [CrossRef]
2. Zhang, K.; Kang, L.; Chen, X.; He, M.; Zhu, C.; Li, D. A Review of Intelligent Unmanned Mining Current Situation and Development Trend. *Energies* **2022**, *15*, 513. [CrossRef]
3. Wang, G.; Pang, Y.; Ren, H. Research and Development Path of Smart Mine Technology System. *Metal Mine* **2022**, *51*, 1–9.
4. Ellmann, A.; Kütimets, K.; Varbla, S.; Väli, E.; Kanter, S. Advancements in underground mine surveys by using SLAM-enabled handheld laser scanners. *Surv. Rev.* **2021**, *54*, 363–374. [CrossRef]
5. Liu, M.; Xu, G.; Tang, T.; Qian, X.; Geng, M. Review of SLAM Based on Lidar. *Comput. Eng. Appl.* **2024**, *60*, 1–14.
6. Bolkas, D.; Guthrie, K.; Durrutya, L. sUAS LiDAR and Photogrammetry Evaluation in Various Surfaces for Surveying and Mapping. *J. Surv. Eng.* **2024**, *150*, 4023021.1–4023021.14. [CrossRef]
7. Cui, S.; Bao, J.; Hu, D.; Yuan, X.; Zhang, K.; Yin, Y.; Wang, S.; Zhu, C. Research status and development trends of SLAM technology in autonomous mining field. *Ind. Mine Autom.* **2024**, *50*, 38–52.
8. Stroner, M.; Urban, R.; Kremen, T.; Braun, J.; Michal, O.; Jirikovsky, T. Scanning the underground: Comparison of the accuracies of SLAM and static laser scanners in a mine tunnel. *Measurement* **2025**, *242*, 115875. [CrossRef]
9. Gong, Z.; Li, J.; Luo, Z.; Wen, C.; Wang, C.; Zelek, J. Mapping and semantic modeling of underground parking lots using a backpack LiDAR system. *IEEE Trans. Intell. Transp. Syst.* **2019**, *22*, 734–746. [CrossRef]
10. Cai, N.; Bi, Y.; Pan, K. The application of SLAM laser scanning technology in the completion of the subway tunnel. *Bull. Surv. Mapp.* **2024**, 44–48. [CrossRef]
11. Kang, J.; Li, M.; Mao, S.; Fan, Y.; Wu, Z.; Li, B. A coal mine tunnel deformation detection method using point cloud data. *Sensors* **2024**, *24*, 2299. [CrossRef] [PubMed]
12. Nie, J.; Liu, J.; Li, H. Roadway Section Extraction and Deformation Detection Method Based on 3D Laser Point Cloud. *Coal Technol.* **2024**, *43*, 105–108.
13. Zhao, B.; Ma, B.; He, B.; Xu, S. Deformation Extraction Method for Roadway 3D Laser Point Cloud Data. *J. Harbin Univ. Sci. Technol.* **2024**, *29*, 107–115.
14. Xu, J.; Shan, S.; Yu, R.; Wu, L.; Chen, G. Extraction of curved metro tunnel section based on point cloud. *Bull. Surv. Mapp.* **2023**, 80–84. [CrossRef]

15. Wang, T.; Wei, G.; Wang, Y. Tunnel Point Cloud Section Extraction Method Based on Orthogonal Rotation of the Central Axis. *J. Geo-Inf. Sci.* **2024**, *26*, 2759–2771.
16. Yan, Y.; Zhang, J.; Li, J.; Liu, Q.; Cui, F. A review of Ground filtering algorithms for airborne lidar point cloud data. *Beijing Surv. Mapp.* **2024**, *38*, 1513–1520.
17. Jing, Z.; Guan, H.; Zang, Y.; Ni, H.; Li, D.; Yu, Y. Survey of Point Cloud Semantic Segmentation Based on Deep Learning. *J. Front. Comput. Sci. Technol.* **2021**, *15*, 26.
18. Cui, H.; Li, J.; Mao, Q.; Hu, Q.; Dong, C.; Tao, Y. STSD: A large-scale benchmark for semantic segmentation of subway tunnel point cloud. *Tunn. Undergr. Space Technol.* **2024**, *150*, 105829. [CrossRef]
19. Wu, Z.; Li, M.; Han, Y.; Feng, X. Semantic segmentation of 3D point cloud for sewer defect detection using an integrated global and local deep learning network. *Measurement* **2025**, *253*, 117434. [CrossRef]
20. Qi, C.R.; Li, Y.; Hao, S.; Guibas, L. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5105–5114.
21. Li, Y.; Bu, R.; Sun, M.; Wu, W.; Di, X.; Chen, B. PointCNN: Convolution On *mathcalX*-Transformed Points. *arXiv* **2018**, arXiv:1801.07791.
22. Hu, Q.; Yang, B.; Xie, L.; Rosa, S.; Guo, Y.; Wang, Z.; Trigoni, N.; Markham, A. Randla-net: Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2020; pp. 11108–11117.
23. Wu, X.; Jiang, L.; Wang, P.S.; Liu, Z.; Liu, X.; Qiao, Y.; Ouyang, W.; He, T.; Zhao, H. Point Transformer V3: Simpler, Faster, Stronger. In *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: New York, NY, USA, 2024; pp. 4840–4851.
24. Huang, Y.; Dong, Z.; Li, H.; Chen, Z.; Liu, C.; Zhang, X. Analysis of the applicability of PointNet++ deep learning model for semantic segmentation of typical elements of urban as-built mapping. *Bull. Surv. Mapp.* **2024**, 85–89. [CrossRef]
25. Yue, Y.; Yang, B. Point cloud segmentation and recognition method of coal yards based on PointNet++. *Chin. J. Constr. Mach.* **2023**, *21*, 199–203.
26. Huang, Z.; Gu, X.; Wang, H.; Zhang, X.; Zhang, X. Semantic Segmentation Model for Transmission Tower Point Cloud Based on Improved PointNet++. *Electr. Power* **2023**, *56*, 77–85.
27. Lu, L.; Wang, L.; Wu, S.; Zu, S.; Ai, Y.; Song, B. Semantic Segmentation of Key Categories in Transmission Line Corridor Point Clouds Based on EMAFL-PTv3. *Electronics* **2025**, *14*, 650. [CrossRef]
28. Liu, H.; Tian, S. Deep 3D point cloud classification and segmentation network based on GateNet. *Vis. Comput.* **2024**, *40*, 971–981. [CrossRef]
29. Chen, G.; Gu, C. Enhancing Point Cloud Classification and Segmentation With Attention-Enhanced SO-PointNet++. *IEEE Access* **2024**, *12*, 195986–195995. [CrossRef]
30. Wang, G.; Wang, L.; Wu, S.; Zu, S.; Song, B. Semantic segmentation of transmission corridor 3D point clouds based on CA-PointNet++. *Electronics* **2023**, *12*, 2829. [CrossRef]
31. Wang, X.; Zhang, Y.; Shao, R.; Chen, S.; Li, H.; Chen, X. SAT-Former: An Efficient 3-D Transformer with Semantic Aggregated Point Tokenizer for Point Cloud Semantic Segmentation in Urban Scenes. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 28326–28342. [CrossRef]
32. Peng, X.; Wang, M.; Huang, B.; Shen, H.; Zhong, H. EDA-TCNet: A dual-attention enhanced network for precise point cloud segmentation in tunnel construction. *Undergr. Space* **2025**, *25*, 350–367. [CrossRef]
33. Peng, X.; Wang, M. Deep learning-based point cloud semantic segmentation for tunnel face excavation areas in drilling and blasting tunnels. *Tunn. Undergr. Space Technol.* **2025**, *162*, 106605. [CrossRef]
34. Chen, X.; Chen, L.; Li, M.; Hu, C.; Song, L.; Yuan, P. A method for extracting axis and constructing section in long roadway deformation monitoring. *J. Mine Autom.* **2024**, *50*, 35–41.
35. Zhou, Y.; Chen, H.; Chen, Y.; Ma, Q. Extraction and characterization of irregular tunnel cross-sections based on point clouds. *Measurement* **2025**, *250*, 117065. [CrossRef]
36. Girardeau-Montaut, D. CloudCompare. 2016. Available online: http://www.cloudcompare.org/ (accessed on 15 January 2026).
37. Li, Q.; Cheng, X. Investigating Hand-held Laser Scanner Accuracy. *Bull. Surv. Mapp.* **2016**, *5*, 65–68+96.
38. Lin, T.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*; IEEE: New York, NY, USA, 2017; pp. 2980–2988.
39. Li, X.; Sun, X.; Meng, Y.; Liang, J.; Wu, F.; Li, J. Dice loss for data-imbalanced NLP tasks. *arXiv* **2019**, arXiv:1911.02855.
40. Si, Y.; Xu, H.; Zhu, X.; Zhang, W.; Dong, Y.; Chen, Y.; Li, H. SCSA: Exploring the synergistic effects between spatial and channel attention. *Neurocomputing* **2025**, *634*, 129866. [CrossRef]

41. Lu, Y.; Zhang, Y.; Fan, X.; Cai, D.; Gong, R. Fusing Residual and Cascade Attention Mechanisms in Voxel–RCNN for 3D Object Detection. *Sensors* **2025**, *25*, 5497. [CrossRef]
42. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2018; pp. 7132–7141.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.