buildings

MDPI

Article

PSDNet: A Bio-Inspired Perception-Based Method for Multi-Scale Semantic Segmentation of Non-Uniform Point Clouds of Precast Components

Xinpeng You [1], Xiankai Xue [1], Yongwei Wang [1], Yakun Xie [2,*], Tao Jiang [2], Mingqing Zhu [1], Hao Zhu [1] and Hao Li [1]

[1] CCCC Second Harbor Engineering Co., Ltd., Wuhan 430040, China; youxinpeng@ccccltd.cn (X.Y.); xuexiankai@ccccltd.cn (X.X.); wangyongwei@ccccltd.cn (Y.W.); zhumingqing@ccccltd.cn (M.Z.); zhuhao2@ccccltd.cn (H.Z.); lihao21@ccccltd.cn (H.L.)

[2] Faculty of Geosciences and Engineering, Southwest Jiaotong University, Chengdu 610031, China; jiangtao@my.swjtu.edu.cn

* Correspondence: yakunxie@my.swjtu.edu.cn

Abstract

Three-dimensional laser-scanned point clouds of precast components are often affected by non-uniform density, large object-scale variation, and construction-site noise. Inspired by the dual-pathway perception of eagle vision, this study proposes PSDNet, an end-to-end semantic segmentation network for precast-component point clouds. An Eagle-Focused Density Intensification module (EFDI) enhances panel-edge details in sparsely sampled regions, while a Multi-Scale Global Fusion Module (MSGM) combines a one-time-built spatial index with global dual-pooling attention to distinguish objects ranging from centimeter-scale embedded components to meter-scale fixtures. Experiments were conducted on a dataset collected from a precast-component project in Riyadh, Saudi Arabia. PSDNet achieved a mean intersection-over-union of 73.6%, panel extraction accuracy of 90.9%, and embedded-component identification rate of 91.8%. Ablation experiments confirmed the complementary contributions of EFDI and MSGM. Comparisons with PointNet++, PointCNN, RandLA-Net, and KPConv showed consistently superior segmentation under non-uniform point densities and large-scale variations. The results demonstrate the potential of PSDNet for digital inspection of precast components.

Keywords: precast components; 3D laser-scanned point clouds; bio-inspired perception; multi-scale semantic segmentation

1. Introduction

Driven by China's "Dual-Carbon" strategy and policies promoting new-type industrialization and the large-scale development of prefabricated construction, planar precast components, including precast concrete wall panels, composite slabs, and precast balconies, have become essential building products in residential buildings, public buildings, and transportation infrastructure. Factory-based production of precast components substantially shortens on-site construction periods while reducing pollution associated with cast-in-place construction and the demand for manual labor [1,2]. Although precast production lines are highly efficient, the large number of embedded components and fixtures distributed throughout individual components presents considerable challenges for comprehensive pre-delivery quality inspection [3].

Check for updates

Academic Editor: Vipul Patel

Received: 1 September 2026
Revised: 9 October 2026
Accepted: 14 October 2026
Published: 16 October 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.

Currently, precast factories primarily rely on straightedges and feeler gauges for sampling-based inspection. This inspection approach has several inherent limitations, including its sampling-based nature [4], low inspection efficiency [5], substantial subjective error [6], limited data traceability [7], and considerable occupational safety risks [8]. Consequently, it cannot provide full-coverage quality inspection of entire component surfaces, making it difficult to mitigate potential structural safety risks [9]. As the requirements for standardization and refined quality management in prefabricated construction continue to increase, conventional manual spot-checking systems are no longer adequate for the large-scale and intelligent pre-delivery inspection of precast panels. Owing to its millimeter-level measurement accuracy, non-contact and non-destructive operation, and synchronous acquisition of dense full-field point clouds, 3D laser scanning has become a mainstream technical solution for intelligent geometric inspection in engineering applications [10]. This technology can rapidly capture the complete three-dimensional spatial coordinates of precast panels and reconstruct their comprehensive geometric information, thereby providing a data foundation for efficient digital inspection of entire component surfaces [11–13].

Point-cloud semantic segmentation of precast components enables the effective classification of different structural elements within a component and is therefore a core technology for efficient digital inspection of precast components.

Early semantic segmentation methods for point clouds of engineering components relied primarily on manually designed geometric features and rule-based inference. Rocío et al. developed a clustering algorithm based on geometric rules derived from the geometric characteristics of indoor building structures, enabling the segmentation of floor slabs, ceilings, walls, and other structural elements [14]. Xu et al. proposed two approaches for component segmentation. The first segmented components by registering point clouds against a prior design model, whereas the second used information such as point-cloud density and component dimensions to perform segmentation [15]. Building on prior-model-based point-cloud segmentation, Han-Shuo Zhang et al. established a template library containing standard geometric information for different component types. The segmented point clouds were then compared with the template library using template matching, thereby achieving high-accuracy semantic segmentation [16]. L. Truong-Hong et al. proposed a region-growing segmentation algorithm integrating spatial information and structural knowledge, which enabled the accurate and fully automated extraction of floor slabs, ceilings, walls, columns, primary beams, and secondary beams in concrete buildings [17]. Yang et al. combined contextual prior knowledge of structural elements with a region-growing algorithm to achieve accurate segmentation of similar bridge structures [18]. Dong Wang et al. proposed a two-stage automated point-cloud processing framework incorporating RANSAC-based plane extraction, Gaussian mixture model clustering, and DBSCAN clustering. The framework successfully enabled the automated, low-cost, and high-accuracy dimensional assessment of reinforcement spacing and length in precast bridge components [18]. However, under conditions involving non-uniform point-cloud density, cross-scale objects, viewpoint occlusion, and combinations of multiple component types in precast factories, low-level geometric information is generally insufficient to capture the discriminative characteristics of precast panel components, resulting in limited algorithmic robustness.

The introduction of shallow machine-learning methods improved the feature discrimination capabilities of conventional approaches. Mohamed Mansour et al. optimized multi-scale neighborhood features in point-cloud data and employed a hierarchical support vector machine (hierarchical SVM) algorithm to segment bridge structural elements [19]. Giuseppe Antuono et al. enhanced the conventional random forest (RF) algorithm through hyperparameter calibration and cross-validation and applied it to architectural heritage conservation, achieving accurate classification of building structural elements [20]. Zhao

Xu et al. developed an automated classification and coding system based on building information modeling (BIM) and random forests. The system generated a unique code for each component according to attributes such as type, dimensions, and material recorded in the BIM model, thereby enabling component information traceability [21]. Fundamentally, although shallow-learning algorithms automate the decision-making process, they still depend on manually designed features. In the complex environments of precast factories, designing discriminative features for visually or geometrically similar components is often difficult, substantially increasing the difficulty of segmentation.

In recent years, deep-learning methods have achieved major advances in point-cloud semantic segmentation. PointNet is widely regarded as a pioneering architecture for deep learning on point clouds because it was the first to enable end-to-end learning directly from unordered point sets, although it lacked the ability to model local structures. To address this limitation, PointNet++ introduced hierarchical sampling and local feature-learning mechanisms, thereby improving its adaptability to point clouds with non-uniform densities. PointCNN and KPConv sought to implement convolution-like operations on point clouds from the perspectives of permutation-invariant processing and kernel-point convolution, respectively, thereby improving the representation of complex shapes. For large-scale outdoor scenes, RandLA-Net introduced lightweight random sampling and local feature aggregation modules, achieving a balance between computational efficiency and segmentation accuracy.

In engineering applications, Luping Li et al. proposed AttTransNet, a novel point-cloud semantic segmentation network. Through an innovative attention mechanism and transfer-learning framework, the network effectively addressed the limited generalizability and high training costs of indoor point-cloud semantic segmentation models and enabled automated Scan-to-BIM processing [22]. Liang et al. proposed a complete active-learning-driven weakly supervised segmentation framework. By integrating an active-learning sample-selection strategy based on high loss and high uncertainty with geometric feature embedding and a class-prototype dictionary, the framework maintained high segmentation accuracy while substantially reducing annotation requirements [23]. Vassilev et al. incorporated Bayesian neural networks (BNNs) into the KPConv point-cloud segmentation network to address the lack of predictive confidence estimation and completed component-level segmentation of bridge point clouds. However, Bayesian inference requires multiple forward passes, potentially making inference speed a bottleneck in real-time applications [24]. Lee et al. proposed a semantic segmentation method for unmanned-aerial-vehicle point clouds based on an improved dynamic graph convolutional neural network (DGCNN). By introducing a spatial attention module into DGCNN, the method improved the recognition of slender components in point-cloud semantic segmentation [25]. Nevertheless, its advantages were primarily demonstrated for specific categories such as utility poles, and its adaptability to complex multi-scale scenes requires further verification.

Guo et al. proposed a multitask learning framework that integrated semantic segmentation and 3D bounding-box estimation into a single network, providing an efficient automated solution for the intelligent recognition and virtual assembly of precast components in large and complex scenes. However, the segmentation accuracy for individual components was not reported [26]. To address the limited transferability of point-cloud segmentation models trained on BIM-generated synthetic data, Difeng Hu et al. developed an improved RandLA-Net backbone incorporating an attention-based upsampling module, thereby enhancing robustness to domain shifts. Because the network is sensitive to the accuracy of the BIM model, it is primarily applicable to large indoor structural components and is not well suited to extracting centimeter-scale small components [27]. Wei Lin et al. proposed LiningNet, a 3D deep-learning network designed for the point-cloud segmentation

of shield-tunnel lining segments. By explicitly encoding the inherent geometric regularities of tunnel segments into the network architecture, LiningNet substantially improved segmentation accuracy. However, the method is applicable only to geometrically consistent and regular structures and is difficult to extend to component semantic segmentation under multi-object and class-imbalanced conditions [28].

In summary, existing deep-learning models for point-cloud semantic segmentation have achieved substantial progress in general-purpose scenarios. However, they continue to exhibit clear performance bottlenecks under the complex conditions encountered in precast factories, including non-uniform point-cloud density, complex multi-scale interference, and severe class imbalance. A new network architecture combining density-adaptive perception with multi-scale global fusion is therefore urgently needed to achieve accurate semantic segmentation of precast panels.

Existing studies in biological vision have demonstrated that raptor vision exhibits a distinctive dual-pathway visual perception mechanism that closely matches the dual requirements of precast-component point-cloud processing: fine-grained local defect detection and global recognition of multi-scale objects [29]. Eagle-eye vision has two principal physiological characteristics. First, the shallow fovea provides a wide field of view at relatively low resolution, enabling the simultaneous acquisition of global scene information, rapid differentiation between foreground objects and complex backgrounds, and the suppression of irrelevant interference through global contextual constraints [30]. Second, the deep fovea serves as a high-resolution perception region with an exceptionally high density of retinal photoreceptors, allowing the long-range and high-precision capture of small-object boundaries and subtle surface irregularities and textures, as well as the accurate localization of fine target details in cluttered backgrounds [31–34].

This bio-inspired dual-pathway perception mechanism is highly consistent with the requirements of point-cloud processing for precast panels in factory environments. The deep fovea corresponds to the extraction of fine local features, such as panel edges and embedded components, whereas the shallow fovea corresponds to global semantic discrimination and spatial-relation modeling for objects at different scales, including centimeter-scale embedded components, decimeter-scale electrical boxes, and meter-scale fixtures. However, eagle-eye vision is inherently a two-dimensional imaging and perception system. It relies primarily on imaging by fixed photoreceptors and planar luminance and texture features, lacks the ability to perceive three-dimensional topological structures, and cannot capture variations in spatial point-density distributions. To address these limitations, this study incorporates multidimensional geometric and statistical features of 3D point clouds and develops a dual-branch bio-inspired network adapted to three-dimensional topology and density variation, thereby compensating for the perceptual limitations of two-dimensional biological vision in three-dimensional scenes [35–39].

Although recent studies have advanced point-cloud semantic segmentation [22–28,35–42], a clear research gap remains for precast component inspection in factory environments. First, laser-scanned point clouds exhibit strong near-dense/far-sparse density variation, which degrades local feature learning. Second, precast components contain objects spanning centimeter to meter scales, such as embedded components and fixtures, causing multi-scale interference. Third, severe class imbalance makes small components difficult to recognize. To bridge this gap, this study proposes PSDNet with four objectives: (1) to develop a density-adaptive encoder–decoder network for precast-panel point clouds; (2) to design the EFDI module for density equalization and local geometric enhancement; (3) to develop the MSGM module for multi-scale and global context fusion; and (4) to validate the method on a 3D laser-scanned precast-panel dataset from Riyadh, Saudi Arabia. The

novelty lies in the bio-inspired dual-pathway design that couples eagle-eye deep-fovea local perception with shallow-fovea global awareness in a 3D point-cloud network.

A novel network architecture is proposed in this study, with the following principal contributions:

(1) A novel network, PSDNet, is proposed specifically for point-cloud data acquired through 3D laser scanning in precast factories. The network achieves state-of-the-art (SOTA) performance in precast-panel extraction.
(2) An Eagle-Focused Density Intensification (EFDI) is designed to emulate the high-resolution local-gaze characteristics of the eagle eye's deep fovea. The module primarily addresses variations in point-cloud density and provides high-quality local features for subsequent global contextual modeling and multi-scale feature fusion.
(3) A Multi-Scale Global Fusion Module (MSGM) is developed to emulate the wide-field global scene-awareness characteristics of the eagle eye's shallow fovea. The module is designed to accommodate the substantial variation in object scales within a scene, thereby improving the network's scale adaptability and robustness to global interference.

2. Materials and Methods

2.1. Framework Overview

Given an input point cloud $P = \{p_i \in \mathbb{R}^3\}_{i=1}^{N}$, the network predicts a semantic label $y_i \in \{0, 1, \ldots, 5\}$ for each point, corresponding to panels, fixtures, embedded components, electrical boxes, temporary structures, and corbels. The raw point cloud is first normalized and converted into a unified format before being fed into the encoder.

Inspired by the differentiated perception mechanism formed by the two foveae of the eagle eye, this study develops PSDNet (Precast Panel Segmentation Network), an encoder–decoder architecture designed for 3D laser-scanned point clouds acquired in precast factories. The overall network architecture is illustrated in Figure 1. The encoder incorporates the Eagle-Focused Density Intensification (EFDI) module to emulate the fine-grained perceptual capability of the eagle eye's deep fovea. This module enhances the extraction of edge and surface details from precast panels while adaptively correcting variations in point-cloud density. On this basis, a Multi-Scale Global Fusion Module (MSGM) is introduced to extract multi-scale features through a multi-branch architecture and dilated convolutions and to integrate global contextual information. This design overcomes the limitations of two-dimensional biological vision in scale adaptation and global-relation modeling for 3D point clouds. By balancing local feature precision with global scene awareness, the network enables efficient segmentation of multi-scale objects in precast components.

The overall network adopts an encoder–decoder architecture, with the raw 3D laser-scanned point cloud serving as the input. The EFDI module is integrated into the initial stage of the encoder, while the MSGMs are embedded within the encoder. The decoder output layer comprises six segmentation branches corresponding to panels, fixtures, embedded components, electrical boxes, temporary structures, and corbels. A corbel is a concrete cantilevered supporting component projecting from the side of a column. Feature skip connections are employed throughout the network to preserve shallow local detail features and deep global semantic features, thereby ensuring the accuracy and completeness of panel extraction.

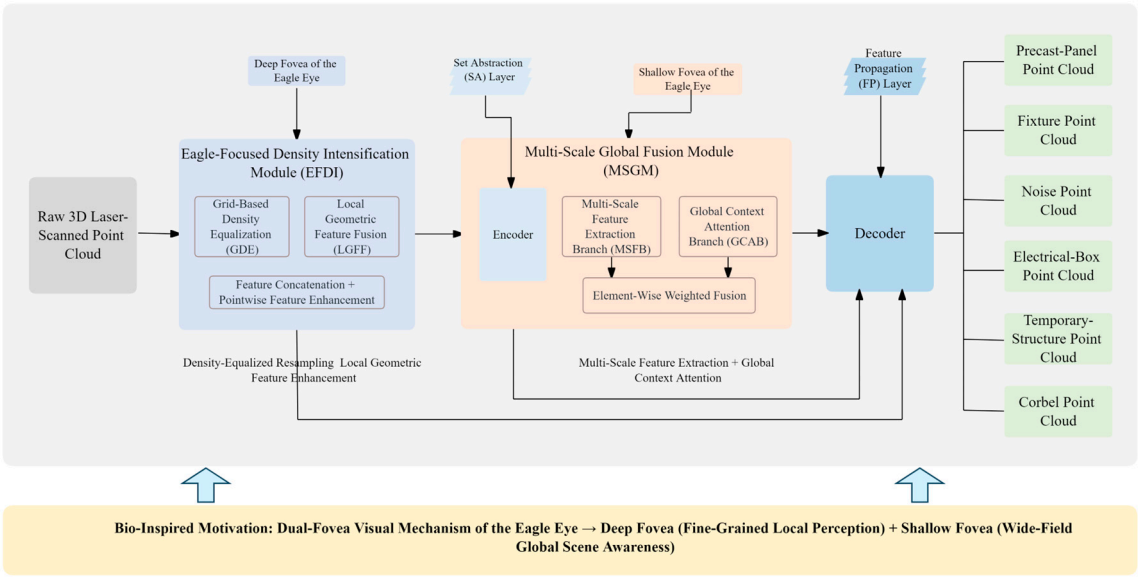


Figure 1. Schematic architecture of PSDNet.

2.2. Eagle-Focused Density Intensification Module (EFDI)

Inspired by the high-resolution fixation mechanism of the eagle eye's deep fovea, this study develops an Eagle-Focused Density Intensification (EFDI) module, whose architecture is shown in Figure 2. The module consists of a Grid-Based Density Equalization (GDE) module and a Local Geometric Feature Fusion (LGFF) module, jointly emulating the eagle eye's precise perception of fine-scale objects. Through a density-adaptive strategy, the EFDI accommodates the non-uniform point-cloud density commonly encountered in precast factory environments. By incorporating point-cloud feature enhancement, it improves the extraction of details from sparsely sampled regions of large planar precast components, thereby providing a reliable basis for subsequent precise classification.

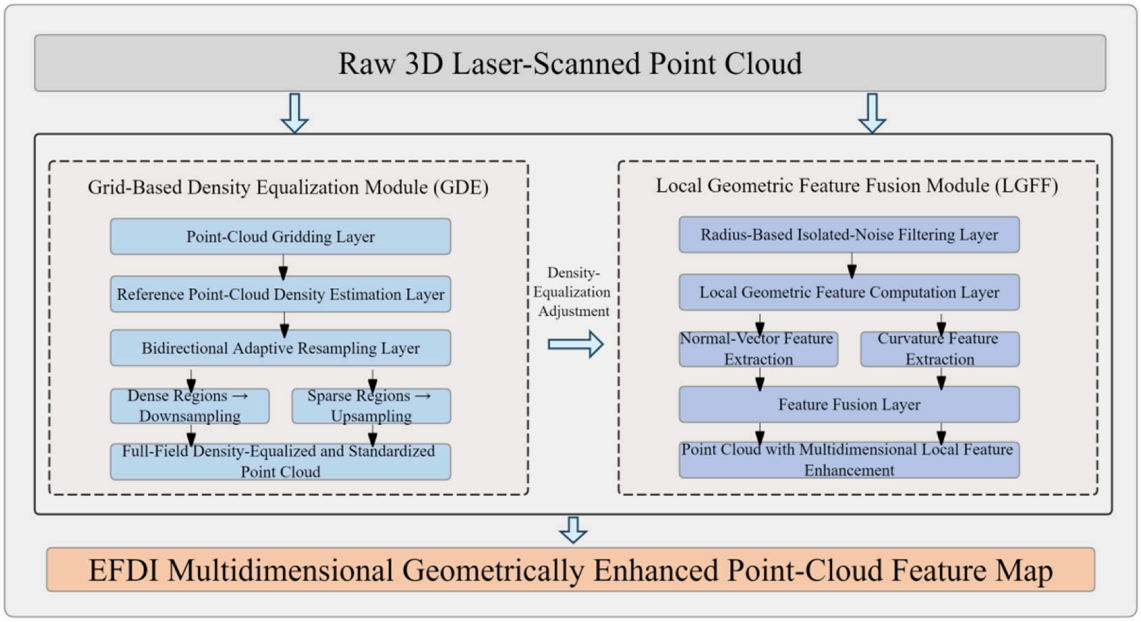


Figure 2. Eagle-Focused Density Intensification (EFDI) module.

Specifically, the Grid-Based Density Equalization (GDE) module processes the input point cloud to accommodate the characteristic density distribution in which points acquired at shorter scanning distances are denser, whereas those acquired at longer distances are

sparser. Given an input point-cloud set $P \in R^{N\times 3}$, where (N) denotes the total number of points and 3 represents the three-dimensional spatial coordinates, the panel point cloud is partitioned into grids. The number of points within each grid is then counted and ranked, and the median point count, k_{avg}, is selected as the reference value. When the number of points in a grid exceeds k_{avg}, farthest point sampling is applied to maintain a consistent point-cloud density. When the number of points is lower than k_{avg}, a novel point-cloud diffusion upsampling algorithm is introduced to balance the point counts across grids. Specifically, the covariance matrix of the point set P_i within grid G_j is first computed to obtain the eigenvalues $\lambda_1 \le \lambda_2 \le \lambda_3$ and the eigenvector n_j corresponding to the minimum eigenvalue. The local tangent plane is then jointly defined by the centroid and the normal vector as follows $\prod$:

$$\prod_j : \left\{ X \in \mathrm{R}^3 \middle| \left(X - \overline{p}_j \right) \cdot \mathrm{n_j} = 0 \right\} \tag{1}$$

The position of the plane is constrained according to the spatial location of the grid, after which uniform sampling is performed on the plane to construct the point set Pavg, containing k_{avg} points. For each point in P_i, its nearest neighbor in Pavg is identified and removed. The remaining points in $\overline{P}_{avg}$ are then merged with P_i to generate a new point-cloud set $\overline{P}$.

In the Local Geometric Feature Fusion (LGFF) module, isolated noise points are first removed using radius-based filtering, as expressed in Equation (2):

$$\mathrm{P_{filter}} = \{ \overline{\mathrm{p}}_\mathrm{i} | \mathrm{k_i} \ge \mathrm{k_{min}} \} \tag{2}$$

where $\mathrm{k_{min}}$ denotes the predefined minimum number of neighboring points. Points with fewer neighboring points than this threshold are identified as isolated noise and removed. For isolated-noise removal, the fixed neighborhood radius is set to 2 cm, and $\mathrm{k_{min}} = 5$.

For each point i, the covariance matrix of its local neighborhood is computed to obtain the eigenvalues $\lambda_1 \le \lambda_2 \le \lambda_3$ and the eigenvector n_i corresponding to the smallest eigenvalue, which is taken as the point normal. Different from the grid-level normal n_j in GDE, n_i here denotes the point-wise normal. The curvature of each point in the point cloud is then calculated as follows:

$$\sigma_\mathrm{i} = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \tag{3}$$

The point normal n_i and curvature σ_i are subsequently fused to obtain the final output of the EFDI, F_{PDM}, as expressed in Equation (4):

$$F_{PDM} = Cat\ (n_i, \sigma_i) \tag{4}$$

where *Cat* denotes a channel-wise concatenation operation used for feature fusion. Here, n_i encodes the local surface orientation, while σ_i describes the local geometric variation; their concatenation provides complementary geometric cues for the subsequent MSGM.

2.3. Multi-Scale Global Fusion Module (MSGM)

To address the limitations of the two-dimensional eagle-eye visual mechanism in adapting to scale variations in 3D point-cloud tasks, as well as the multi-scale interference caused by centimeter-scale embedded components and decimeter- to meter-scale fixtures in precast factory environments, this study proposes a Multi-Scale Global Fusion Module (MSGM). as shown in Figure 3. The module employs multi-scale dilated convolutions to enlarge the effective receptive field and incorporates a global–local feature fusion paradigm to strengthen the interaction between global and local representations, thereby forming an "evolved eagle-eye vision" architecture. Building on the high-resolution local perception capability provided by the EFDI, this architecture substantially enhances the modeling of

global multi-scale contextual information. Experimental results demonstrate that, compared with conventional multi-scale fusion methods, MSGM achieves higher computational efficiency and stronger discrimination between objects of different scales.

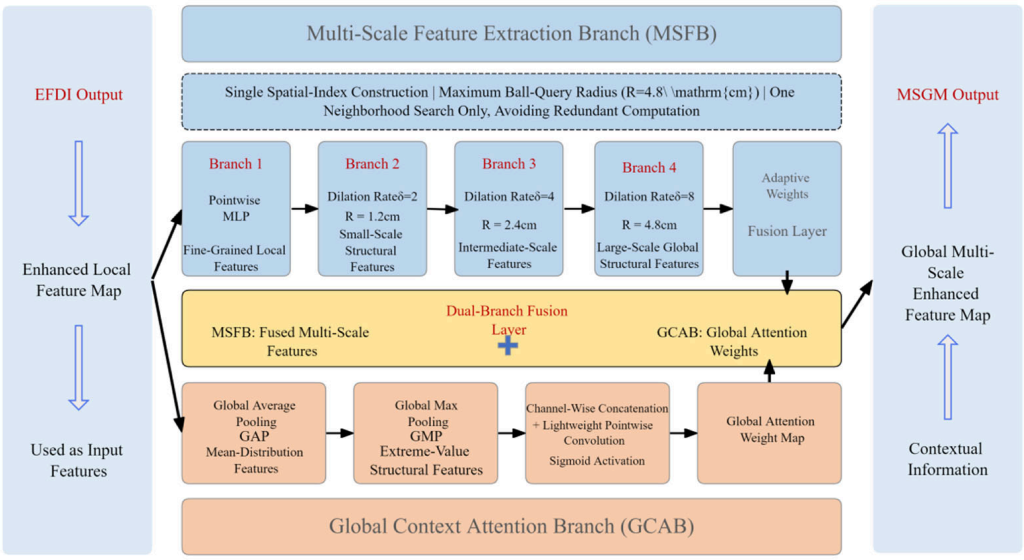


Figure 3. Multi-Scale Global Fusion Module (MSGM).

The Multi-Scale Feature Extraction Branch (MSFB) constructs four parallel feature branches using a ball-query operation combined with a dilation-rate strategy [43,44]. A spatial index with a maximum search radius of 4.8 cm is first constructed, after which multi-scale neighborhoods are generated using different dilation rates. The first branch employs a pointwise multilayer perceptron (MLP) to extract fine-grained local features. The second branch uses a dilation rate of 2 to generate neighborhoods with a radius of 1.2 cm and extracts features of small-scale structures. The third branch uses a dilation rate of 4 to generate neighborhoods with a radius of 2.4 cm and captures features at intermediate scales. The fourth branch uses a dilation rate of 8 to generate neighborhoods with a radius of 4.8 cm and captures large-scale global structural features. This design requires only a single spatial-index construction, thereby avoiding the computational overhead associated with repeated neighborhood searches.

$$\begin{cases} F_1 = \text{PointMLP}(F_{SA}) \\ F_2 = \text{BallQueryConv}(F_{SA}, r = 1.2\text{cm}) \\ F_3 = \text{BallQueryConv}(F_{SA}, r = 2.4\text{cm}) \\ F_4 = \text{BallQueryConv}(F_{SA}, r = 4.8\text{cm}) \end{cases} \tag{5}$$

where F_{SA} denotes the output features of the encoder abstraction layer; PointMLP denotes a pointwise multilayer perceptron; and BallQueryConv denotes a point-cloud convolution operation based on the ball-query mechanism.

To overcome the limited adaptability of scale-specific features caused by fixed-weight fusion, an adaptive scale-weight allocation mechanism is introduced. This mechanism dynamically adjusts the weight assigned to each scale branch according to the distribution characteristics of objects within the scene, thereby achieving optimal fusion of multi-scale features. The fusion process is expressed in Equation (6).

$$F_{ms} = w_1 \cdot F_1 + w_2 \cdot F_2 + w_3 \cdot F_3 + w_4 \cdot F_4 \tag{6}$$

where w_1, w_2, w_3, w_4 are learnable adaptive weights assigned to the four branches and satisfy the constraint $w_1 + w_2 + w_3 + w_4 = 1$. Through iterative training, the model adaptively optimizes these weight parameters, automatically enhancing informative features at relevant object scales while suppressing irrelevant interference features. This process improves the robustness of multi-scale feature extraction in complex and cluttered scenes.

The Multi-Scale Feature Extraction Branch focuses primarily on geometric information within local neighborhoods and therefore lacks semantic constraints derived from the global scene. Consequently, it is susceptible to variations in scanning angle, component occlusion, and interference from local point-cloud geometries, potentially resulting in indistinct precast-panel boundaries and the mis-segmentation of small objects. To address these limitations, a Global Context Attention Branch (GCAB) is developed. This branch employs a dual-pooling architecture to extract global statistical features, model spatial relationships within the point cloud, and enhance salient features.

Specifically, global average pooling (GAP) and global max pooling (GMP) are separately applied to the fused multi-scale feature $\mathrm{F_{ms}}$. These operations simultaneously extract global mean-distribution features and extreme-value structural features, thereby providing a comprehensive representation of the global scene information. The calculation is expressed in Equation (7).

$$\begin{cases} F_{gap} = \mathrm{GAP}(\mathrm{F_{ms}}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{F_{ms}}(\mathrm{i},:) \\ F_{gmp} = \mathrm{GMP}(\mathrm{F_{ms}}) = \underset{i=1}{\overset{N}{max}} F_{ms}(\mathrm{i},:) \end{cases} \tag{7}$$

where F_{gap} denotes the global average-pooled feature, F_{gmp} denotes the global max-pooled feature, and (N) denotes the total number of points at the current layer. The two global feature representations are concatenated along the channel dimension to combine the complementary advantages of global mean and extreme-value features. The concatenated features are then dimensionally reduced using a lightweight pointwise convolution and normalized through a sigmoid activation function to generate a global attention weight map, thereby enabling scene-level semantic modeling. The calculation is expressed in Equation (8).

$$A_{global} = \mathrm{Sigmoid}\left(\mathrm{PointMLP}\left(\mathrm{Cat}\left(F_{gap}, F_{gmp}\right)\right)\right) \tag{8}$$

where A_{global} denotes the global attention weight map, and Cat denotes feature concatenation along the channel dimension. Finally, the global attention weights are fused with the multi-scale feature representation through element-wise weighting, thereby coupling local multi-scale details with global semantic information. The resulting representation constitutes the final output feature of the module, as expressed in Equation (9).

$$F_{MSGM} = F_{ms} \odot A_{global} \tag{9}$$

where F_{MSGM} denotes the final enhanced feature produced by the module, and $\odot$ denotes element-wise multiplication. This output feature compensates for the limitation of the preceding EFDI, which primarily improves point-cloud data quality but lacks global semantic constraints. While preserving cross-scale detail features associated with small components and large fixtures, it strengthens the spatial relationships between the panel and its surrounding environment, thereby substantially improving the model's segmentation accuracy and adaptability to precast factory scenes.

The MSGM adopts a parallel dual-branch architecture comprising the Multi-Scale Feature Extraction Branch (MSFB) and the Global Context Attention Branch (GCAB). Based on a ball-query radius expansion strategy, the MSFB constructs the spatial index only once using the maximum search radius and generates neighborhood features at four scales

through filtering with multiple dilation rates. This design accommodates the extraction of objects across the full range of scales encountered in precast factory scenes, after which the multi-scale features are fused using adaptive weights. The GCAB captures global semantic information through a dual-pooling architecture and generates attention weights for feature enhancement. Finally, the features produced by the two branches are fused to achieve an integrated representation of multi-scale details and global contextual information.

3. Experimental Design

3.1. Dataset

The data were collected at the smart factory for the prefabricated construction of the Sedra Phases 3, 4, and 5 project, a prefabricated housing construction project in Saudi Arabia undertaken by China Harbour Engineering Company Ltd. (Beijing, China). The project adopts a prefabricated wall-panel structural system and mainly comprises precast concrete components such as precast exterior wall panels, precast interior wall panels, and precast parapet walls. Each precast component is composed of units such as fixtures, embedded components, electrical boxes, and temporary structures. In this study, a total of 195 complete point-cloud samples of precast components were collected, containing approximately 5.9×10^7 points. After unified preprocessing, including coordinate normalization and format conversion, the dataset was divided into 155 training samples, 20 validation samples, and 20 test samples. The current dataset covers six principal semantic classes: (0) precast panels, shown in green; (1) fixtures, shown in red; (2) embedded components, shown in blue; (3) electrical boxes, shown in purple; (4) temporary structures, shown in yellow; and (5) corbels, shown in brown.The types and appearances of precast components are illustrated in Figure 4. The color assignments correspond to the manual annotation coloring in Figure 5.

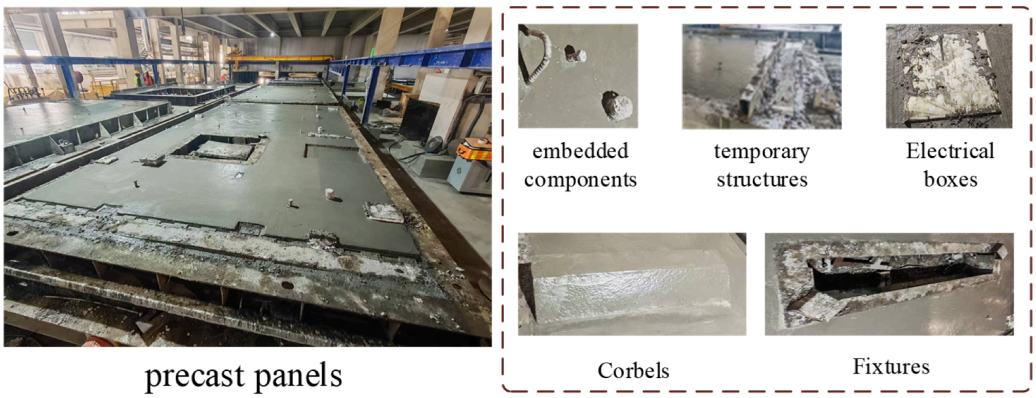


Figure 4. Precast RC panels and multi-scale interference objects in the dataset collected from a precast concrete component project in Riyadh, Saudi Arabia.

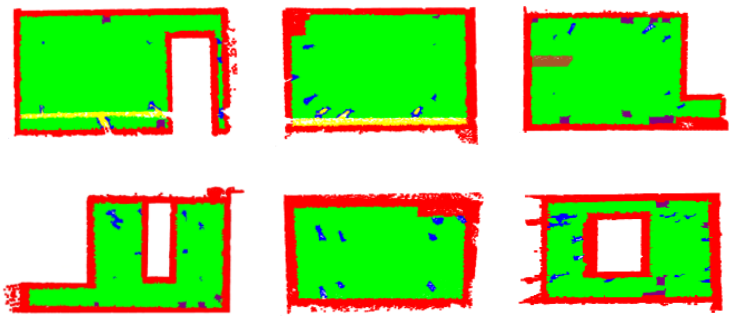


Figure 5. Illustration of the original manually annotated data.

The dataset was acquired using a 3D laser scanner at a project site in Riyadh. The terrestrial laser scanner used was the UCL360PRO series, manufactured by UNRE, Hangzhou, China. Its main parameters are as follows: 16 GB operating memory, laser wavelength of 1550 nm, horizontal field-of-view of 360°, vertical field-of-view of 290°, ranging accuracy of ±1.5 mm, scanning range from 0.6 m to 100 m, and a scanning rate of 500,000 points per second. All precast components were manually annotated. Precast-panel instances account for 59.78% of the annotations, fixtures for 36.52% (decimeter- to meter-scale objects), embedded components for 0.68% (centimeter-scale objects), electrical boxes for 0.88% (decimeter-scale objects), temporary structures for 1.80% (meter-scale objects), and corbels for 0.34% (decimeter-scale objects). The point-cloud density exhibits a pronounced near-dense/far-sparse distribution, with an approximate density of 11 points/cm^2 at short scanning distances and 3 points/cm^2 at long scanning distances. The precast scene therefore faces several core challenges, including non-uniform point-cloud density, multi-scale interference, and diverse combinations of structural types.

3.2. Implementation Details

To ensure the reproducibility and fairness of the experiments, the experimental environment is detailed in Table 1. The hardware configuration consisted of a PC equipped with a 12th-generation Intel® Core™ i9-12900K processor operating at 3.19 GHz, an NVIDIA GeForce RTX 3080 Ti graphics card with 12 GB of video memory for deep-learning acceleration, and 64.0 GB of system memory, of which 63.7 GB was available. This configuration provided sufficient computational capacity for point-cloud preprocessing, model training, and testing.

Table 1. Experimental environment configuration.

Configuration Category	Configuration Parameters
Hardware	Processor: Intel® Core™ i9-12900K, 3.19 GHz GPU: NVIDIA GeForce RTX 3080 Ti with 12 GB of video memory Memory: 64.0 GB of system memory, with 63.7 GB available
Software	Operating system: Windows 10 Pro Programming language: Python 3.12.7 Deep-learning framework: PyTorch 1.11.0 GPU acceleration toolkit: CUDA 11.5 Point-cloud processing library: Open3D 0.18.0

The software environment comprised Windows 10 Pro, version 22H2, with OS build 19045.6456. The deep-learning framework was implemented using PyTorch 1.11.0 in a Python 3.12.7 environment, with CUDA 11.5 employed to leverage GPU-based parallel computing. Open3D 0.18.0 was used for point-cloud loading, preprocessing, and visualization, thereby supporting the efficient and stable execution of the proposed algorithm. The training duration was set to 200 epochs, with a batch size of 8, an initial learning rate of 0.001, and SALoss adopted as the loss function for model optimization.

The model was trained using the Adam optimizer with (beta1 = 0.9, beta2 = 0.999), and the weight-decay coefficient was set to 1×10^{-4}. Training was conducted for 200 epochs with a batch size of 8 and an initial learning rate of 0.001. An exponential learning-rate decay strategy was adopted, whereby the learning rate was multiplied by 0.7 every 10 epochs, i.e., (lr = lr*0.7), with a minimum value of 1×10^{-5}. No early-stopping mechanism was used; instead, all 200 epochs were completed, and the model achieving the highest mean intersection over union (mIoU) on the validation set was retained as the best-performing model.

In addition, the momentum of batch normalization (BN) was dynamically adjusted over the course of training. Its initial value was set to 0.1 and multiplied by 0.5 every 10 epochs, with a lower bound of 0.01. The primary evaluation metric was mIoU, with particular emphasis placed on panel extraction accuracy and embedded-component identification rate. Accurate panel extraction is important because the panel constitutes the final product of precast-component manufacturing and directly supports construction quality management. Moreover, owing to the scanning angle of the laser scanner, diffraction artifacts may occur in the point cloud, causing embedded components such as corrugated ducts and lifting anchors to exhibit distorted geometries. These components may consequently be weakened into noise-like features, thereby affecting the quality assessment of the panels.

4. Experimental Results

4.1. Module Ablation Study

To evaluate the synergistic effects of the Eagle-Focused Density Intensification (EFDI) module and the Multi-Scale Global Fusion Module (MSGM), three progressively enhanced variants were constructed using PointNet++ as the baseline model, and a quantitative ablation study was conducted. In addition to standard semantic segmentation metrics, particular attention was given to precast-panel extraction accuracy and centimeter-scale object recognition under conditions involving non-uniform point-cloud density and multi-scale interference. These evaluations were used to assess the effectiveness of the proposed modules and their synergistic benefits.

As shown in Table 2, incorporating EFDI substantially improved all evaluation metrics. The mIoU increased from 65.9% to 70.6%, while the panel extraction accuracy increased from 78.3% to 88.4%. The embedded-component identification rate also increased from 80.1% to 89.0%. These improvements can be attributed to the adaptive resampling of non-uniform point clouds performed by EFDI. By downsampling densely distributed points acquired at short scanning distances and upsampling sparsely distributed points acquired at longer distances, EFDI effectively mitigates feature loss caused by variations in point-cloud density. It also performs preliminary removal of isolated noise points and enhances local representations using geometric features such as normal vectors and curvature, thereby improving the separability of panels, embedded components, and other structures. Notably, these performance gains require only an additional 0.21 million parameters, imposing minimal additional computational burden while maintaining a favorable balance between accuracy and efficiency.

Table 2. Quantitative results of the ablation study.

Model	mIoU (%)	Panel Extraction Accuracy (%)	Embedded-Component Identification Rate (%)	Additional Parameters
PointNet++	65.9	78.3	80.1	0
PointNet++ + EFDI	70.6	88.4	89.0	+0.21 M
PointNet++ + EFDI + MSGM (Proposed)	73.6	90.9	91.8	+0.43 M

Note: Additional parameters refer to the number of parameters added relative to the baseline model. mIoU denotes the mean intersection over union. Panel extraction accuracy is defined as the proportion of ground-truth panel points correctly identified by the model, and the embedded-component identification rate is defined as the proportion of ground-truth embedded-component points correctly identified by the model. In the following table, OA denotes overall accuracy.

The qualitative results in Figure 6 further demonstrate the effectiveness of EFDI. In Regions 1 and 2, the point clouds are sparse on the left and dense on the right, whereas Region 3 exhibits the opposite distribution. In sparsely sampled regions, the baseline model

tends to misclassify panel points, produce fragmented and discontinuous boundaries, and fail to distinguish the boundaries and spatial relationships between embedded components and panels, resulting in incomplete panel extraction. By contrast, the variant incorporating EFDI generates more continuous and complete panel structures and effectively recovers details in sparsely sampled regions, thereby providing a stronger foundation for subsequent feature fusion and accurate extraction.

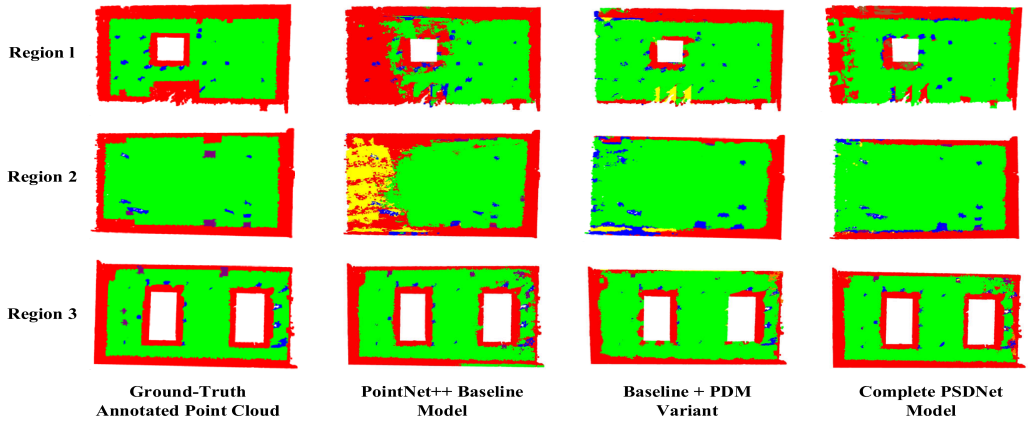


Figure 6. Qualitative visualization of the ablation study.

After MSGM was further integrated to form the complete PSDNet model, all principal metrics improved again. The mIoU reached 73.6%, the panel extraction accuracy reached 90.9%, and the embedded-component identification rate reached 91.8%, while the total parameter increase relative to the baseline remained only 0.43 million. The visualization results in Figure 6 further show that the complete PSDNet model produces clearer and more regular panel boundaries. Background clutter caused by embedded components, fixtures, and scanning-angle effects is identified more effectively, and the distinction between centimeter-scale embedded-component noise and decimeter- to meter-scale fixtures is substantially improved. The continued increase in panel extraction accuracy indicates that MSGM supplements the limited receptive field and scale adaptability of EFDI through multi-scale feature extraction and global contextual enhancement. Consequently, the two modules operate synergistically, preserving fine local panel details while substantially improving the model's adaptability to complex precast factory scenes.

The original point cloud contains precast panels, fixtures, electrical boxes, and other objects and exhibits a near-dense/far-sparse density distribution. In sparsely sampled regions, the baseline model produces fragmented panel contours, irregular boundaries, and misclassified areas. It also fails to distinguish embedded components effectively from panels and fixtures, resulting in poor panel completeness. After EFDI is introduced, panel completeness improves markedly, and features in sparsely sampled regions are effectively recovered; however, the model remains limited in distinguishing multi-scale components from fixtures. The complete PSDNet model produces the clearest and most regular panel boundaries, accurately identifies small components, and more effectively distinguishes components and fixtures at different scales. It achieves the highest panel extraction accuracy and completeness, demonstrating the synergistic benefits of EFDI and MSGM.

4.2. Performance Comparison

To evaluate the overall performance of PSDNet in the semantic segmentation of 3D laser-scanned point clouds of precast components, comparative experiments were conducted against four representative point-cloud segmentation baselines: PointNet++, PointCNN, RandLA-Net, and KPConv. As shown in Table 3, PSDNet outperformed all comparison

models in terms of mIoU, OA, embedded-component identification accuracy, and panel identification accuracy. Compared with KPConv, the strongest baseline, PSDNet achieved improvements of 8.8 percentage points in mIoU, 8.5 percentage points in OA, 11.1 percentage points in embedded-component identification accuracy, and 10.0 percentage points in panel identification accuracy.

Table 3. Quantitative results of the comparative experiments.

Model	mIoU (%)	Panel Identification Accuracy (%)	Embedded-Component Identification Accuracy (%)	OA (%)
PointNet++	65.9	78.3	80.1	81.2
PointCNN	59.8	72.7	67.5	76.4
KPConv	64.8	80.9	80.7	83.6
RandLA-Net	46.2	58.2	57.3	65.7
PSDNet (Proposed)	73.6	90.9	91.8	92.1

These improvements have practical engineering value under the challenging conditions of precast factories, including non-uniform point-cloud density, mixtures of components at different scales, and severe class imbalance. PointNet++, PointCNN, RandLA-Net, and KPConv are constrained by fixed sampling strategies or limited convolutional scales and therefore tend to produce incomplete panels in sparsely sampled regions, inaccurate fixture detection, and confusion between embedded components and noise. In contrast, PSDNet adapts more effectively to non-uniform point clouds and captures the fine geometric characteristics of panels more completely.

The qualitative results further show that PSDNet is particularly effective in recognizing centimeter-scale objects, achieving higher identification accuracy for small components than the four baseline models. Because the conventional networks lack density-adaptive processing and multi-scale global modeling mechanisms, they are more prone to omitted or incorrect segmentation of embedded components, electrical boxes, and other small objects. Through the synergistic operation of EFDI and MSGM, PSDNet improves the recognition of these small-scale targets.

Moreover, PSDNet introduces only a modest increase in the number of parameters relative to PointNet++. It therefore achieves substantial accuracy improvements without imposing excessive computational overhead, balancing segmentation accuracy with the requirements of real-time on-site inspection.The qualitative comparison results are visualized in Figure 7. In Regions 1 and 2, the point clouds are sparse on the left and dense on the right, whereas Region 3 exhibits the opposite density distribution. Although the four baseline models can approximately segment the main panel body, erroneous regions remain on the panel surface, and fixture boundaries are fragmented. Under non-uniform point-cloud distributions, the baseline models generally achieve more accurate segmentation in dense regions but produce fragmented boundaries and class misalignment in sparse regions. This issue is particularly evident for PointCNN and PointNet++. RandLA-Net exhibits extensive discontinuities and missing regions in the panel and fixture contours, together with low embedded-component identification accuracy. Although KPConv largely preserves the main panel structure, its ability to separate small components remains limited, and residual segmentation artifacts are still present. These prediction results demonstrate the advantages of PSDNet in processing non-uniform point clouds, distinguishing components across different scales, and recognizing small objects.

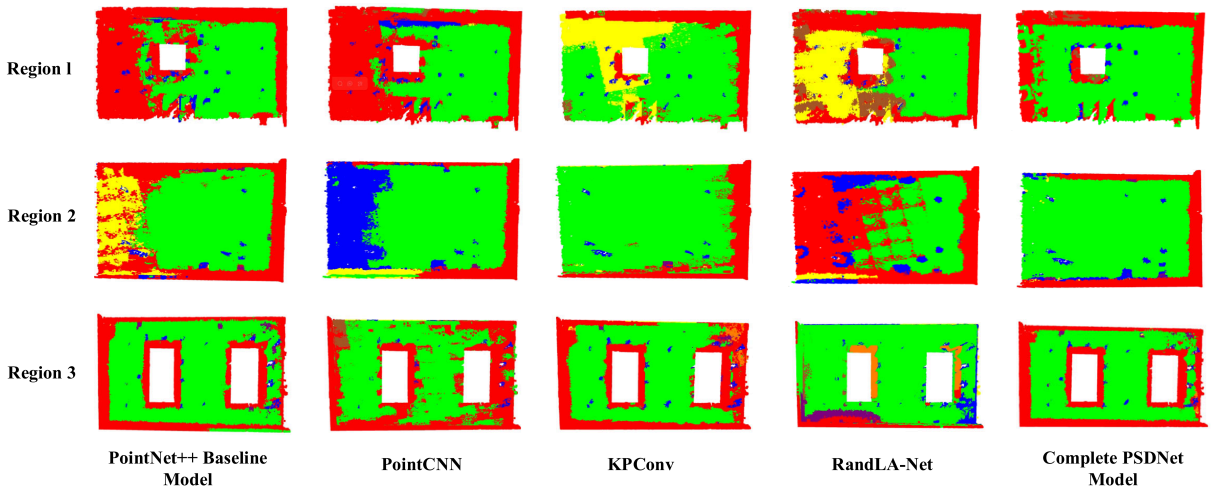


Figure 7. Qualitative visualization of the comparative study.

4.3. Discussion

The superiority of PSDNet can be attributed to the complementary roles of EFDI and MSGM. EFDI mitigates the near-dense/far-sparse density variation by grid-based density equalization and local geometric feature fusion. As shown in the ablation study, EFDI increased mIoU from 65.9% to 70.6% and panel extraction accuracy from 78.3% to 88.4%. MSGM further enhances multi-scale and global context modeling. After jointly incorporating EFDI and MSGM, the full PSDNet achieved 73.6% mIoU, 90.9% panel extraction accuracy, and 91.8% embedded-component identification rate, while adding only 0.43 million parameters relative to PointNet++.

Compared with reported methods for engineering point-cloud segmentation [19,22–29], PSDNet is specifically designed for precast factory scenes with non-uniform density, cross-scale objects, and class imbalance. For example, AttTransNet [22] improves transferability for indoor Scan-to-BIM, LiningNet [28] is tailored to geometrically regular tunnel linings, and BIM-to-scan methods [27] are sensitive to BIM accuracy and less effective for centimeter-scale components. However, because the datasets and annotation protocols differ, direct numerical comparison is not appropriate; the comparison here is mainly qualitative. The main failure cases of PSDNet occur in extremely sparse point clouds and in regions where panels are heavily occluded by multiple layers of fixtures.

Algorithmic rationale. Traditional methods rely on hand-crafted features or fixed sampling/convolution scales, which are sensitive to non-uniform density and multi-scale interference. PSDNet instead combines EFDI, which equalizes density and fuses normals and curvature for local detail (Equations (1)–(4)), with MSGM, which uses a single spatial index, multi-dilation-rate neighborhoods, adaptive weights, and global dual-pooling attention to model multi-scale context (Equations (5)–(9)). This restructuring of feature learning, rather than simply increasing depth, explains its higher accuracy. Ablation confirms: EFDI raises mIoU from 65.9% to 70.6%, and the full model reaches 73.6% with only 0.43 M additional parameters.

5. Conclusions

To address the engineering challenges associated with the non-uniform density distribution and substantial object-scale variation in 3D laser-scanned point clouds acquired in precast factories, this study proposed PSDNet, a semantic segmentation network for precast panels inspired by the dual visual mechanism of eagle vision, which combines fine-grained local fixation with wide-field global scene awareness. Ablation experiments

and comparisons with mainstream algorithms were conducted using the self-constructed PPD on-site point-cloud dataset. The principal conclusions are as follows:

(1) In precast-component scenes characterized by non-uniform point-cloud density and the coexistence of objects at multiple scales, PSDNet achieved the best performance among PointNet++, PointCNN, RandLA-Net, KPConv, and the proposed method in terms of mIoU, panel extraction accuracy, embedded-component identification rate, and overall accuracy (OA). In particular, the panel extraction accuracy reached 90.9%.

(2) The proposed EFDI effectively mitigated the near-dense/far-sparse density distribution inherent in point clouds acquired using 3D laser scanners, thereby compensating for the intrinsic inability of two-dimensional eagle-eye vision to perceive three-dimensional density variations. In addition, curvature and normal-vector feature enhancement was introduced for large planar precast components. With an increase of only 0.21 million parameters, the module substantially reduced incomplete and fragmented panel boundaries under long-range scanning conditions. Incorporating EFDI increased the mIoU by 4.7 percentage points.

(3) The proposed Multi-Scale Global Fusion Module (MSGM) enhanced the network's global contextual modeling capability and compensated for the limited receptive field of EFDI and its insufficient ability to distinguish cross-scale objects. After EFDI and MSGM were jointly incorporated, the model's mIoU increased by a further 3.0 percentage points, while the total number of parameters increased by only 0.43 million relative to the PointNet++ baseline. The resulting model therefore achieved a favorable balance between segmentation accuracy and on-site inference efficiency.

The proposed PSDNet still has several limitations. Its feature-extraction stability remains insufficient in extremely sparse point clouds and in scenes where panels are completely occluded by multiple layers of fixtures. Future research will optimize the density-adaptive regulation strategy of the EFDI and strengthen the ability of MSGM to recover features in heavily occluded regions. Multimodal fusion between PSDNet and building information modeling (BIM) data will also be investigated to further extend the engineering applications of the proposed method to intelligent life-cycle monitoring of precast components and digital construction acceptance.

For practical deployment in precast component factories, it is recommended to maintain a scanning point density of no lower than 3 points/cm^2, supplement sparse regions through multi-view scanning, and perform manual verification in heavily occluded or extremely sparse regions.

Author Contributions: Conceptualization, X.Y. and Y.X.; Methodology, X.Y., Y.X. and T.J.; Software, T.J.; Validation, X.X., T.J. and H.L.; Formal analysis, Y.X. and T.J.; Investigation, X.Y., X.X., Y.W., M.Z. and H.L.; Resources, X.X., Y.W., M.Z., H.Z. and H.L.; Data curation, X.X. and M.Z.; Writing—original draft, X.Y.; Writing—review & editing, X.Y., X.X. and Y.X.; Visualization, Y.W.; Supervision, Y.X. and H.Z.; Project administration, Y.W. and H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Research and Development Project of China Communications Construction Group, grant number JSKF-20250110-002.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: Authors Xinpeng You, Xiankai Xue, Yongwei Wang, Mingqing Zhu, Hao Zhu, and Hao Li were employed by CCCC Second Harbor Engineering Co. Ltd., Wuhan, China. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Kim, M.K.; Sohn, H.; Chang, C.C. Automated dimensional quality assessment of precast concrete panels using terrestrial laser scanning. *Autom. Constr.* **2014**, *45*, 163–177. [CrossRef]
2. Seo, W.; Choi, B.; Shin, D.; Kim, J. Development of a quality management system for precast concrete factories. *J. Civ. Eng. Manag.* **2023**, *29*, 475–486. [CrossRef]
3. Akinci, B.; Boukamp, F.; Gordon, C.; Huber, D.; Lyons, C.; Park, K. A formalism for utilization of sensor systems and integrated project models for active construction quality control. *Autom. Constr.* **2006**, *15*, 411–423. [CrossRef]
4. Skrzypczak, I. Statistical quality inspection methodology in production of precast concrete elements. *Materials* **2023**, *16*, 431. [CrossRef] [PubMed]
5. Li, P.; Wei, Z.Y.; Chen, H.H.; Yan, X.F.; Wei, M.Q. Revisiting tradition and beyond: A customized bilateral filtering framework for point cloud denoising. *ACM Trans. Graph.* **2025**, *44*, 47. [CrossRef]
6. Shu, J.; Li, W.; Zhang, C.; Gao, Y.; Xiang, Y.; Ma, L. Point cloud-based dimensional quality assessment of precast concrete components using deep learning. *J. Build. Eng.* **2023**, *70*, 106391. [CrossRef]
7. Ma, Z.L.; Liu, Y.; Li, J.Y. Review on automated quality inspection of precast concrete components. *Autom. Constr.* **2023**, *150*, 104828. [CrossRef]
8. Helle, R.H.; Lemu, H.G. A case study on use of 3D scanning for reverse engineering and quality control. *Mater. Today Proc.* **2021**, *45*, 5255–5262. [CrossRef]
9. Yang, Z.T.; Tang, D.B.; Li, D.S.; Chen, T.Z.; Liu, J.P.; Qi, H.T.; Wu, Z.; Zhou, J.W. Automatic dimensional quality inspection system for regular precast concrete elements based on 3D structure lighting scan technology. *Autom. Constr.* **2025**, *296*, 104828. [CrossRef]
10. Luo, J.Q.; Yang, Z.X.; Shi, P.C.; Ye, Q. Indoor scan-to-BIM automation: From mobile perception to 3D building modelling. *Autom. Constr.* **2026**, *182*, 106731. [CrossRef]
11. Hong, C.; Nie, S.; Lao, J.Y.; Wang, C.; Wang, K.X.; Yu, J.; Jiang, Y.; Wang, P.; Xi, X.H.; Zhu, X.X. A photon point cloud denoising algorithm based on multi-feature combination and machine learning. *Int. J. Digit. Earth* **2026**, *19*, 2645302. [CrossRef]
12. Ren, H.; Fu, Z.; Zhang, Z.; Ji, B.; Zhu, J.; Zhang, H. Geometric quality inspection of steel structures assisted by point cloud data. *Measurement* **2025**, *250*, 117160. [CrossRef]
13. Li, X.; Song, Z.P.; Zhi, B.; Pu, J.Y.; Meng, C. Intelligent identification of rock mass structural based on point cloud deep learning technology. *Constr. Build. Mater.* **2024**, *456*, 139340. [CrossRef]
14. Romero-Jarén, R.; Arranz, J.J. Automatic segmentation and classification of BIM elements from point clouds. *Autom. Constr.* **2021**, *124*, 103576. [CrossRef]
15. Xu, M.N.; Sun, L.M.; Liu, Y.F.; Li, B.L.; Zhou, M.; Chen, X. Member separation and deformation recognition of spatial grid structures in-service. *Eng. Struct.* **2024**, *304*, 117642. [CrossRef]
16. Zhang, H.S.; Fan, J.S.; Duan, X.J.; Nie, J.G.; Huang, Y.B.; Chai, T.T.; Liu, Y.F. Point cloud segmentation and construction verification for large-span modular steel structures. *J. Constr. Steel Res.* **2025**, *226*, 109288. [CrossRef]
17. Truong-Hong, L.; Lindenbergh, R. Extracting structural components of concrete buildings from laser scanning point clouds from construction sites. *Adv. Eng. Inform.* **2022**, *51*, 101490. [CrossRef]
18. Yang, T.; Zou, Y.; Yang, X.F.; Castillo, E.d.R. Domain knowledge-enhanced region growing framework for semantic segmentation of bridge point clouds. *Autom. Constr.* **2024**, *165*, 105572. [CrossRef]
19. Mansour, M.; Martens, J.; Blankenbach, J. Hierarchical SVM for Semantic Segmentation of 3D Point Clouds for Infrastructure Scenes. *Infrastructures* **2024**, *9*, 83. [CrossRef]
20. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives. *Proceedings of the 2025 Conference on Geo-Digitalization for Sustainable Development Goals (ISDE 2025)*; International Society of Photogrammetry and Remote Sensing (ISPRS): Hannover, Germany, 2026.
21. Xu, Z.; Xie, Z.; Wang, X.R.; Niu, M. Automatic Classification and Coding of Prefabricated Components Using IFC and the Random Forest Algorithm. *Buildings* **2022**, *12*, 688. [CrossRef]
22. Li, L.; Chen, J.; Su, X.; Han, H.; Fan, C. Deep learning network for indoor point cloud semantic segmentation with transferability. *Autom. Constr.* **2024**, *168*, 105806. [CrossRef]
23. Liang, Z.X.; Lai, X.D.; Zhang, L. Active learning-driven semantic segmentation for railway point clouds with limited labels. *Autom. Constr.* **2025**, *171*, 106016. [CrossRef]
24. Vassilev, H.; Laska, M.; Blankenbach, J. Uncertainty-aware point cloud segmentation for infrastructure projects using Bayesian deep learning. *Autom. Constr.* **2024**, *164*, 105419. [CrossRef]
25. Lee, J.S.; Park, J.; Ryu, Y.M. Semantic segmentation of bridge components based on hierarchical point cloud model. *Autom. Constr.* **2021**, *130*, 103847. [CrossRef]
26. Guo, M.; Guo, S.; Meng, Q.L.; Liu, M.; Wang, J.; Cui, C.; Zhang, X. Intelligent virtual trial assembly of prefabricated frame structures for large and complex construction scenes. *Autom. Constr.* **2025**, *172*, 106047. [CrossRef]

27. Hu, D.F.; Dong, Y.; Li, M.K.; Wang, H.; Wang, T. Margin-aware maximum classifier discrepancy for BIM-to-scan semantic segmentation of building point clouds. *Autom. Constr.* **2026**, *183*, 106799. [CrossRef]
28. Lin, W.; Sheil, B.; Zhang, P.; Li, K.; Xie, X. Structural geometry-informed 3D deep learning for segmental tunnel lining analysis in point clouds. *Autom. Constr.* **2025**, *176*, 106281. [CrossRef]
29. Yuen, K.V.; Ye, G. Adaptive feature expansion and fusion model for precast component segmentation. *Comput.-Aided Civ. Infrastruct. Eng.* **2025**, *40*, 3483–3502. [CrossRef]
30. Chen, X.D.; Li, J.; Liu, P.R.; Sun, Q.; Hao, S.W.; Cui, H. An automatic concrete crack-detection method fusing point clouds and images based on improved Otsu's algorithm. *Sensors* **2021**, *21*, 1581. [CrossRef] [PubMed]
31. Miñano, S.; Golodetz, S.; Cavallari, T. Through Hawks' Eyes: Synthetically reconstructing the visual field of a bird in flight. *Int. J. Comput. Vis.* **2023**, *131*, 1497–1531. [CrossRef] [PubMed]
32. Zhang, Z.C.; Chen, J.; Xu, X.Y.; Liu, C.; Han, Y. Hawk-eye-inspired perception algorithm of stereo vision for obtaining orchard 3D point cloud navigation map. *CAAI Trans. Intell. Technol.* **2023**, *8*, 987–1001. [CrossRef]
33. Zhang, Z.C.; Chen, J. Orchard 3D point cloud reconstruction based on binocular eagle-eye vision and super-resolution. *J. Jilin Univ. Eng. Technol. Ed.* **2024**, *54*, 1469–1481.
34. Wang, D.; Gao, Z.; Fang, J.; Li, Y.; Xu, Z. Eagle-YOLOv8: UAV object detection inspired by the eagle-eye vision system. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 9432–9447. [CrossRef]
35. Hu, R.; Cao, F.L.; Wang, W.J. A novel complementary dual-aware network for point cloud classification. *Eng. Appl. Artif. Intell.* **2024**, *138*, 109224. [CrossRef]
36. Sun, L.; Xie, B.; Liu, Y.F.; Shi, H.; Wang, T.C.; Cao, J.L. GeoVLA: Empowering 3D representation in vision-language-action models. *arXiv* **2025**, arXiv:2508.09071.
37. Qu, K.L.; Gao, P.; Dai, Q.; Sun, Y.H. Point-focused attention meets context-scan state space: Robust biological visual perception for point cloud representation. In Proceedings of the International Conference on Learning Representations (ICLR), Rio de Janeiro, Brazil, 23–27 April 2026.
38. Dang, J.S.; Deng, D.L.; Wang, B.M.; Wu, J.; Zhang, H.; Li, H.; Jiao, J.; Pan, D.; Xie, M.; Liu, J. Primary visual cortex inspired point cloud analysis framework. *Proc. AAAI Conf. Artif. Intell.* **2026**, *40*, 3506–3514. [CrossRef]
39. Yu, Z.X.; Qian, X.L.; Gao, S.H.; Jiang, Q.; Gao, Y.; Yang, K.; Wang, K. Seeing clearly and deeply: An RGBD imaging approach with a bio-inspired monocentric design. *arXiv* **2025**, arXiv:2510.25314.
40. Jiang, Z.; Hu, N.; Jiang, Z.; Ma, N. PG-DGNN: A Physics-Guided Dynamic Graph Neural Network for Robust Point Cloud Semantic Segmentation in Prefabricated Construction. In *Proceedings of the 2025 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD), Changchun, China, 26–28 September 2025*; IEEE: Piscataway, NJ, USA, 2025.
41. Yuan, C.; Li, J.; Gao, F.; Wang, Z.; Hu, M. Geometric assessment of precast concrete components using structurally informed semantic point clouds. *Eng. Struct.* **2026**, *366*, 119962. [CrossRef]
42. Li, F.; Zhao, S.; Feng, Z.; Li, Q.; Xu, X. Point cloud data-based edge detection of precast concrete components for dimensional quality assessment using self-attention mechanisms. *Adv. Eng. Inform.* **2025**, *65*, 103383. [CrossRef]
43. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. PointNet++: Deep hierarchical feature learning on point sets in a metric space. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5099–5108.
44. Yang, N.; Wang, Y.; Zhang, L.; Jiang, B. Point cloud semantic segmentation network based on graph convolution and attention mechanism. *Eng. Appl. Artif. Intell.* **2025**, *141*, 109790. [CrossRef]

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.