Proceedings of CONF-CDS 2026 Symposium: Computer Vision-Based Multimodal Cognitive Load Estimation for Adaptive Media Communication
DOI: 10.54254/2755-2721/2026.GL34134

Research on the Technical Architecture and Multi-Scenario Applications of 3D Point Cloud Processing

Sheng Zeng

College of Computer and Artificial Intelligence, Guangzhou Xinhua University, Dongguan, China
1026334914@qq.com

Abstract. In 3D vision, point cloud data functions as a fundamental representation that preserves the geometric structure of objects without loss of spatial information. It plays a key role in tasks like autonomous driving, robotic perception, and 3D reconstruction, and supports understanding and modeling real-world environments. This paper reviews the development of 3D point cloud processing technology. By reviewing the relevant research literature in recent years, it introduces the classification system, key methods, application scenarios, common datasets of point cloud processing, as well as current challenges and future directions. The results demonstrate that current point cloud processing technology has established a complete workflow from low-level data processing to high-level industrial applications, has become the mainstream solution to traditional point cloud challenges, and offers guidance for research, development, and engineering applications in related fields.

Keywords: Point Cloud Processing, Deep Learning, Classification System, Application Scenarios, Challenges and Directions

1. Introduction

In the digital age, 3D vision technology is developing rapidly. Point cloud data, as the fundamental form of 3D vision, describes the surface information of an object through a large number of discrete points, retains the geometric features and spatial structure of the object, and is widely used in fields such as autonomous driving, robot navigation, and virtual reality [1, 2]. With the development of depth sensors and laser scanning technologies, point cloud acquisition has become more convenient, making point cloud processing a research hotspot in computer vision [3, 4]. However, point cloud data is unordered, sparse, unstructured, and often incomplete, limiting the effectiveness of conventional neural networks and processing methods [3, 4]. These characteristics introduce challenges in practical applications, including high computational and memory costs in large-scale urban scenes and unclear segmentation boundaries and degraded small-object performance in sparse environments, limiting real-world deployment [4, 5]. Thus, core tasks in point cloud processing include classification, segmentation, registration, object detection, completion, and denoising [1, 2]. In response to these issues, this paper investigates recent studies on point cloud processing and presents a systematic analysis. It examines the characteristics, advantages, and limitations of existing methods from both task-oriented and technical-route perspectives, followed by an overview

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.GL34134

of application scenarios and evaluation criteria. In addition, it also summarizes current challenges, corresponding solutions, and future development directions in point cloud processing.

2. Classification and technical routes of point cloud processing methods

2.1. Task-based classification

According to task objectives, point cloud processing methods can be systematically categorized into four primary tasks: classification, segmentation, registration, and completion. And this task-oriented taxonomy reflects the different levels of understanding and processing required for 3D point cloud data. Specifically, point cloud classification assigns a single semantic label to an entire point cloud for high-level scene understanding. Meanwhile, point cloud segmentation provides finer-grained labeling at the point level (semantic) or instance level (object differentiation) [1, 2, 4-9]. In addition, point cloud registration aims to estimate spatial transformations that align multiple point clouds into a unified coordinate system. While traditional optimization-based methods are generally stable, they are often sensitive to noise and initial alignment. By contrast, deep learning-based methods improve robustness by learning discriminative features and correspondence matching [10-12]. Furthermore, point cloud completion addresses the challenge of incomplete or partially observed 3D data, aiming to reconstruct complete geometric shapes from sparse inputs. For example, GRNet and AdaPoinTr leverage deep generative models to infer missing structures [13-15].

2.2. Technical-route-based classification

From the perspective of data representation and processing paradigms, point cloud methods can be categorized into three main technical routes: voxel-based methods, point-based direct methods, and graph-based methods. As shown in Figure 1, voxel-based methods convert irregular point clouds into regular 3D grids, hence enabling 3D convolutional neural networks for feature extraction and learning. However, these methods are computationally and memory intensive at high resolutions and may lose fine geometric details. For example, VoxNet and other voxel models work with 3D data using voxels [1, 2, 4].

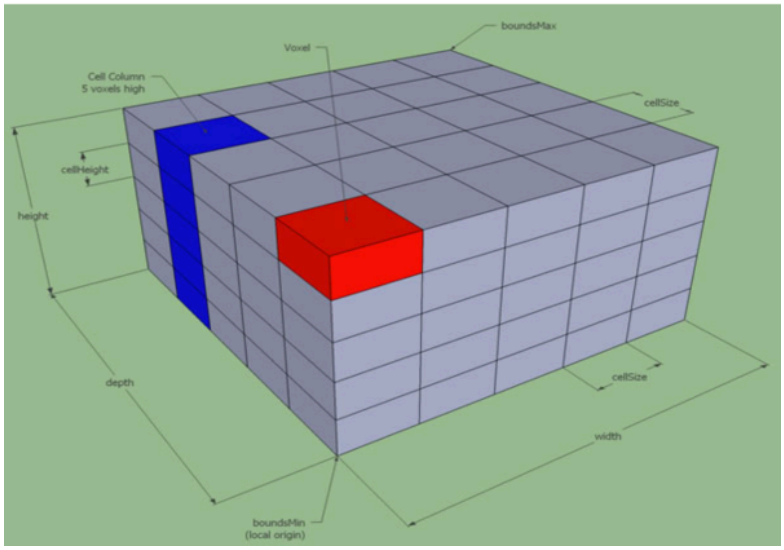

Figure 1. Voxel-based method

DOI: 10.54254/2755-2721/2026.GL34134

In contrast, point-based direct methods operate directly on raw point cloud data without explicit discretization, as shown in Figure 2. These methods preserve geometric information and extract local and global features using PointNet-style and Transformer-based architectures such as PointNet and Point Transformer [2-5, 7, 9].

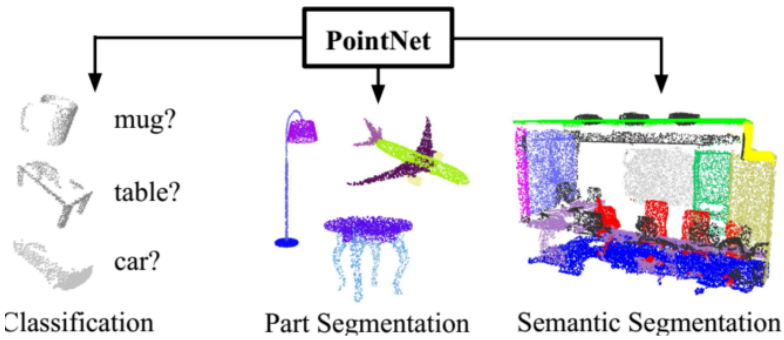

Figure 2. Point-based direct method

Furthermore, graph-based methods represent point clouds as graphs, where points are treated as nodes and spatial relationships define edges, and then apply graph convolutional networks to model geometric dependencies. For example, PointGCN and 3D-GCN improve local structure modeling through message passing [6, 9, 16].

3. Key applications and evaluation criteria of point cloud processing methods

3.1. Intelligent perception and automation applications

In autonomous driving, point cloud data facilitates environmental perception, path planning, and obstacle avoidance by delivering real-time 3D information of the surroundings, including roads, vehicles, and pedestrians, to ensure safe driving [7, 17]. End-to-end LiDAR-based autonomous driving methods further directly map point clouds to vehicle control commands. In robotics, point cloud data supports positioning, navigation, object recognition, and grasping, boosting environmental understanding and task adaptability [1].

To enable these applications, commonly used datasets include nuScenes, KITTI, and ScanNet, as shown in Figure 3. The nuScenes provides multimodal data, including point clouds, images, and vehicle signals, for autonomous driving tasks. In contrast, KITTI focuses on urban driving scenarios and supports object detection, semantic segmentation, and registration. For indoor environments, ScanNet provides large-scale point cloud data with annotations for scene understanding [18]. At the same time, these systems are typically evaluated using detection accuracy, path planning accuracy, real-time performance, and obstacle avoidance success rate. Specifically, detection accuracy is measured by mean average precision (mAP) and intersection over union (IoU). Path planning accuracy uses trajectory error and computational cost. Meanwhile, real-time performance reflects processing speed and latency, while obstacle avoidance success rate measures the system's ability to avoid collisions during operation [17].

DOI: 10.54254/2755-2721/2026.GL34134

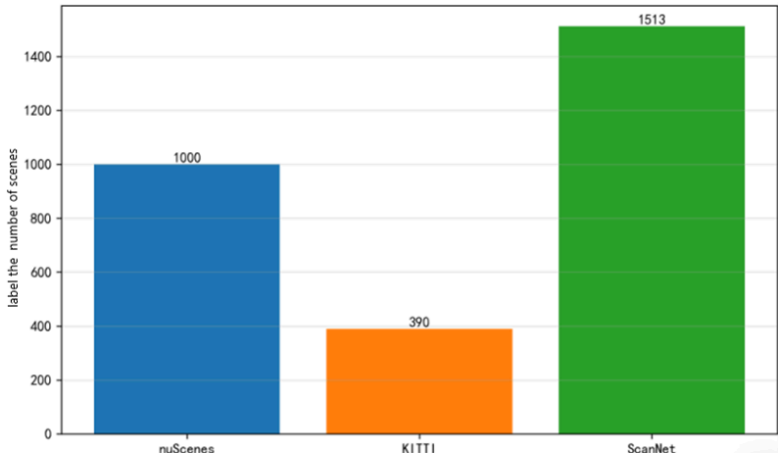

Figure 3. Comparison of the size of mainstream datasets

3.2. Immersive experiences and virtual reconstruction

Point cloud processing becomes a supporting technology in the fields of augmented reality (AR) and virtual reality (VR), modeling in three-dimensional space, handling virtual object interactions, enhancing immersion, and allowing users to experience the virtual world more comfortably [1]. For example, point cloud data can be used to rapidly construct 3D scene models for integrating virtual objects into real environments. And it also supports interactions with virtual objects, such as gesture recognition [15, 18].

To support these tasks, mainstream datasets in immersive experiences and virtual reconstruction include Matterport3D, ScanNet, and 3DIndoor, as shown in Figure 4. Specifically, Matterport3D provides large-scale indoor point cloud and image data for 3D reconstruction and scene analysis. ScanNet contains dense indoor point cloud data with annotations, enabling tasks such as semantic segmentation and object detection. In contrast, 3DIndoor is designed for indoor scene understanding and provides richly annotated point cloud data for various indoor environments [18].

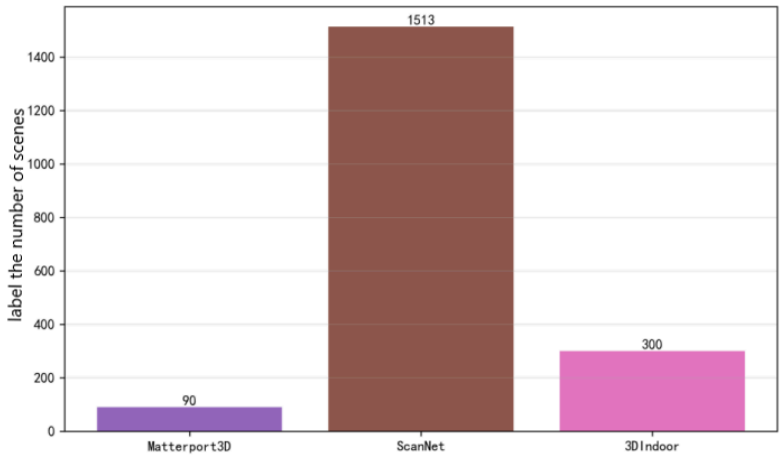

Figure 4. Comparison of indoor reconstruction datasets

In immersive experiences and virtual reconstruction, evaluation focuses on rendering accuracy, real-time performance, interaction accuracy, and user experience. Rendering accuracy is measured by model error and level of detail, while real-time performance depends on processing speed and

DOI: 10.54254/2755-2721/2026.GL34134

rendering frame rate. Interaction accuracy reflects the precision and smoothness of user-object interactions, and user experience is typically assessed through subjective feedback [9, 19, 20].

3.3. Spatial modeling and digital management

In spatial modeling and digital management, point cloud processing plays an important role in smart cities, digital twins, cultural heritage conservation, and geographic information systems (GIS) , as shown in Figure 5. High-precision digital elevation model (DEM) construction from point clouds provides reliable support for terrain analysis and urban planning [21, 22]. In smart cities, point clouds enable urban modeling, resource monitoring, and traffic management, improving operational efficiency. Digital twin applications use point clouds to build virtual representations of real-world environments for simulation and analysis. Cultural heritage conservation relies on point cloud data for 3D reconstruction and digital preservation. In GIS, point clouds support topographic mapping and land use analysis, providing accurate geographic information [1, 2, 22].

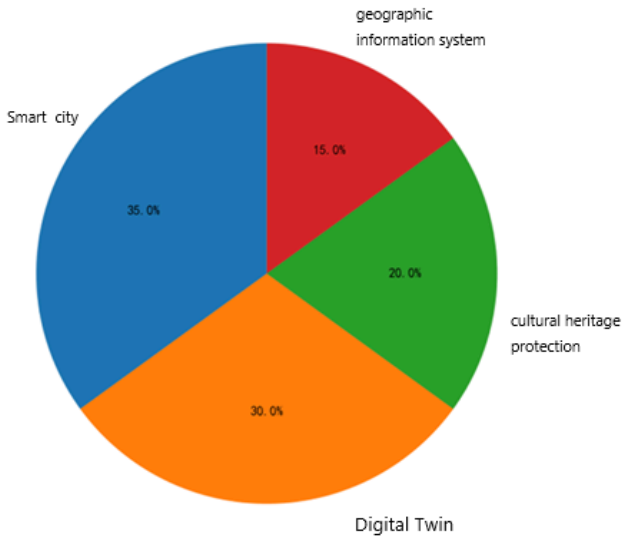

Figure 5. Distribution of application areas

For these applications, common datasets include CityScapes, OpenStreetMap, Cultural Heritage Datasets, and LiDAR data. CityScapes provides urban road scene data for tasks such as semantic segmentation and object detection. OpenStreetMap offers large-scale open geographic information for various GIS applications. And cultural heritage datasets contain annotated point cloud data for preservation and analysis. LiDAR data is widely used in topographic mapping and urban modeling [18]. For evaluation, model accuracy, update frequency, data fusion quality, and reconstruction quality are considered. Specifically, model accuracy reflects error magnitude and detail preservation. Beisdes, update frequency measures how well models capture real-world changes over time. Data fusion quality assesses multi-source integration consistency and accuracy, while reconstruction quality evaluates cultural heritage restoration effectiveness [1].

DOI: 10.54254/2755-2721/2026.GL34134

4. Limitations of existing methods and potential solutions

4.1. Data sparsity and noise occlusion

Point cloud data often suffers from sparsity and uneven distribution. This is because acquisition devices and environmental factors lead to sparse, uneven, or partially missing point clouds, which increases the complexity of subsequent analysis and processing, as shown in Figure 6 [3, 4]. Besides, noise points and object occlusions reduce overall point cloud quality, affecting the accuracy of tasks like classification, segmentation, and registration [10, 12, 22]. In large-scale scenarios, traditional methods typically demand significant computational resources, and the scarcity and high cost of high-quality labeled data further complicate processing [4, 8, 13, 16]. Multimodal fusion combines point clouds with RGB images or depth images to provide richer information, thereby enhancing processing accuracy and robustness [13]. And denoising and point cloud completion algorithms are effective for handling noise and missing data. For example, filtering removes noise points, while completion approaches reconstruct missing regions [10]. For limited labeled data, weakly supervised or unsupervised methods, such as contrastive and self-supervised learning, train models with fewer labels [8].

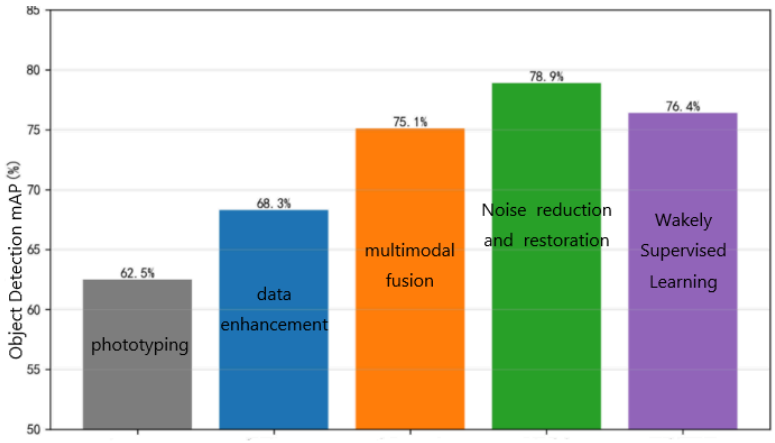

Figure 6. Data sparsity and noise problems

4.2. Segmentation accuracy and registration performance

Point cloud segmentation mainly suffers from class imbalance, small object segmentation, and boundary refinement. Rare classes are often overlooked, while small objects contain sparse points that make feature extraction difficult. At the same time, boundary segmentation requires the model to accurately identify object contours [4, 16, 18]. Similarly, point cloud registration is affected by large initial pose differences, repetitive or symmetrical structures, and partial overlaps. Large pose gaps may cause misregistration, repetitive or symmetric objects can lead to matching errors, and limited overlap reduces reliable correspondences, decreasing registration accuracy [10, 12]. In addition, point cloud completion is challenging due to the difficulty of reconstructing complex topological structures and preserving fine details [13-15].

To address segmentation issues, class imbalance is reduced using oversampling, undersampling, or focal loss. Meanwhile, small object segmentation can be improved using multi-scale feature fusion and attention mechanisms. In addition, boundary refinement can be boosted using dedicated modules that integrate edge detection with contour extraction [18]. For registration, coarse-to-fine

DOI: 10.54254/2755-2721/2026.GL34134

strategies are commonly adopted, where coarse alignment estimates the initial pose followed by fine refinement. Feature matching and geometric constraints help handle repetitive or symmetric structures, while global registration or improved descriptors improve performance under partial overlap [10]. For completion tasks, generative adversarial networks and diffusion models improve structure generation and detail preservation, and multimodal inputs like RGB or depth data further enhance reconstruction quality [13].

5. Conclusion

This paper reviews the development of 3D point cloud processing technology, classifies methods along two dimensions: tasks and technical routes, and explores the characteristics, applications, and challenges of different methods. Point cloud processing has evolved from traditional manual feature extraction to deep learning-based approaches, achieving significant improvements in accuracy and efficiency, and has been widely used in areas like autonomous driving, urban modeling, industrial optimization, and ecological environment monitoring.

However, challenges such as data sparsity and noise occlusion, the trade-off between model accuracy and computational efficiency, and segmentation and registration performance remain unresolved. Future work includes designing more efficient network architectures, reducing reliance on labeled data, further exploring cross-modal fusion, and improving real-time performance. As the technology matures, point cloud processing supports intelligent perception and spatial modeling and is applied in urban management, resource optimization, and environmental monitoring. Its potential to improve urban efficiency, optimize industrial resource use, and enable precise ecological monitoring highlights its relevance to long-term sustainable development. This demonstrates how technological advancement can address societal needs and indicates strong prospects for future intelligent and environmentally conscious growth.

References

[1] Li, J., Sun, H., Dong, Y., Zhang, R., & Sun, X. (2022). Review of 3D point cloud processing based on deep learning. Journal of Computer Research and Development, 59(05), 1160-1179.
[2] Siddiqui, M. Y., & Ahn, H. (2022). Deep learning-based 3D instance and semantic segmentation: A review. Journal on Artificial Intelligence, 4(2), 99–114.
[3] Zhao, J., Zhou, Y., He, X., Bu, Y., Yao, R., & Guo, R. (2022). Analysis of research progress in point cloud segmentation based on deep learning. Journal of Electronics & Information Technology, 44(12), 4426-4440.
[4] Shu, J., Wang, S., Yu, S., & Zhang, J. (2023). CFSA-Net: Efficient large-scale point cloud semantic segmentation based on cross-fusion self-attention. Computers, Materials & Continua, 77(3), 2677-2697.
[5] Li, Y., Chen, F., Lin, Q., Li, Z., Gao, D., & Yang, J. (2025). TFNet: Point cloud semantic segmentation network based on triple feature extraction. Geocarto International, 40(1)
[6] Chen, S. (2025). Research on point cloud classification and segmentation methods based on graph convolutional neural networks [M]. Guilin University of Technology.
[7] Xue, J., Zhao, B., Zhao, C., Li, Y., & Cao, Y. (2025). Classification of point cloud data in road scenes based on PointNet++. Sensors, 26(1), 153.
[8] Huang, H., Bu, Y., Xu, H., & Wang, X. (2025). Point cloud segmentation algorithm based on contrastive learning and label mining. Journal of Computer Research and Development, 62(01), 132–143.
[9] Liu, S., Chi, J., Wu, C., Xu, F., & Yu, X. (2024). SGT-Net: A transformer-based stratified graph convolutional network for 3D point cloud semantic segmentation. Computers, Materials & Continua, 79(3), 4471-4489.
[10] Wan, H., & Wu, X. (2025). A survey of 3D point cloud registration methods. Journal of Image and Graphics, 30(10), 3255–3274.
[11] Li, Z., Hu, X., Fan, W., Guo, B., Pan, F., & Luo, Z. (2025). Research on 3D point cloud registration technology based on SIFT-3DSC and GICP. Acta Metrologica Sinica, 46(06), 830–838.

DOI: 10.54254/2755-2721/2026.GL34134

[12] Zhu, Y., Tian, L., Ye, F., Sun, G., & Fang, X. (2023). Automatic extraction of the sparse prior correspondences for non-rigid point cloud registration. Computer Modeling in Engineering & Sciences, 136(2), 1835-1856.
[13] Hu, F., Li, C., Zhou, T., Cheng, H., & Gu, M. (2025). Survey of deep learning-based 3D point cloud completion algorithms. Journal of Image and Graphics, 30(02), 309–333.
[14] Md, I. S., Kwon, K. C., F. M., F. H., Md, B. H., Shinde, R. K., Gil, S. K., & Kim, N. (2023). Enhanced 3D point cloud reconstruction for light field microscopy using U-Net-based convolutional neural networks. Computer Systems Science and Engineering, 47(3), 2921–2937.
[15] Gankhuyag, U., & Han, J.-H. (2021). Automatic BIM indoor modelling from unstructured point clouds using a convolutional neural network. Intelligent Automation & Soft Computing, 28(1), 133–152.
[16] Xuhan, Yang, H., Shen, Q., Yang, J., Liang, H., Bao, C., & Cang, S. (2020). Automatic terrain debris recognition network based on 3D remote sensing data. Computers, Materials & Continua, 65(1), 579–596.
[17] Zhang, J. (2023). Research on end-to-end autonomous driving technology based on 3D LiDAR (Master's thesis, Nanjing University of Science and Technology).
[18] Qiu, Z., Zhang, L., Yao, Y., Feng, X., & Gao, J. (2025). Advances in 3D point cloud scene semantic segmentation. Journal of Image and Graphics, 30(07), 2325-2342.
[19] Nguyen, T. L., Chu, H. N., Hua, T. T., Phung, T. N., & Le, V. H. (2026). TQU-GraspingObject: 3D common objects detection, recognition, and localization on point cloud for hand grasping in sharing environments. Computers, Materials & Continua, 87(2).
[20] Momynkulov, Z., Tursynova, A., Olzhayev, O., Ikramov, A., Ibrayev, S., & Omarov, B. (2025). Three-dimensional trajectory planning for robotic manipulators using model predictive control and point cloud optimization. Computer Modeling in Engineering & Sciences, 145(1), 891–918.
[21] Zeng, W., Zhang, X., Gu, S., Zhang, R., Yu, L., Zhou, Y., & Zhuo, Y. (2025). A method of urban high-precision DEM construction based on point cloud model. Geocarto International, 40(1).
[22] Chen, B., Feng, J., Zhou, Z., Xu, C., Yan, L., & Xu, L. (2025). Point cloud filtering algorithm of water and soil loss in the disturbed area of the power transmission and transformation project based on CSF. Geocarto International, 40(1).