Industrial3D: A Large-Scale Industrial MEP Point Cloud Dataset and Benchmark for Complex Industrial Scene Understanding

Chao Yin[1,2], Bing Sun[1], Difeng Hu[3], Boyu Wang[4], Mingkai Li[5], Jack C.P. Cheng[1]

[1]Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, China
[2]Guangzhou Institute of Geography, Guangdong Academy of Sciences, Guangzhou, China
[3]Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, China
[4]New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
[5]Department of the Built Environment, National University of Singapore, Singapore

Abstract -

Automated point cloud understanding is critical for modern digital construction workflows, yet existing datasets focus predominantly on architectural elements, leaving complex mechanical, electrical, and plumbing (MEP) systems severely underrepresented. This paper introduces Industrial3D, the first large-scale, high-fidelity dataset specifically designed for industrial MEP scene understanding. Industrial3D comprises over 610 million expertly labeled points across 13 water treatment facilities, captured at 6mm terrestrial LiDAR resolution with 12 specialized MEP and structural classes, representing 4× the scale of existing construction datasets. The dataset exhibits extreme class imbalance (215:1 ratio) and geometric ambiguity where over 86% of tail-class points share cylindrical primitives with the dominant pipe class, creating systematic challenges absent in architectural benchmarks. We evaluate representative methods across learning paradigms, including fully-supervised approaches (KPConv, ResPointNet++, PosPool, and a class-balanced variant) and unsupervised methods (GrowSP). Our evaluation reveals that current state-of-the-art methods face significant limitations in specialized industrial domains, particularly for small, functionally critical components such as valves, strainers, and reducers. Fully-supervised methods achieve approximately 53–54% mIoU, while unsupervised GrowSP reaches only 11.7% mIoU and fails completely on most tail classes. These results point to research opportunities in long-tailed learning and geometry-aware segmentation for industrial scenes. Industrial3D provides a resource for advancing automated construction verification and Scan-to-BIM workflows. Dataset and code will be publicly available at: `https://github.com/pointcloudyc/Industrial3D`.

Keywords -

Point Cloud Dataset; MEP Systems; Semantic Segmentation; Benchmark; Terrestrial Laser Scanning; Industrial Scene Understanding

1 Introduction

Automated point cloud understanding is essential for modern digital construction workflows, including as-built verification, progress monitoring, and automated Building Information Model (BIM) generation [1, 2]. Terrestrial laser scanning (TLS) has become the standard for rapid, high-fidelity 3D data acquisition, capturing billions of geometric measurements of complex mechanical, electrical, and plumbing (MEP) systems. However, a critical bottleneck persists: transforming raw point clouds into actionable semantic understanding remains challenging, particularly for complex industrial environments where MEP systems exhibit intricate geometric relationships, extreme scale variations, and severe occlusion.

Consider a typical water treatment facility: pipes with identical cylindrical geometry serve different functions, small components like valves are embedded within larger structural systems, and dense equipment arrangements create challenging visibility conditions. Current point cloud understanding methods achieve strong results on architectural datasets like S3DIS [3] and ScanNet [4], but struggle with the specialized demands of industrial MEP environments. Existing datasets exhibit three fundamental limitations: (1) *domain mismatch*, focusing on furniture and architectural elements rather than MEP systems; (2) *insufficient scale*, with inadequate point density and scene complexity for fine-grained industrial analysis; and (3) *limited class diversity*, lacking specialized MEP components critical for construction applications.

To address these limitations, we introduce Industrial3D, the first large-scale, high-fidelity dataset specifically designed for industrial MEP scene understanding. Industrial3D comprises over 610 million expertly labeled points across 13 water treatment facilities, captured at 6mm TLS resolution with 12 specialized MEP and structural classes. We also evaluate representative methods across learning paradigms, including fully-supervised approaches (KPConv [5], ResPointNet++ [6], PosPool [7], and a class-balanced variant using CB loss [8]) and the unsupervised method GrowSP [9], providing baseline performance metrics and analysis of domain-specific challenges.

2 Related Work

2.1 Point Cloud Datasets

Large-scale annotated datasets have driven progress in point cloud understanding. S3DIS [3] provides 273 million points across 271 rooms with 13 classes focused on architectural elements and furniture. ScanNet [4] expands to 1,500 indoor scans with 20 classes but maintains similar architectural bias. Semantic3D [10] offers outdoor scenes with 4 billion points but limited class diversity. Construction-specific datasets remain scarce: CLOI [11] includes industrial objects but has limited public availability, while PSNet5 [6] provides a medium-scale MEP-specific publicly available dataset with 80 million points across 5 classes, yet remains limited in scale and class diversity. Table 1 summarizes the comparison.

These datasets share critical limitations for MEP understanding: domain mismatch (architectural focus versus MEP complexity), insufficient resolution for fine-grained components, and absence of specialized classes like valves, flanges, and pumps. Industrial3D addresses these gaps as the first large-scale public dataset targeting MEP understanding, with 610M+ points at 6mm resolution across 12 specialized classes—4× the scale of existing construction datasets.

Table 1. Comparison of Industrial3D with existing point cloud datasets. M = million, B = billion. * indicates limited public availability.

Dataset	Type	Points	Classes	Imbal.
S3DIS	Indoor	273M	13	62:1
ScanNet	Indoor	242M	20	~25:1
Semantic3D	Outdoor	4B	8	75:1
CLOI*	Industrial	140M	10	–
PSNet5	Industrial	80M	5	~50:1
Industrial3D	**Industrial**	**610M**	**12**	**215:1**

2.2 Point Cloud Semantic Segmentation

Deep learning has driven significant progress in point cloud semantic segmentation [12]. PointNet [13] pioneered neural architectures for 3D point clouds, while PointNet++ [14] extended this with hierarchical feature learning. Recent advances include kernel-based convolutions (KPConv [5]), efficient random sampling strategies (RandLA-Net [15]), and local aggregation operators (PosPool [7]).

Construction-specific applications have further explored these architectures. Our prior work ResPointNet++ [6] integrated residual learning into PointNet++ for industrial MEP segmentation with 5 classes on the PSNet5 dataset, and SE-PseudoGrid [16] introduced squeeze-and-excitation attention for piping component classification. Beyond supervised learning, unsupervised methods like GrowSP [9] learn semantic groupings without labels through primitive clustering and superpoint growing. Industrial3D enables systematic evaluation of these representative methods on real-world industrial MEP scenes with 12 classes.

3 Industrial3D Dataset

3.1 Data Acquisition and Annotation

The Industrial3D dataset was captured from 13 large-scale rooms within operational water treatment facilities in Hong Kong. These scenes were selected for their diverse and complex arrangements of MEP and structural components. The total scanned area covers approximately 20,000 m^2, encompassing over 2.3 billion raw 3D points. Table 2 provides statistics of seven representative scenes.

Table 2. Statistics of seven representative scenes in Industrial3D, showing the scale and diversity of industrial environments captured.

Scene	Length (m)	Width (m)	Area (m^2)
Chiller	45.1	19.2	865.92
OGB	86.2	42.5	3,663.50
OSCG	15.3	38.2	584.46
PBF2_1	58.0	22.0	1,276.00
PBF2_2	68.4	38.2	2,612.88
SPH	36.1	26.9	971.09
WRT2	53.3	23.7	1,263.21
Total	-	-	**11,237.06**

High-fidelity point cloud data was acquired using a Leica BLK360 terrestrial laser scanner [17], as shown in Figure 1. This instrument captures 360,000 points per second with 6mm accuracy at 10-meter range, which is critical for capturing small MEP components. Multiple scans were performed within each room from different locations to ensure comprehensive coverage. Raw scan data was processed using Autodesk ReCap software for registration and format conversion.

The annotation process required approximately 754 person-hours to complete, with five domain experts labeling over 610 million points. We defined 12 semantic classes encompassing MEP and structural elements, as detailed in Table 3. Using CloudCompare [18], annotators followed a systematic pipeline: tiling large scenes into manageable chunks, extracting point clusters for each class, merging results, and refining boundaries with outlier removal. Figure 2 illustrates representative instances of each semantic category. Notably, the geometric similarity among pipe-related classes (pipe, elbow, tee, reducer) creates significant geometric ambiguity, where over 86%

(a) (b)

Figure 1. Data acquisition setup for Industrial3D (source from [6]). (a) Partial view of the WRT2 scene showing complex industrial piping systems. (b) Leica BLK360 terrestrial laser scanner during data acquisition with iPad controller for scan registration and preview.

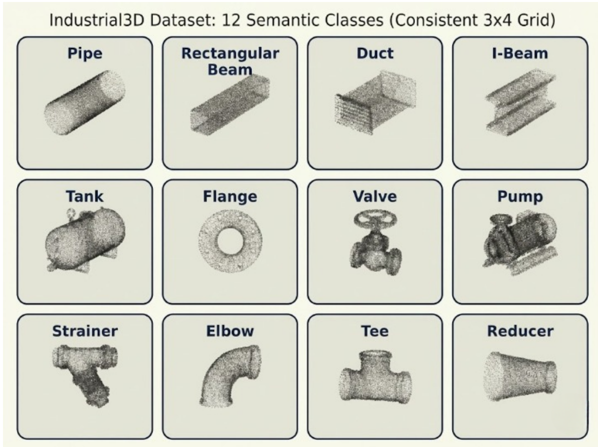

Figure 2. Representative instances of each semantic category in Industrial3D, visualized as synthetic point clouds sampled from BIM models. Note: these are idealized representations; actual LiDAR scans exhibit significant occlusion, noise, and incomplete coverage due to terrestrial scanning geometry.

of tail-class points share cylindrical primitives with the dominant pipe class.

Table 3. The 12 semantic classes in Industrial3D, grouped by frequency. Head classes account for 77% of total points; tail classes each represent less than 3%.

Class	**Description**
Head classes (77% of points)	
Pipe	Cylindrical conduits for fluid transport
RectangularBeam	Rectangular structural support members
Duct	Rectangular air handling conduits
Common classes	
Ibeam	I-shaped structural steel beams
Tank	Large storage vessels
Tail classes (<3% each)	
Flange	Circular connecting components
Valve	Flow control devices
Pump	Mechanical fluid movement devices
Strainer	Filtration devices in pipe systems
Elbow	Angular pipe connectors
Tee	T-shaped pipe junction fittings
Reducer	Pipe diameter transition fittings

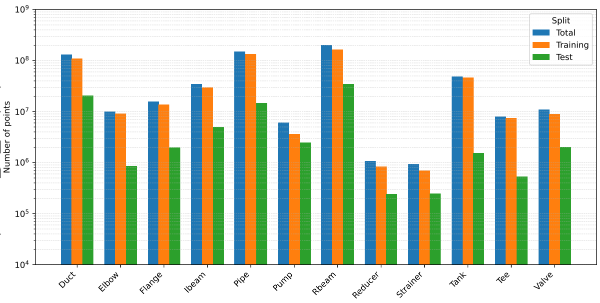

Figure 3. Category-wise point distribution in Industrial3D (logarithmic scale). The severe class imbalance (215:1 ratio) between head classes (Pipe, RectangularBeam, Duct) and tail classes (Reducer, Strainer, Tee) is characteristic of real-world industrial environments.

3.2 Dataset Statistics and Visualization

The final dataset comprises over 610 million expertly labeled points. Each point includes 3D coordinates (X, Y, Z), intensity, and RGB color values. The dataset exhibits significant class imbalance with a 215:1 ratio between the most and least frequent classes—3.5× more severe than S3DIS (62:1). Figure 3 shows the category-wise point distribution on a logarithmic scale. Figure 4 presents the distribution of semantic categories across different industrial areas. Figure 5 provides qualitative visualizations of two representative scenes.

3.3 Dataset Characteristics

Industrial3D differs from existing datasets in several ways: (1) **Scale**, 4× larger than existing industrial datasets with 610M+ labeled points; (2) **Public availability**, as the first large-scale publicly available industrial MEP dataset; (3) **Class diversity**, with 12 specialized MEP and structural classes absent from architectural benchmarks; and (4) **Challenge severity**, the most extreme class imbalance (215:1) among comparable benchmarks, combined with geometric ambiguity where functionally different components share similar cylindrical shapes.

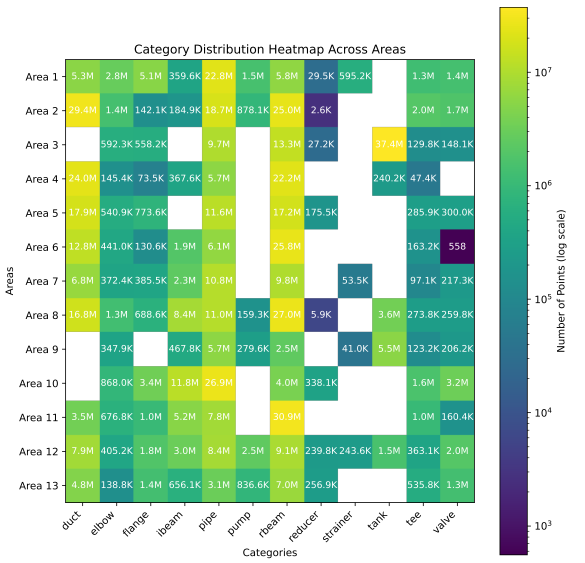

Figure 4. Heatmap showing distribution of semantic categories across industrial areas in Industrial3D. Values are displayed in millions (M) or thousands (K) of points. Structural elements are distributed across all areas, while specialized functional components are concentrated in operation-specific zones.

4 Benchmark Experiments

4.1 Experimental Setup

We partition Industrial3D into a test set (Area 6 and Area 12) and a training set (remaining areas), following the S3DIS area-based evaluation protocol [3]. Area 6 (PBF2_2, Primary Biological Filter) and Area 12 (SPH, Sludge Press House) were selected as test sets because they represent spatially separated facilities with distinct operational characteristics: Area 6 features dense piping arrangements typical of biological filtration systems, while Area 12 contains all diverse mechanical equipment including presses and pumps. This split ensures evaluation on unseen facility types, testing generalization capability across different industrial environments. We evaluate representative methods across two learning paradigms: (1) fully-supervised approaches including ResPointNet++ [6], KPConv [5], PosPool [7], and ResPointNet++ with class-balanced (CB) loss [8]; and (2) the unsupervised method GrowSP [9]. We report overall accuracy (OA), mean Intersection-over-Union (mIoU), and per-class IoU as evaluation metrics.

4.2 Fully-Supervised Results

Table 4 presents benchmark results for fully-supervised methods. All methods achieve similar overall performance (mIoU $\approx$ 53–54%) but exhibit substantial gaps on tail classes. ResPointNet++ achieves 53.31% mIoU

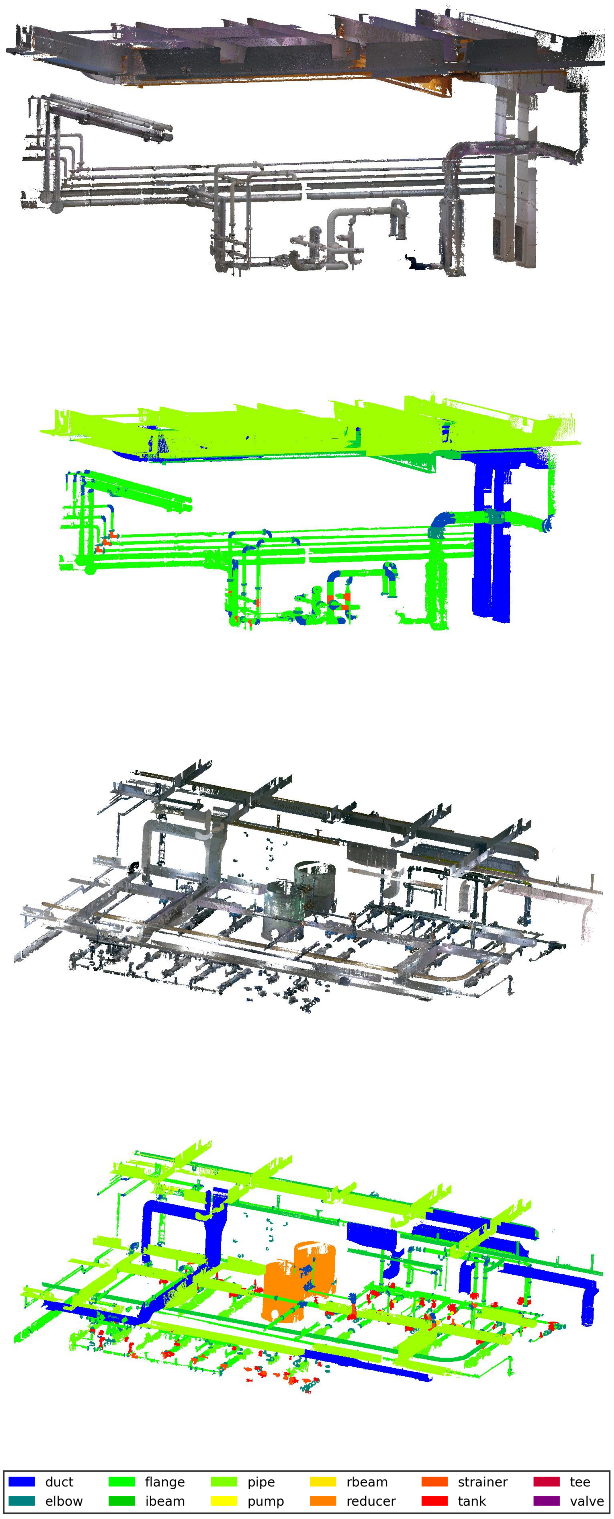

Figure 5. Qualitative visualization of two representative scenes from Industrial3D. Each scene shows: (1) raw point cloud data with RGB color; (2) semantic annotations with class-specific coloring. Rows from top to bottom: Area 6-1 (dense piping region) raw data and annotated data, and Area 12 (Sludge Press House - SPH scene) raw data and annotated

with strong performance on head classes (Duct: 91.36%, Ibeam: 99.41%, RectBeam: 95.19%) but struggles with rare components (Reducer: 0.00%, Strainer: 0.00%). KPConv shows similar patterns with 53.65% mIoU. PosPool achieves non-zero performance on Reducer (34.48%) but comparable overall mIoU at 53.18%. The class-balanced variant (ResPointNet++ w. CB) improves Reducer performance to 14.92% and achieves the highest mIoU at 54.05%, demonstrating that addressing class imbalance provides modest gains.

The results reveal a "dual crisis" in industrial point cloud segmentation: statistical rarity combined with geometric ambiguity. Critical functional components (valves, pumps, strainers) achieve significantly lower IoU than structural elements. The complete failure on Strainer (0.00% for all methods) results from multiple factors: (1) insufficient training samples (Strainer represents only 0.6% of training points), (2) geometric similarity to pipes (both are cylindrical), and (3) small physical size making these components difficult to distinguish from surrounding clutter. For Reducer and Tee, geometric ambiguity is the primary issue, these components share cylindrical primitives with pipes, causing systematic confusion. The class-balanced variant shows modest improvement on Reducer (14.92% IoU versus 0.00%), confirming that frequency-based re-weighting helps but cannot fully resolve geometric ambiguity. For an in-depth treatment including proposed solutions, we refer readers to our long-tailed learning methods paper [19].

4.3 Unsupervised Results

Table 5 presents results for GrowSP [9], which learns semantic groupings through unsupervised primitive clustering and superpoint growing. We evaluate three stages of the GrowSP pipeline: initial primitive extraction, superpoint formation, and final primitive merging to understand performance at each processing level.

GrowSP achieves 84.57% OA and 42.06% mIoU at the superpoint stage, demonstrating reasonable performance on geometrically distinct classes such as Tank (90.61%), RectBeam (86.23%), and Pipe (73.46%). However, performance degrades substantially for fine-grained MEP components: Elbow (4.19%), Reducer (0.00%), and Strainer (0.00%) are nearly or completely unrecognized. The primitive merging stage leads to severe performance collapse (mIoU drops to 11.73%), indicating that the merging heuristics fail to preserve semantic boundaries in complex industrial scenes.

These results demonstrate that while GrowSP can capture coarse geometric structure without supervision, it fundamentally struggles with the fine-grained distinctions required for MEP components. The 11-point mIoU gap between GrowSP (42.06%) and supervised methods (~54%) indicates that domain-specific supervision remains essential for accurate industrial scene understanding.

5 Conclusion and Future Work

We introduced Industrial3D, the first large-scale, high-fidelity point cloud dataset for industrial MEP scene understanding. With over 610 million labeled points across 12 specialized classes from 13 water treatment facilities, Industrial3D provides 4× the scale of existing construction datasets while exhibiting authentic industrial challenges: severe class imbalance (215:1), geometric ambiguity where 86% of tail-class points share primitives with head classes, and complex occlusion patterns. Our benchmark evaluation reveals that fully-supervised methods achieve approximately 53–54% mIoU but struggle significantly with tail classes. The unsupervised GrowSP method achieves only 11.7% mIoU and fails completely on multiple tail classes (Reducer, Strainer, Tee at the merged stage). These results point to fundamental research gaps in handling geometrically ambiguous and statistically rare components.

Addressing these challenges requires multi-faceted approaches. For extreme class imbalance, techniques such as class-balanced loss functions [8], re-sampling strategies, and two-stage training can help models learn from tail classes. However, as our class-balanced experiments show (Reducer improving from 0.00% to 14.92% IoU), re-weighting alone is insufficient due to geometric ambiguity. Effectively distinguishing functionally different but geometrically similar components (e.g., pipes versus reducers) requires geometry-aware segmentation that leverages topological relationships, cylindrical primitive fitting, and multi-modal feature fusion combining geometric (XYZ), appearance (RGB), and intensity information. Industrial3D provides the testbed for developing such specialized approaches.

Future work will expand benchmarking to include weakly-supervised methods (SQN [20], CPCM [21], and WSSS-ST [22]), self-supervised approaches (PointContrast [23], OcCo [24]), and foundation models (PointSAM [25], Reason3D [26]) to evaluate their effectiveness under sparse supervision and zero-shot settings. Beyond semantic segmentation, Industrial3D enables research in instance segmentation, geometric primitive fitting, and topological relationship inference, critical steps toward fully automated Scan-to-BIM workflows. We will release Industrial3D publicly to support research in automated construction verification and digital twin generation.

Acknowledgments

This work was supported by the China Postdoctoral Science Foundation (No. 2023M740761).

Table 4. Benchmark results of fully-supervised methods on the Industrial3D test set. OA = overall accuracy (%), mIoU = mean IoU (%), per-class IoU (%). CB indicates class-balanced weighting loss. Best results per column are in bold.

Method	mIoU	Duct	Elbow	Flange	Ibeam	Pipe	Pump	RectBeam	Reducer	Strainer	Tank	Tee	Valve
ResPointNet++	53.31	91.36	46.05	50.92	**99.41**	76.23	35.95	95.19	0.00	0.00	97.06	12.03	35.56
KPConv	53.65	91.35	**45.44**	50.95	98.35	75.96	42.60	95.04	0.00	0.00	97.72	7.72	**38.69**
PosPool	53.18	90.25	32.93	42.88	98.93	**76.35**	31.27	94.18	**34.48**	0.00	95.59	9.17	32.14
ResPointNet++ w. CB	**54.05**	**91.30**	41.88	**50.71**	99.08	75.69	**43.81**	**95.07**	14.92	0.00	**99.04**	4.68	32.76

Table 5. Benchmark results of the unsupervised method GrowSP on the Industrial3D test set. OA = overall accuracy (%), mIoU = mean IoU (%), per-class IoU (%). The three rows represent different stages of the GrowSP pipeline.

Stage	OA	mIoU	Duct	Elbow	Flange	Ibeam	Pipe	Pump	RectBeam	Reducer	Strainer	Tank	Tee	Valve
Superpoints	84.57	42.06	68.07	4.19	18.64	66.15	73.46	59.14	86.23	0.00	0.00	90.61	12.98	25.18
Primitives	84.00	40.29	67.58	4.19	15.35	66.15	71.70	49.20	85.99	0.00	0.00	90.61	12.98	19.75
Merged Primitives	41.81	11.73	16.37	6.66	6.18	17.56	55.02	0.00	39.00	0.00	0.00	0.00	0.00	0.00

References

[1] Pingbo Tang, Daniel Huber, Burcu Akinci, Robert Lipman, and Alan Lytle. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction*, 19(7):829–843, 2010.

[2] Frédéric Bosché, Mahmoud Ahmed, Yelda Turkan, Carl T Haas, and Ralph Haas. The value of integrating scan-to-bim and scan-vs-bim techniques for construction monitoring using laser scanning and bim: The case of cylindrical mep components. *Automation in Construction*, 49:201–213, 2015.

[3] Iro Armeni, Ozan Sener, Amir R Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, and Silvio Savarese. 3d semantic parsing of large-scale indoor spaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1534–1543, 2016.

[4] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5828–5839, 2017.

[5] Hugues Thomas, Charles R Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, François Goulette, and Leonidas J Guibas. Kpconv: Flexible and deformable convolution for point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6411–6420, 2019.

[6] Chao Yin, Boyu Wang, Vincent JL Gan, Mingzhu Wang, and Jack CP Cheng. Automated semantic segmentation of industrial point clouds using respointnet++. *Automation in Construction*, 130:103874, 2021.

[7] Ze Liu, Han Hu, Yue Cao, Zheng Zhang, and Xin Tong. A closer look at local aggregation operators in point cloud analysis. In *Computer Vision–ECCV 2020*, pages 326–342. Springer, 2020.

[8] Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, and Serge Belongie. Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[9] Zihui Zhang, Bo Yang, Bing Wang, and Bo Li. Growsp: Unsupervised semantic segmentation of 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17619–17629, 2023.

[10] Timo Hackel, Nikolay Savinov, Lubor Ladicky, Jan D Wegner, Konrad Schindler, and Marc Pollefeys. Semantic3d.net: A new large-scale point cloud classification benchmark. *arXiv preprint arXiv:1704.03847*, 2017.

[11] Eva Agapaki and Ioannis Brilakis. Cloi-net: Class segmentation of industrial facilities' point cloud datasets. *Advanced Engineering Informatics*, 45: 101121, 2020.

[12] Yulan Guo, Hanyun Wang, Qingyong Hu, Hao Liu, Li Liu, and Mohammed Bennamoun. Deep learning for 3d point clouds: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(12): 4338–4364, 2020.

[13] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 652–660, 2017.

[14] Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on

point sets in a metric space. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

[15] Qingyong Hu, Bo Yang, Linhai Xie, Stefano Rosa, Yulan Guo, Zhihua Wang, Niki Trigoni, and Andrew Markham. Randla-net: Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11108–11117, 2020.

[16] Chao Yin, Jack CP Cheng, Boyu Wang, and Vincent JL Gan. Automated classification of piping components from 3d lidar point clouds using se-pseudogrid. *Automation in Construction*, 139: 104300, 2022.

[17] Leica Geosystems. Leica blk360 imaging laser scanner. On-line: `https://leica-geosystems.com/products/laser-scanners/scanners/blk360`, 2018.

[18] Daniel Girardeau-Montaut. Cloudcompare: 3d point cloud and mesh processing software. On-line: `https://www.cloudcompare.org/`, 2023.

[19] Chao Yin, Qing Han, Zhiwei Hou, Yue Liu, Anjin Dai, Hongda Hu, Ji Yang, and Wei Yao. Resolving primitive-sharing ambiguity in long-tailed industrial point cloud segmentation via spatial context constraints. *arXiv preprint arXiv:2601.19128*, 2026.

[20] Qingyong Hu, Bo Yang, Guangchi Fang, Yulan Guo, Ales Leonardis, Niki Trigoni, and Andrew Markham. Sqn: Weakly-supervised semantic segmentation of large-scale 3d point clouds. In *Computer Vision–ECCV 2022*, pages 600–619. Springer, 2022.

[21] Lizhao Liu, Zhuangwei Zhuang, Shangxin Huang, Xunlong Xiao, Tianhang Xiang, Cen Chen, Jingdong Wang, and Mingkui Tan. Cpcm: Contextual point cloud modeling for weakly-supervised point cloud semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023.

[22] Chao Yin, Bo Yang, Jack C.P. Cheng, Vincent J.L. Gan, Boyu Wang, and Ji Yang. Label-efficient semantic segmentation of large-scale industrial point clouds using weakly supervised learning. *Automation in Construction*, 148:104757, 2023. ISSN 0926-5805. doi:https://doi.org/10.1016/j.autcon.2023.104757. URL `https://www.sciencedirect.com/science/article/pii/S0926580523000171`.

[23] Saining Xie, Jiatao Gu, Demi Guo, Charles R Qi, Leonidas J Guibas, and Or Litany. Pointcontrast: Unsupervised pre-training for 3d point cloud understanding. In *Computer Vision–ECCV 2020*, pages 574–591. Springer, 2020.

[24] Hanchen Wang, Qi Liu, Xiangyu Yue, Joan Lasenby, and Walterio W Mayol-Cuevas. Unsupervised point cloud pre-training via occlusion completion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9782–9792, 2021.

[25] Yuchen Zhou, Jiayuan Gu, Tung Yen Chiang, Fanbo Xiang, and Hao Su. Point-SAM: Promptable 3d segmentation model for point clouds. In *The Thirteenth International Conference on Learning Representations*, 2025.

[26] Kuan-Chih Huang, Xiangtai Li, Lu Qi, Shuicheng Yan, and Ming-Hsuan Yang. Reason3d: Searching and reasoning 3d segmentation via large language model. In *International Conference on 3D Vision (3DV)*, 2025.