

A review of weakly supervised semantic segmentation for airborne laser scanning point clouds

Ziwei Luo¹, Jun Jiang¹, Yuan Wu¹, Xiaoyun Yan¹, Yangjun Ou¹, Zhaojing Wang¹

¹Wuhan Textile University, Wuhan, China

Airborne Laser Scanning (ALS) point clouds provide accurate three-dimensional geospatial information and are closely related to various remote sensing applications, such as real-scene 3D modeling, infrastructure informatization, and urban mapping. Deep learning has greatly improved ALS semantic segmentation, but fully supervised methods require dense point-wise manual labels, which are expensive to obtain for large-area airborne LiDAR surveys. Under this requirement, weakly supervised ALS segmentation has attracted increased attention because it can use limited manual labels to guide dense semantic interpretation of large area airborne LiDAR point clouds. This paper reviews weakly supervised semantic segmentation methods for ALS point clouds. We first summarize the development of deep learning-based ALS interpretation, commonly used datasets, and evaluation metrics. Existing methods are then reviewed from three main dimensions. First, weak annotation acquisition and sample selection are discussed, including sparse point labels, coarse subcloud-level annotations, click or scribble-based labels, and active learning strategies for selecting informative ALS samples. Second, reliability modeling of weak supervision is analyzed by focusing on pseudo-label confidence, entropy, uncertainty, consistency constraints, and feature-space separability, which are used to identify reliable supervision signals from incomplete or noisy labels. Third, supervision optimization is summarized from the perspectives of self-training, label propagation, pseudo-label refinement, curriculum learning, and robust loss design, aiming to improve the stability and effectiveness of weakly supervised ALS segmentation. Finally, we discuss future directions, including standardized weak-label benchmarks, cross-scene and cross-sensor generalization, promptable foundation models and multimodal remote sensing fusion.

Keywords – Airborne laser scanning, LiDAR point clouds, remote sensing, weak supervision, semantic segmentation.

1. INTRODUCTION

As a representative LiDAR acquisition technique, Airborne Laser Scanning (ALS) has become an important source of large-scale three-dimensional geospatial data [1]. By directly capturing the geometric structure of terrain surfaces, buildings, vegetation, roads, power lines, and other man-made or natural objects, ALS point clouds provide precise elevation and structural information that is difficult to obtain from passive optical imagery alone. Owing to these advantages, ALS data has been widely used in urban mapping [2], land-cover analysis [3], forest monitoring [4], infrastructure management [5], digital twin construction [6], and real-scene 3D modeling [7].

From the perspective of geospatial scene understanding, semantic segmentation plays a central role in ALS point cloud interpretation. Raw ALS point clouds consist of massive, unordered, and irregularly distributed 3D points, which cannot be directly used for high-level geographic analysis. Semantic segmentation aims to assign a semantic category to each point, thereby transforming raw laser measurements into point-wise thematic information. This process provides the basis for subsequent tasks such as object extraction, urban structure analysis, vegetation monitoring, terrain modeling, and infrastructure inventory. With the rapid development of deep learning, point-based networks [8],[9], graph-based networks [10], kernel convolutions [11], efficient sampling models [12], voxel-based pipelines [13], and transformer-style architectures [14],[15] have substantially improved the representation learning capability of ALS point clouds. These methods enable more effective modeling of local geometry, hierarchical context, and long-range semantic relationships, leading to significant progress in fully supervised ALS semantic segmentation.

However, fully supervised ALS semantic segmentation still faces several critical challenges. First, dense point-wise annotation is costly and difficult to scale. ALS scenes usually cover large geographic areas and contain millions or even hundreds of millions of points. Manually assigning accurate labels to each point requires specialized software, professional knowledge, and substantial labor. Second, ALS point clouds exhibit uneven point density, vertical overlap, ambiguous object boundaries, and severe class imbalance. Dominant classes such as ground, buildings, and vegetation often occupy most points, whereas small or thin objects such as power lines, poles, vehicles, and street facilities are difficult to annotate and learn. Third, ALS data collected from different sensors, flight heights, scanning angles, seasons, and geographic regions may show significant distribution shifts. Models trained on one scene or sensor may therefore generalize poorly to unseen cities, land-cover types, or acquisition conditions. These factors make fully supervised learning difficult to deploy in large-scale and long-term ALS applications.

In response to these limitations, Weakly Supervised Learning (WSL) provides a practical alternative to exhaustive point-wise annotation. Following the established taxonomy of incomplete, inexact, and inaccurate supervision [111], this review treats sparse point labels as incomplete supervision, subcloud or segment labels as inexact supervision, and noisy labels as inaccurate supervision. Clicks and scribbles may combine incomplete and inexact characteristics, depending on their coverage and annotation interface. Active learning is treated as a sample-selection strategy, whereas pseudo-labeling is treated as a mechanism for expanding supervision during training. Semi-supervised and active learning studies are included when limited human labels guide dense point-wise prediction. Self-supervised pre-training, few-shot learning, and domain adaptation are included only when they directly support the acquisition, reliability, or transfer of weak supervision. This boundary separates the source of human supervision from the strategies used to select, expand, or transfer it.

Existing surveys provide important but broader perspectives. General 3D reviews summarize architectures and segmentation pipelines [17], [18], while Xiao et al. [19] organize label-efficient learning across indoor, outdoor, and synthetic point clouds. Gao et al. [112] recently connected weak supervision to LiDAR interpretation and parameter inversion across multiple sensing platforms. These studies establish the wider field, but they do not center on point-wise ALS semantic segmentation or compare the full supervision pipeline against ALS-specific label costs, computational evidence, and rare-class behavior. Table 1 positions the present review relative to these representative surveys.

The literature considered in this review centers on methods that use limited, coarse, or imperfect supervision to produce point-wise semantic predictions from ALS point clouds. It covers different forms of human annotation, including sparse points, subcloud or segment labels, clicks, scribbles, and noisy labels, together with sample-selection strategies, reliability modeling, pseudo-label generation, label

propagation, and supervision optimization. Studies from adjacent point-cloud domains are considered only when they provide mechanisms or evidence that can reasonably inform ALS applications, with their original data settings and transfer assumptions stated explicitly. Work restricted to two-dimensional segmentation, object detection, or instance segmentation falls outside the main scope unless it directly supports prompt-based or cross-modal supervision for ALS semantic segmentation.

Against this background, the present review asks how limited supervision can be constructed and used reliably for point-wise ALS semantic segmentation. Existing studies are organized by three interacting functions: annotation acquisition and sample selection, reliability modeling, and supervision optimization. These functions are not mutually exclusive method categories. Instead, they describe where supervision is acquired, assessed, and used, allowing one method to occupy several stages. This framework connects annotation cost, pseudo-label quality, and model training, as illustrated in Figure 1.

Table 1 – Positioning of this review relative to representative surveys.

Review	Primary scope	Weak-supervision focus	ALS semantic-segmentation focus
Guo et al. [17]	General 3D point-cloud learning	Limited	No
He et al. [18]	3D segmentation and multimodal pipelines	Limited	No
Xiao et al. [19]	Label-efficient 3D point-cloud learning	Primary	Partial
Gao et al. [112]	LiDAR interpretation and inversion	Primary	One component
This review	Point-wise ALS semantic segmentation	Primary	Primary; supervision-pipeline lens

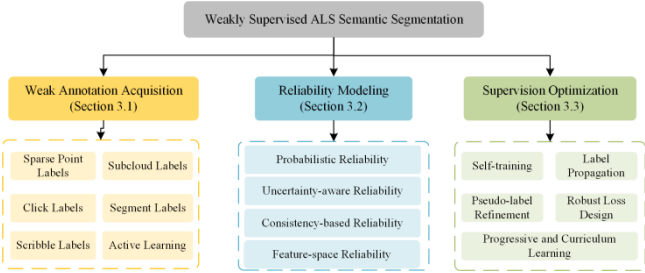


Figure 1 – Functional framework for weakly supervised ALS semantic segmentation. Annotation acquisition and sample selection determine which points or regions receive human supervision. Reliability modeling evaluates or refines human annotations and model-generated labels, while supervision optimization uses these signals to update the segmentation model. The three functions are analytical dimensions rather than mutually exclusive method categories.

The main contributions of this paper are summarized as follows:

1. We provide a critical review of weakly supervised semantic segmentation for ALS point clouds, with particular attention to large airborne scenes, uneven point density, complex vertical structures, and the cost of dense point-wise annotation.
2. We propose a supervision-oriented taxonomy that categorizes existing studies into weak annotation acquisition and sample selection, reliability modeling of weak supervision, and supervision optimization, thereby linking annotation cost reduction, pseudo-label reliability, and robust model training into a unified framework.
3. We summarize representative datasets, evaluation metrics, and benchmark results for ALS and related large-scale point cloud semantic segmentation, and discuss future directions such as standardized weak-label evaluation, cross-domain generalization, foundation-model-assisted pseudo-label generation, and multimodal remote sensing fusion.

The remainder of this paper is organized as follows. Section 2 introduces the preliminaries, including the development of deep learning for ALS point cloud segmentation, the task definition, commonly used datasets, and evaluation metrics. Section 3 reviews weak supervision mechanisms for ALS point cloud semantic segmentation from the perspectives of weak annotation acquisition, reliability modeling, and supervision optimization. Section 4 discusses future research directions, including standardized benchmarks, cross-domain generalization, foundation model-assisted pseudo-label generation, and multimodal remote sensing fusion. Finally, Section 5 concludes this paper.

2. PRELIMINARIES

2.1 The development of deep learning in ALS

Early ALS classification methods mainly relied on hand-crafted geometric, radiometric, and contextual features, combined with classical machine learning methods such as random forests [20],[21], support vector machines [22], conditional random fields, and JointBoost [23]. Although these methods are interpretable, their performance depends heavily on feature design and neighborhood scale selection, which limits their ability to handle large-scale and complex urban scenes.

With the development of deep learning, ALS segmentation has gradually shifted from feature engineering to data-driven representation learning. General 3D point cloud network architectures, including PointNet [8], PointNet++ [9], DGCNN [10], KPConv [11], RandLA-Net [12], and transformer-based models [14],[15],[24], provide effective tools for learning local geometry, hierarchical context, and long-range semantic relationships. In ALS applications, these backbones are often combined with height, intensity, return information, multi-scale neighborhood features, or attention mechanisms to distinguish complex categories such as buildings, ground, vegetation, roads, power lines, and

small man-made objects. Fully supervised deep networks have therefore achieved significant progress on ALS benchmarks.

However, the success of fully supervised methods largely depends on dense point-wise annotations, which are expensive and difficult to obtain for ALS data. To tackle these limitations, weakly supervised learning has emerged as a practical and scalable paradigm for ALS point cloud interpretation. Instead of requiring exhaustive point-level annotations, weakly supervised methods learn from sparse, coarse, or noisy labels. By reducing the dependence on manual labeling and improving the utilization of limited annotations, weak supervision can better support large-scale ALS semantic segmentation under spatial heterogeneity, class imbalance, and cross-scene variation. The methodological evolution of ALS semantic segmentation is summarized in Figure 2.

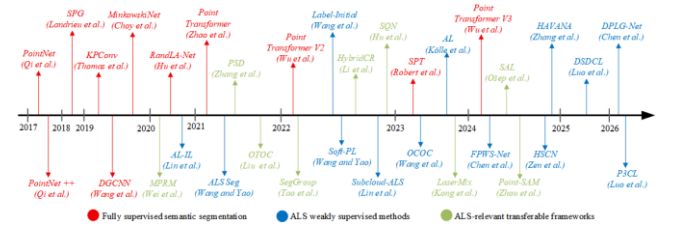


Figure 2 – Timeline of representative developments relevant to ALS semantic segmentation. Red denotes fully supervised segmentation backbones, blue denotes weakly supervised methods evaluated on ALS data, and green denotes candidate mechanisms evaluated primarily in adjacent point-cloud domains. The timeline is illustrative rather than exhaustive and indicates publication sequence, not direct comparability or method superiority.

2.2 Task definition

Given an ALS point cloud $P = \{p_i\}_{i=1}^N$, where each point p_i usually contains 3D coordinates (x_i, y_i, z_i) and may also include auxiliary attributes such as intensity, return number, RGB color, multispectral information, height-normalized features, or local geometric descriptors, semantic segmentation aims to learn a mapping function

$$f_{\theta}: p_i \rightarrow \hat{y}_i, \quad \hat{y}_i \in \{1, 2, \dots, C\},$$

where C denotes the number of semantic classes and $\hat{y}_i$ is the predicted class label of point p_i . In fully supervised learning, each training point is associated with a ground-truth label y_i , and the model is optimized using dense point-wise supervision. The standard objective is usually formulated as a point-wise classification loss, such as cross-entropy loss, optionally combined with class-balanced weights, focal loss, Lovasz-Softmax loss, or boundary-aware losses to handle class imbalance and structural ambiguity.

In weakly supervised ALS semantic segmentation, dense point-wise labels are replaced by limited or imperfect supervision. Human annotations may consist of sparse point labels, coarse subcloud or segment labels, clicks, scribbles, or noisy labels, while the remaining points are unlabeled. Active learning may determine which samples are annotated, but it does not constitute a separate supervision type. The model must therefore learn dense semantic predictions from

the available annotations and unlabeled points. A general training objective can be written as

$$\mathcal{L} = \mathcal{L}_s(P_l, Y_l) + \lambda \mathcal{L}_u(P_u, \tilde{Y}_u) + \gamma \mathcal{R},$$

where P_l denote the labeled points or weakly annotated units, Y_l represents their available human annotations and P_u denotes unlabeled points. The supervised term $\mathcal{L}_s$ learns from the available annotations, whereas $\mathcal{L}_u$ exploits unlabeled points using pseudo-labels or soft predictions $\tilde{Y}_u$. The regularization term $\mathcal{R}$ may impose consistency, entropy, contrastive, or graph-based constraints. The weights λ and γ control the contributions of the unlabeled-data and regularization terms, respectively. Not all methods instantiate every component; either weight may be set to zero when the corresponding term is absent.

2.3 Datasets and evaluation metrics

2.3.1 Datasets

Public benchmark datasets play an important role in developing and evaluating semantic segmentation for ALS and related outdoor point clouds. They differ in acquisition platform, point density, coverage, class definitions, and auxiliary attributes. These differences affect label sparsity, class imbalance, pseudo-label reliability, and cross-scene transfer. Table 2 summarizes representative datasets and shows the contrast between low-density manned-aircraft ALS, high-density UAV data, and multi-platform urban collections.

ISPRS Vaihingen 3D. The ISPRS Vaihingen 3D semantic labeling dataset [2] is one of the most widely used benchmark datasets for ALS point cloud classification. It was acquired over an urban area in Vaihingen, Germany, and contains approximately 1.166 million points with a point density of about 4-7 points/m². The dataset provides point-wise annotations for nine semantic classes, including buildings, low vegetation, impervious surfaces, trees, cars, fences, shrubs, powerlines, and facades. Owing to its moderate size and well-established evaluation protocol, this dataset has been widely used for testing supervised, weakly supervised, and active learning-based ALS segmentation methods.

LASDU. The LASDU [25] is a large-scale aerial LiDAR dataset designed for semantic labeling in dense urban areas. It covers approximately 1 km² and contains more than 3 million points annotated into five object classes. Compared with ISPRS Vaihingen, LASDU represents a more complex high-density urban scene with dense buildings and heterogeneous urban structures. Therefore, it is useful for evaluating whether ALS segmentation models can handle compact urban morphology and class imbalance under limited supervision.

DALES. The DALES [26] is a large-scale aerial LiDAR dataset for semantic segmentation. It contains over a half-billion manually labeled points covering approximately 10 km² and includes eight object categories. DALES was designed to provide a large public ALS benchmark for evaluating deep learning methods on aerial LiDAR scenes.

H3D. The H3D [27] is a high-resolution UAV LiDAR and multi-view stereo dataset for semantic segmentation. It depicts the village of Hessigheim in Germany and provides dense point clouds with a mean point density of about 800 points/m². The dataset includes manually labeled point clouds with at least 11 semantic classes and also provides textured 3D meshes derived from oblique imagery. Compared with conventional ALS datasets, H3D has much higher point density and richer geometric details, making it useful for evaluating fine-grained segmentation and multi-modal 3D scene understanding.

DublinCity. The DublinCity [28] dataset is a city-scale annotated ALS point cloud dataset for urban scene understanding. It was collected over Dublin, Ireland, and provides manually annotated labels for approximately 260 million points selected from a much larger airborne LiDAR point cloud. The dataset has a very high point density, ranging from about 250 to 348 points/m², and contains 13 semantic classes organized in hierarchical levels of detail, including coarse urban categories such as buildings, vegetation, and ground, as well as finer elements such as windows, doors, and trees.

WHU-Urban3D. The WHU-Urban3D [29] dataset is a large-scale urban LiDAR point cloud dataset designed for semantic instance segmentation. It integrates airborne laser scanning point clouds, mobile laser scanning point clouds, and panoramic images, covering large-scale road and urban scenes from different cities. The dataset contains more than 200 million points with rich point-wise semantic instance annotations, and provides additional attributes such as reflected intensity and return information in addition to 3D coordinates. Compared with single-platform ALS datasets, WHU-Urban3D is more suitable for studying cross-platform, cross-scene, and multimodal point cloud understanding. Therefore, it provides an important benchmark for evaluating whether ALS segmentation models can generalize to complex urban environments and benefit from complementary MLS and image information.

Table 2 – Summary of representative ALS point cloud datasets for semantic segmentation.

Dataset	Year	Density	#Classes	#Points
ISPRS [2]	2012	4-7pts/m ²	9	1.166M
LASDU [25]	2020	3-4pts/m ²	5	3.12M
DALES [26]	2020	50 pts/m ²	8	500M
H3D [27]	2021	800 pts/m ²	11	74M
DublinCity [28]	2019	250-384pts/m ²	13	260M
WHU-Urban3D [29]	2024	60pts/m ²	9	213M
ÉCLAIR [30]	2024	50pts/m ²	11	582M
FRONTAL [31]	2024	37pts/m ²	7	9.261B

2.3.2 Evaluation metrics

Commonly used metrics for evaluating the performance of 3D semantic segmentation include Overall Accuracy (OA), mean Intersection-over-Union (mIoU), and F1-score.

Assuming that K is the total number of classes, and TP , FP , FN , and TN denote True Positive, False Positive, False Negative, and True Negative, respectively, the three evaluation metrics can be calculated as follows:

$$OA = \frac{TP}{TP + FP + FN + TN}, \quad (1)$$

$$mIoU = \frac{1}{K} \sum_{k=1}^K \frac{TP_k}{TP_k + FP_k + FN_k}, \quad (2)$$

$$F1 = \frac{1}{K} \sum_{k=1}^K \frac{2TP_k}{2TP_k + FP_k + FN_k}. \quad (3)$$

Here, OA measures the fraction of correctly classified points, whereas mIoU and mean F1-score are macro-averaged over the K semantic classes. Reported values should therefore be interpreted together with class definitions and benchmark-specific evaluation protocols.

3. WEAK SUPERVISION MECHANISMS FOR ALS POINT CLOUD SEMANTIC SEGMENTATION

Weakly supervised semantic segmentation of ALS point clouds aims to recover dense point-wise labels from incomplete, inexact, or inaccurate supervision. ALS scenes add strong variation in density, height, object scale, and class frequency. A useful framework must therefore explain not only how few labels are used, but also how those labels cover the scene, how uncertain supervision is recognized, and how errors are prevented from propagating during training.

The three dimensions form a feedback loop rather than a sequence of independent techniques. Annotation acquisition determines which classes and structures enter the training set. Reliability modeling converts these labels and model predictions into selected, weighted, or refined supervision. Optimization updates the model and produces new predictions that can change both reliability estimates and subsequent sample selection. A weakness at one stage therefore constrains the others. Missing rare classes cannot be recovered by loss reweighting alone, and a strong optimizer can amplify biased pseudo-labels when reliability calibration fails.

Throughout this section, evidence is distinguished by validation domain. Studies evaluated on ALS benchmarks are treated as direct evidence, whereas studies based on indoor scans, mobile or automotive LiDAR, and general point-cloud benchmarks are used only to identify candidate mechanisms. For the latter, transferability is assessed according to the mechanism's dependence on object completeness, neighborhood topology, point density, temporal redundancy, and auxiliary imagery; their reported performance is not taken as evidence of effectiveness on ALS data.

3.1 Weak annotation acquisition and point selection

The first problem in weakly supervised ALS segmentation is how to obtain informative supervision under a limited annotation budget. Dense point-wise labeling is difficult to scale because ALS scenes may contain millions of irregularly distributed points and because annotators must distinguish complex categories such as terrain, roofs, vegetation, roads, power lines, poles, vehicles, and other small infrastructure objects. A useful weak-label strategy should therefore reduce annotation cost while still preserving sufficient category coverage, spatial coverage, and structural cues for later pseudo-label generation and network training.

Sparse point supervision preserves the point-wise localization required by the final task. ALS-validated studies [32]–[34] use sparse labels as anchors and expand them through confidence filtering, spatial constraints, or iterative pseudo-labeling. Their main advantage is compatibility with standard segmentation backbones. Their main risk is incomplete semantic coverage. A nominal label ratio can hide whether power lines, poles, vehicles, and class boundaries were sampled at all. The results of Wang and Yao [33] also show that initialization changes convergence and pseudo-label quality. Sparse labels are therefore most effective when sampling is class-aware and spatially distributed, rather than uniformly random.

Coarse region or subcloud supervision reduces annotation interactions by assuming local semantic homogeneity. Direct ALS evidence from Lin et al. [35] shows that subcloud labels can seed point-wise predictions, but the loss of localization disproportionately affects mixed regions and thin classes. Region mining [36], clustering-based densification [37], and hypergraph relations [38] provide adjacent-domain mechanisms rather than ALS validation. Their most plausible ALS use is to connect large, internally coherent ground, roof, or vegetation regions after constructing density and height-aware superpoints. Transfer becomes unreliable near roof-tree interfaces, terrain-object boundaries, and power lines, where dominant-class leakage can erase small or vertically overlapping structures. These mechanisms are therefore conditionally transferable and require localized point cues or boundary constraints to protect rare classes.

Segments, clicks, and scribbles occupy a middle ground between sparse points and coarse scene labels. SegGroup [39] and OTOC [40], [41] were developed for indoor scenes, whereas scribble supervision [42] and temporal graph propagation [43] were validated on automotive LiDAR sequences. Their transfer potential differs. Segment relations may be useful for high-density UAV data if superpoints preserve vertical discontinuities. Scribbles are promising for elongated roads or power lines because they provide spatially extended cues, whereas object-centric one-click expansion is less suited to incomplete top-down ALS objects. Temporal propagation requires repeated acquisitions and accurate strip registration, which are unavailable in most ALS datasets. These approaches should therefore be treated as targeted ALS annotation hypotheses rather than interchangeable label-efficient solutions.

Active learning has both direct ALS and adjacent-domain evidence. ALS studies [44], [45], [49]–[52] support entropy, spatial, and class-aware selection, while general point-cloud and MLS studies [46]–[48], [53]–[55] introduce region diversity, structural reasoning, and hierarchical querying. These criteria are comparatively transferable because they operate on model uncertainty and spatial coverage rather than on complete object instances. However, ALS adaptation requires tile or strip-level querying, density-normalized scores, broad spatial coverage, and class-aware safeguards. Without these changes, the query budget may concentrate on noisy boundaries, low-density artifacts, or dominant classes.

The preferred annotation unit should follow acquisition geometry rather than nominal label percentage. In low-density manned-aircraft ALS, class-aware sparse points can anchor rare infrastructure, while active tiles provide broad coverage across flight strips. In high-density UAV data, segment, click, or scribble supervision becomes more plausible after density normalization and topology-preserving superpoint construction. Coarse subcloud labels remain suitable for ground, roof, and vegetation, but mixed boundaries and thin structures require additional point or scribble cues. The most credible transfer path is a hybrid annotation design that combines broad regional coverage with localized supervision for rare or elongated classes.

3.2 Reliability modeling of weak supervision

After weak annotations are obtained, the next problem is reliability: which weak labels, pseudo labels, predictions, or features should be trusted during training. Pseudo-labeling [56] is widely used to expand limited supervision into denser point-level labels, but pseudo labels may contain severe errors when the initial labels are sparse, coarse, or spatially-biased. In ALS scenes, reliability modeling is especially important because uneven density, vertical overlap, ambiguous terrain-object boundaries, and severe class imbalance can cause early pseudo-label errors to be reinforced during self-training or consistency learning.

Output confidence is the simplest reliability signal. ALS methods [16], [32], [33], [57] use prediction probabilities, entropy, adaptive thresholds, or soft gates to select and weight pseudo-labels. These signals are inexpensive and model agnostic, but confidence is not equivalent to correctness. Class imbalance can make dominant categories confidently wrong and rare categories persistently uncertain. A single global threshold may therefore increase pseudo-label precision while removing the few available power-line or vehicle samples. Class-conditional calibration is more appropriate when label frequency and point density vary strongly across categories.

Uncertainty-aware methods estimate prediction instability rather than relying on a single confidence score. Evidence spans sparsely annotated point clouds [34], scribble or temporal LiDAR [42], [43], and adjacent-domain uncertainty-aware methods [59]–[65]. The mechanism is potentially transferable because uncertainty is derived from predictive disagreement rather than from a specific sensor. However, thresholds learned under indoor or automotive

class priors cannot be reused directly for ALS. Density and class-conditional calibration is needed, and uncertain thin classes should be weighted or queried rather than discarded. Ensemble or Bayesian variants may improve calibration but add training and memory cost.

Consistency treats stable predictions under perturbations, model updates, or neighboring observations as reliable. Direct airborne evidence includes current-versus-ensemble agreement [16] and feature-perturbation consistency for airborne multispectral LiDAR [67]. Self-purification [66] and driving, mobile, or general point-cloud studies [62], [63], [68]–[74] extend consistency across regions, views, teachers, and time. Teacher-student and feature-level consistency are plausible for ALS when augmentations preserve elevation order, return attributes, local density, and the connectivity of thin structures. Region mixing and temporal matching are more restrictive because cropping can remove power lines and temporal propagation requires repeated, accurately registered acquisitions. Feature-level consistency is therefore a nearer-term ALS option, whereas driving-derived temporal consistency remains conditional on data availability and registration quality.

Feature-space reliability asks whether weakly labeled samples form coherent semantic neighborhoods. Direct ALS studies use contrastive or prototype-based learning under subcloud, self-supervised, few-shot, and semi-supervised settings [35], [75]–[77], [84], while related methods [78]–[83], [85], [86] are validated mainly on adjacent point-cloud benchmarks. Contrastive and prototype mechanisms are plausible for ALS because they do not depend on image appearance, but their neighborhoods and augmentations must reflect airborne geometry. Transfer should use density-normalized sampling, height, intensity, or return-aware features, multiple prototypes for heterogeneous classes, and class-aware construction of positive and negative pairs. Otherwise, dominant ground or vegetation clusters can absorb scarce infrastructure samples. These methods are therefore promising but not directly portable without ALS-specific sampling and prototype design.

For reliability modeling, the transfer burden differs across mechanisms. Confidence weighting and teacher-student consistency are relatively portable but require ALS-specific calibration and perturbations. Feature-space constraints are conditionally portable because their semantic neighborhoods depend on density and geometry. Temporal matching is least portable because it assumes repeated observations and accurate registration. This hierarchy separates candidate mechanisms from demonstrated ALS performance and identifies calibration, structural preservation, and registration as the main validation targets.

3.3 Supervision optimization

The third problem is how to optimize the segmentation model once weak or reliability-filtered supervision is available. Even reliable pseudo-labels are usually incomplete, imbalanced, and spatially uneven. Therefore, weakly supervised ALS segmentation requires training strategies that can gradually enlarge supervision, refine noisy labels,

control the order of sample introduction, and reduce the dominance of frequent classes such as ground, vegetation, and buildings.

Self-training converts limited labels into dense targets, but its transferability depends on how the initial seeds are defined. Direct ALS studies [32], [35] start from sparse points or subcloud labels, whereas region mining [36] and general sparse-label propagation [87] provide adjacent-domain mechanisms. The iterative pseudo-label loop is portable in principle, but confidence thresholds, update frequency, and spatial expansion must be recalibrated for ALS density and rare-class priors. Point anchors preserve localization but may omit classes, while coarse seeds cover more area but amplify mixed-region errors. Gains should therefore be attributed only after controlling the backbone, label initialization, and pseudo-label quantity.

Label propagation uses spatial or feature neighborhoods, whereas pseudo-label refinement updates supervision through evolving model predictions. Superpoint, graph, local-query, and dense-propagation methods [66], [88]-[92] are supported mainly by adjacent point-cloud benchmarks, while DPLG-Net [93] provides direct ALS evidence. Neighborhood propagation can transfer to ALS where surfaces are smooth and locally homogeneous. It becomes unreliable near vertical overlaps and thin structures unless the graph uses height discontinuities, return or intensity cues, anisotropic neighborhoods, and boundary-aware stopping. Propagation is therefore plausible for ground and roofs, conditional for vegetation, and high risk for power lines and fences.

Progressive and curriculum strategies control when pseudo-labels enter optimization. P3CL [57] provides direct ALS evidence, while feature-perturbation learning on airborne multispectral LiDAR [67] and staged pseudo-label training on MLS data [94] provide related geospatial evidence. Scheduling is more portable than spatial propagation because it changes when a label is used rather than imposing a fixed neighborhood. However, easy-to-hard rankings must be calibrated by class and density so that rare classes do not enter too late. Stabilization-first schedules reduce early confirmation bias but extend training and introduce additional schedule parameters.

Robust losses change the influence of weak labels but cannot repair missing semantic coverage. Lin et al. [35] use an overlap-region loss for mixed subclouds, while Gao et al. [95] model category-dependent ALS label noise. Other ALS studies apply class balancing, focal terms, or combinations of cross-entropy, affinity, and Lovasz-Softmax losses [67],[96]-[98]. Reweighting can reduce domination by ground, buildings, or vegetation. It does not make an incorrect pseudo-label correct, and focal loss can amplify mislabeled samples because they appear difficult. Loss comparisons are therefore meaningful only under fixed sampling and pseudo-label protocols.

Optimization mechanisms also differ in portability. Confidence scheduling and robust losses can be adapted through ALS-specific calibration, whereas graph propagation requires ALS-aware neighborhoods and

temporal or object-level expansion depends on acquisition conditions. None can recover a class that is absent from the initial supervision. Adjacent-domain optimization methods should therefore be presented as design candidates, and their ALS value should be established through controlled class-wise and pseudo-label analyses rather than assumed from non-ALS benchmarks.

3.4 Comparative evidence and practical trade-offs

Reported benchmark values require separate treatment from comparative evidence. Table 3 and Table 4 summarize representative results on ISPRS Vaihingen, LASDU, H3D, and DALES, but they do not constitute a leaderboard because backbones, label initialization, training protocols, and supervision units differ. Table 5 therefore focuses exclusively on ALS-validated studies with a traceable supervision unit and at least one matched performance, computational, or annotation-design result. Table 6 then examines whether similar aggregate scores conceal different outcomes for power lines and fences.

Table 3 – Performance comparison of representative semantic segmentation methods on ISPRS and LASDU ALS benchmarks.

Method	Sup.	ISPRS		LASDU	
		OA	F1	OA	F1
PointNet++ [9]	100%	81.2	65.5	–	–
DGCNN [10]		95.1	84.2	85.5	73.1
RandLA-Net [12]		83.1	70.1	–	–
KPConv [11]		84.2	71.2	85.9	72.3
Wang et al. [33]	15 points	80.4	63.1	–	–
	30 points	81.1	65.4	–	–
	60 points	83.0	68.6	–	–
	100 points (0.2%)	83.7	70.2	–	–
Wang and Yao [16]	0.05%	78.2	61.9	87.3	74.6
	0.1%	83.0	70.0	87.6	77.4
	0.2%	83.4	70.7	87.7	78.1
	0.5%	84.0	71.9	87.6	78.3
PSD [72]	0.1%	75.5	51.5	85.1	70.3
	1%	79.1	61.4	86.7	74.7
SQN [90]	0.1%	74.5	52.1	81.2	72.7
	1%	80.8	61.8	85.1	75.6
DR-Net [99]	0.1%	67	51.8	80.4	70.2
	1%	77.5	62.1	83.2	73.7
Lin et al. [35]	0.5%(SC)	80	63.1	–	–
	1%(SC)	81.4	65.9	–	–
MPRM [36]	1%(SC)	75.2	59.6	–	–
HSCN [84]	0.1%	79.5	68.7	85.2	75.8
	1%	82.9	72.8	86.5	77.7
DSDCL [76]	0.1%	79.7	65.7	85.4	73.4
	1%	81.8	70.1	87.3	77.0
DPLG-Net [93]	0.05%	81.8	68.8	85.6	76.8
	0.1%	84.1	71.8	87.3	78.6
	0.2%	84.4	72.0	87	78.7
	0.5%	84.5	72.5	87.5	79.2
Dai et al. [49]	AL(1%)	–	70.4	–	–
Kolle et al. [50]	AL(RF)	81.6	67.5	–	–
	AL(SCN)	80.9	66.2	–	–
Thr-MPRNet [77]	CD	41.4	–	–	–
Yuan et al. [100]	0.1%(CD)	81.8	65.1	–	–
	1%(CD)	83.7	68.4	–	–

Table 4 – Performance comparison of representative semantic segmentation methods on H3D and DALES ALS benchmarks.

Method	Sup.	H3D		DALES	
		OA	F1	OA	mIoU
PointNet++ [9]	100%	–	–	95.7	68.3
RandLA-Net [12]		–	–	97.1	80.0
KPConv [11]		–	–	97.8	81.1
Wang and Yao [16]	0.01%	85.8	72.1	–	–
	0.05%	86.7	75	–	–
	0.1%	88.2	76.3	–	–
	0.2%	87.5	76.6	–	–
OCOC [51]	750 SC (0.023%)	86.5	74.2	–	–
Kolle [50]	AL(RF)	84.8	70.4	–	–
	AL (SCN)	85.4	70.9	–	–
Thr-MPRNet [77]	CD	44.1	–	76.8	65.3

Note: Sup. denotes the supervision setting. SC denotes subcloud-level supervision; AL denotes active learning; CD denotes cross-domain evaluation. RF denotes random forest, and SCN denotes submanifold sparse convolutional Network. "–" indicates that the corresponding result was not reported in the original paper.

Matched within study comparisons show that weak supervision retained 89.4–99.1% of the fully supervised mean F1-score across four representative ALS settings. These results should not be interpreted as a ranking because the underlying supervision budgets include sparse points, subcloud labels, and queried subclouds, which involve different annotation procedures. Table 5 therefore reports the supervision unit, performance retention, and available cost evidence together. Overall, weak supervision can substantially reduce dense point-wise annotation, although the total saving depends on the additional effort required for sample selection, subcloud inspection, pseudo-label refinement, and iterative training.

Table 5 – Reported label efficiency and cost evidence from ALS-validated studies.

Study	Supervision, performance, reported cost
Wang and Yao [16]	0.1% points; 70.0 vs 71.2 F1 (98.3% retention); 2.3 vs 1.9 h training; inference not separated.
Wang and Yao [33]	0.1% points (330 labels); 56.39 vs 53.77 F1 under random initialization; computation unavailable; isolates label initialization.
Lin et al. [35]	0.5% subclouds; 63.1 vs 70.6 F1 (89.4% retention); per-sphere time reported; human interaction unavailable.
Lin et al. [44]	282 actively selected full-label tiles (31.7%); reached full-training mIoU; 2.9 h selection and training; annotation time excluded.
Kölle et al. [50]	Active point labels, typically <1%; within about 5 percentage points of full labels; RF about 1 min/iteration and SCN about 50 times slower; oracle-dependent.

Study	Supervision, performance, reported cost
OCOC [51]	750 subclouds with one click/class (0.023% points); 74.19 vs 74.86 F1 (99.1% retention); human loop time unavailable; point ratio omits subcloud inspection.
P3CL [57]	0.1% points; 64.9 vs 70.1 F1 (92.6% retention); training increased by 51%; inference remained similar.
DPLG-Net [93]	0.1% points; 71.8 F1 on ISPRS and 78.6 F1 on LASDU; computation, human time, and cross-sensor validation unavailable.

Mean F1-score alone does not reveal whether performance is retained for geometrically-thin classes. Table 6 therefore compares the power-line and fence F1-scores of representative weakly supervised methods with those of their corresponding fully supervised backbones on the ISPRS dataset.

Table 6 – Class-level F1-scores (%) for representative backbone-aligned comparisons between full and weak supervision on the ISPRS dataset.

Method	Supervision	Power line	Fence	Mean F1
KPConv [11]	100% points	74.2	29.6	71.2
Wang and Yao [16]	0.1% points	72.8	31.5	70.0
Lin et al. [35]	0.5% subclouds	44.9	11.8	63.1
Lin et al. [35]	1% subclouds	45.7	28.9	65.9
RandLA-Net [12]	100% points	69.0	36.2	70.1
P3CL [57]	0.1% points	57.7	33.4	64.9

These class-level results show that similar mean F1-scores can mask distinct rare-class behaviour. With KPConv as the reference, the sparse-point method of Wang and Yao [16] largely retains power-line performance and slightly improves fence performance. At 0.5% subcloud supervision, Lin et al. [35] shows substantial losses for both classes. Increasing the supervision ratio to 1% recovers most of the fence performance but leaves a large power-line deficit. Relative to RandLA-Net, P3CL [57] retains most of the fence performance but exhibits a larger power-line deficit. Although these comparisons are illustrative rather than exhaustive, they demonstrate that nominal annotation ratios and mean scores alone are insufficient for assessing label efficiency. Class-level metrics should therefore accompany aggregate performance, particularly for thin and rare objects.

4. PERSPECTIVES AND FUTURE WORK

The main future research directions discussed in this section are summarized in Figure 3, including standardized weak-label benchmarks, cross-scene and cross-sensor generalization, promptable foundation models for pseudo-label generation, and multimodal remote sensing fusion.

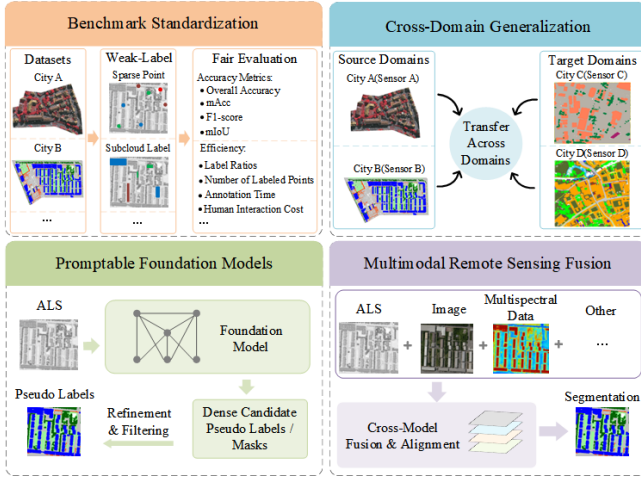


Figure 3 – Evidence-informed research directions for weakly supervised ALS segmentation. Benchmark standardization addresses incomparable annotation budgets and reporting practices, while cross-domain evaluation tests scene and sensor transfer. Promptable foundation models and multimodal fusion represent less mature directions that require controlled evaluation of projection, registration, density, and missing-modality errors.

4.1 Benchmark standardization

Standardized weak-label benchmarks are still insufficient for ALS point cloud semantic segmentation. Although existing datasets, such as ISPRS [2], LASDU [25], DALES [26] and H3D [27], have promoted the development of fully supervised ALS segmentation, most of them were not originally designed for evaluating weakly supervised learning under unified annotation budgets. Recent benchmark-oriented studies have begun to emphasize this problem. Kölle et al. [50] benchmarked active learning strategies for ALS point cloud semantic segmentation, while Kölle et al. [101] further developed a fully automatized active learning framework and highlighted the high cost of constructing reliable ALS annotations. WHU-Urban3D [29] also provides a large-scale urban dataset with ALS, MLS, and panoramic images, showing the potential of cross-platform and multimodal benchmarks.

Future benchmarks should pair dense ground truth with fixed weak-label masks for sparse points, subclouds, clicks, scribbles, and active-learning rounds. Accuracy metrics should be accompanied by absolute label counts, class coverage, human annotation time, and interaction logs. Cost-performance curves based on mean F1-score would be more comparable than percentages alone across different point densities. Reporting training time, peak memory, inference time, and per-class calibration would also expose methods that trade annotation savings for hidden computational cost.

4.2 Cross-domain generalization

Cross-domain generalization in this review involves two distinct questions: whether a trained model transfers across ALS scenes or sensors, and whether a weak supervision mechanism developed outside ALS remains valid under airborne acquisition. The first concerns representation and prediction shift. The second concerns assumptions about annotation units, neighborhood topology, point density,

temporal redundancy, and image alignment. A non-ALS mechanism can therefore be conceptually transferable without its reported accuracy being transferable. Future studies should revalidate these assumptions on ALS rather than treating adjacent-domain results as surrogate performance evidence.

Recent geospatial studies address transfer at different stages. Wang et al. [102] adapt predictions at test time across photogrammetric, airborne, mobile, and synthetic point clouds. APCoTTA [103] extends this setting to continual airborne LiDAR shifts, while Yuan et al. [100] transfer representations to ALS targets with a few labeled samples. These settings are complementary rather than directly comparable. Feature transfer assumes labeled target samples during training, test-time adaptation assumes access to unlabeled target batches, and continual adaptation must resist forgetting. Future evaluations should state the available target supervision and separate cross-city, cross-sensor, few-shot, and continual conditions.

4.3 Promptable foundation models

Promptable foundation models, represented by the Segment Anything Model (SAM) [104], can convert sparse interactions into candidate masks. However, the supporting evidence spans different tasks and sensing domains. Foundation-model-assisted 2D weak supervision [105] benefits from dense image boundaries. SAL [106] projects 2D knowledge into driving LiDAR, Point-SAM [107] distills SAM into a promptable 3D model, and Yu et al. [108] use SAM-derived cues for noisy 3D instance supervision. These studies demonstrate prompt flexibility, but they do not yet establish reliable semantic segmentation of large-area ALS point clouds.

ALS presents a stricter grounding problem because image coverage may be incomplete and point-image registration errors can systematically corrupt roofs, facades, terrain, and thin infrastructure. Native 3D prompting avoids projection but remains sensitive to density and scale shifts in its training data. ALS studies should therefore separate the contribution of prompt generation, 2D-3D alignment, and subsequent pseudo-label filtering. Useful tests include controlled misregistration, missing RGB, density variation, cross-season imagery, and class-aware prompt budgets, with targeted human correction reserved for low-confidence thin structures.

4.4 Multimodal remote sensing fusion

Multimodal fusion can compensate for the limited spectral and texture information of ALS, but existing studies support different claims. Chen et al. [67] combine spectral and geometric cues under weak supervision for airborne multispectral LiDAR classification. EGAFNet [109] uses ALS-derived top-view projections for semantic segmentation, while Cao et al. [110] transfer semantics from posed images without 3D annotations. The first provides direct weak supervision evidence for classification, the second provides ALS segmentation evidence without establishing label efficiency gains, and the third demonstrates cross-modal transfer beyond an ALS-specific

setting. These mechanisms are complementary, not a single validated pipeline.

For weakly supervised ALS segmentation, auxiliary modalities are useful only when their class-specific information exceeds the uncertainty introduced by alignment, resolution, season, occlusion, and missing data. Reliability-aware fusion should therefore be tested under controlled misregistration, temporal offset, and modality dropout. Class-wise gains are essential because a higher overall accuracy may reflect easier ground or vegetation predictions while power lines and boundaries remain unchanged. Reporting annotation and computational cost alongside these tests would show whether the additional modality provides practical value rather than only greater model capacity.

5. CONCLUSION

Weakly supervised semantic segmentation provides a practical way to reduce the dependence of ALS point cloud interpretation on dense point-wise annotations. This review summarized recent progress from three perspectives: weak annotation acquisition and sample selection, reliability modeling of weak labels and pseudo-labels, and supervision optimization through self-training, label propagation, curriculum learning, and robust loss design. The reviewed studies show that weakly supervised ALS segmentation is not simply a problem of using fewer labels; its effectiveness depends on semantic coverage, pseudo-label reliability, and training stability under sparse, imbalanced, and ambiguous supervision. Future research should place more emphasis on standardized weak-label benchmarks, fair annotation-cost evaluation, cross-domain generalization, promptable foundation-model-assisted pseudo-label generation, and reliability-aware multimodal fusion. These directions are essential for making weakly supervised ALS segmentation more scalable, robust, and applicable to real-world geospatial mapping tasks.

REFERENCES

- [1] Wehr, A., and Lohr, U., "Airborne laser scanning-an introduction and overview," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 54, Nos. 2-3, 1999, pp. 68-82.
- [2] Niemeyer, J., Rottensteiner, F., and Soergel, U., "Contextual classification of lidar data and building object detection in urban areas," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 87, 2014, pp. 152-165.
- [3] Yan, W. Y., Shaker, A., and El-Ashmawy, N., "Urban land cover classification using airborne LiDAR data: A review," *Remote Sensing of Environment*, Vol. 158, 2015, pp. 295-310.
- [4] Hyypä, J., Hyypä, H., Leckie, D., Gougeon, F., Yu, X., and Maltamo, M., "Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests," *International Journal of Remote Sensing*, Vol. 29, No. 5, 2008, pp. 1339-1366.
- [5] Xi, S., Zhang, Z., Niu, Y., Li, H., and Zhang, Q., "Power line extraction and tree risk detection based on airborne LiDAR," *Sensors*, Vol. 23, No. 19, 2023, p. 8233.
- [6] Yan, J., Zlatanova, S., Aleksandrov, M., Diakite, A., and Pettit, C., "Integration of 3D objects and terrain for 3D modelling supporting the digital twin," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 4, 2019, pp. 147-154.
- [7] Wang, R., Peethambaran, J., and Chen, D., "Lidar point clouds to 3-D urban models: A review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 11, No. 2, 2018, pp. 606-627.
- [8] Qi, C. R., Su, H., Mo, K., and Guibas, L. J., "Pointnet: Deep learning on point sets for 3d classification and segmentation," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652-660.
- [9] Qi, C. R., Yi, L., Su, H., and Guibas, L. J., "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," *Advances in Neural Information Processing Systems*, Vol. 30, 2017.
- [10] Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., and Solomon, J. M., "Dynamic graph cnn for learning on point clouds," *ACM Transactions on Graphics (tog)*, Vol. 38, No. 5, 2019, pp. 1-12.
- [11] Thomas, H., Qi, C. R., Deschud, J.-E., Marcotegui, B., Goulette, F., and Guibas, L. J., "Kpconv: Flexible and deformable convolution for point clouds," *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6411-6420.
- [12] Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., and Markham, A., "Randla-net: Efficient semantic segmentation of large-scale point clouds," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11108-11117.
- [13] Choy, C., Gwak, J., and Savarese, S., "4d spatio-temporal convnets: Minkowski convolutional neural networks," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 3075-3084.
- [14] Zhao, H., Jiang, L., Jia, J., Torr, P. H., and Koltun, V., "Point transformer," *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 16259-16268.
- [15] Wu, X., Lao, Y., Jiang, L., Liu, X., and Zhao, H., "Point transformer v2: Grouped vector attention and partition-based pooling," *Advances in Neural Information Processing Systems*, Vol. 35, 2022, pp. 33330-33342.

- [16] Wang, P., and Yao, W., "A new weakly supervised approach for ALS point cloud semantic segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 188, 2022, pp. 237-254.
- [17] Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., and Bennamoun, M., "Deep learning for 3d point clouds: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 43, No. 12, 2020, pp. 4338-4364.
- [18] He, Y., Yu, H., Liu, X., Yang, Z., Sun, W., Anwar, S., and Mian, A., "Deep learning based 3D segmentation in computer vision: A survey," *Information Fusion*, Vol. 115, 2025, p. 102722.
- [19] Xiao, A., Zhang, X., Shao, L., and Lu, S., "A survey of label-efficient deep learning for 3D point clouds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 46, No. 12, 2024, pp. 9139-9160.
- [20] Guo, L., Chehata, N., Mallet, C., and Boukir, S., "Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 66, No. 1, 2011, pp. 56-66.
- [21] Weinmann, M., Jutzi, B., Hinz, S., and Mallet, C., "Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 105, 2015, pp. 286-304.
- [22] Zhang, J., Lin, X., and Ning, X., "SVM-based classification of segmented airborne LiDAR point clouds in urban areas," *Remote sensing*, Vol. 5, No. 8, 2013, pp. 3749-3775.
- [23] Guo, B., Huang, X., Zhang, F., and Sohn, G., "Classification of airborne laser scanning data using JointBoost," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 100, 2015, pp. 71-83.
- [24] Wu, X., Jiang, L., Wang, P.-S., Liu, Z., Liu, X., Qiao, Y., Ouyang, W., He, T., and Zhao, H., "Point transformer v3: Simpler faster stronger," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024, pp. 4840-4851.
- [25] Ye, Z., Xu, Y., Huang, R., Tong, X., Li, X., Liu, X., Luan, K., Hoegner, L., and Stilla, U., "Lasdu: A large-scale aerial lidar dataset for semantic labeling in dense urban areas," *ISPRS International Journal of Geo-Information*, Vol. 9, No. 7, 2020, p. 450.
- [26] Varney, N., Asari, V. K., and Graehling, Q., "DALES: A large-scale aerial LiDAR data set for semantic segmentation," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 186-187.
- [27] Kölle, M., Laupheimer, D., Schmohl, S., Haala, N., Rottensteiner, F., Wegner, J. D., and Ledoux, H., "The Hessigheim 3D (H3D) benchmark on semantic segmentation of high-resolution 3D point clouds and textured meshes from UAV LiDAR and Multi-View-Stereo," *ISPRS Open Journal of Photogrammetry and Remote Sensing*, Vol. 1, 2021, p. 100001.
- [28] Zolanvari, S., Ruano, S., Rana, A., Cummins, A., Da Silva, R. E., Rahbar, M., and Smolic, A., "DublinCity: Annotated LiDAR point cloud and its applications," *arXiv preprint arXiv:1909.03613*, 2019.
- [29] Han, X., Liu, C., Zhou, Y., Tan, K., Dong, Z., and Yang, B., "WHU-Urban3D: An urban scene LiDAR point cloud dataset for semantic instance segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 209, 2024, pp. 500-513.
- [30] Melekhov I, Umashankar A, Kim H J, et al. "Eclair: A high-fidelity aerial lidar dataset for semantic segmentation," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024: 7627-7637.
- [31] Gaydon C, Daab M, Roche F. Fractal: "An ultra-large-scale aerial lidar dataset for 3d semantic segmentation of diverse landscapes," *arXiv preprint arXiv:2405.04634*, 2024.
- [32] Wang, P., and Yao, W., "Weakly supervised pseudo-label assisted learning for ALS point cloud semantic segmentation," *arXiv preprint arXiv:2105.01919*, 2021.
- [33] Wang, P., Yao, W., et al., "Exploring label initialization for weakly supervised als point cloud semantic segmentation," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 2, 2022, pp. 151-158.
- [34] Yao, Y., Xu, K., Murasaki, K., Ando, S., and Sagata, A., "Pseudo-labelling-aided semantic segmentation on sparsely annotated 3D point clouds," *IPSJ Transactions on Computer Vision and Applications*, Vol. 12, No. 1, 2020, p. 2.
- [35] Lin, Y., Vosselman, G., and Yang, M. Y., "Weakly supervised semantic segmentation of airborne laser scanning point clouds," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 187, 2022, pp. 79-100.
- [36] Wei, J., Lin, G., Yap, K.-H., Hung, T.-Y., and Xie, L., "Multi-path region mining for weakly supervised 3D semantic segmentation on point clouds," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 4384-4393.
- [37] Xia, S., Yue, J., Kania, K., Fang, L., Tagliasacchi, A., Moo Yi, K., and Sun, W., "Densify your labels: Unsupervised clustering with bipartite matching for weakly supervised point cloud segmentation," *arXiv e-prints*, 2023, pp. arXiv-2312.
- [38] Lu, Z., Zhang, P., Dai, Y., Li, W., and Su, Z., "Hypergraph convolutional network based weakly supervised point cloud semantic segmentation with scene-level annotations," *Neurocomputing*, Vol. 620, 2025, p. 129264.
- [39] Tao, A., Duan, Y., Wei, Y., Lu, J., and Zhou, J., "SegGroup: Seg-level supervision for 3D instance and semantic segmentation," *IEEE Transactions on Image Processing*, Vol. 31, 2022, pp. 4952-4965.

- [40] Liu, Z., Qi, X., and Fu, C.-W., "One thing one click: A self-training approach for weakly supervised 3d semantic segmentation," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 1726-1736.
- [41] Liu, Z., Qi, X., and Fu, C.-W., "You Only Need One Thing One Click: Self-Training for Weakly Supervised 3D Scene Understanding," *World Scientific Annual Review of Artificial Intelligence*, Vol. 2, 2024, p. 2440005.
- [42] Unal, O., Dai, D., and Van Gool, L., "Scribble-supervised lidar semantic segmentation," *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, 2022, pp. 2697-2707.
- [43] Shi, H., Wei, J., Li, R., Liu, F., and Lin, G., "Weakly supervised segmentation on outdoor 4D point clouds with temporal matching and spatial graph propagation," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 11840-11849.
- [44] Lin, Y., Vosselman, G., Cao, Y., and Yang, M. Y., "Active and incremental learning for semantic ALS point cloud segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 169, 2020, pp. 73-92.
- [45] Lin, Y., Vosselman, G., Cao, Y., and Yang, M., "Efficient training of semantic point cloud segmentation via active learning," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 2, 2020, pp. 243-250.
- [46] Shi, X., Cai, L., Chen, K., Foo, C. S., Jia, K., and Xu, X., "Label-Efficient Point Cloud Semantic Segmentation: A Holistic Active Learning Approach," *World Scientific Annual Review of Artificial Intelligence*, Vol. 2, 2024, p. 2440006.
- [47] Shao, F., Luo, Y., Liu, P., Chen, J., Yang, Y., Lu, Y., and Xiao, J., "Active learning for point cloud semantic segmentation via spatial-structural diversity reasoning," *Proceedings of the 30th ACM International Conference on Multimedia*, 2022, pp. 2575-2585.
- [48] Wu, T.-H., Liu, Y.-C., Huang, Y.-K., Lee, H.-Y., Su, H.-T., Huang, P.-C., and Hsu, W. H., "Redal: Region-based and diversity-aware active learning for point cloud semantic segmentation," *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 15510-15519.
- [49] Dai, H., and Zhao, Z., "Spatial Incremental Sampling for Active-Learning-Based Airborne Laser Scanning Point Cloud Segmentation," *2024 IEEE International Conference on Smart Internet of Things (SmartIoT)*, IEEE, 2024, pp. 421-426.
- [50] Kölle, M., Walter, V., Schmohl, S., and Soergel, U., "Learning on the edge: Benchmarking active learning for the semantic segmentation of als point clouds," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 10, 2023, pp. 945-952.
- [51] Wang, P., Yao, W., and Shao, J., "One class one click: Quasi scene-level weakly supervised point cloud semantic segmentation with active learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 204, 2023, pp. 89-104.
- [52] Liang, Z.-X., Lai, X.-D., and Yan, Y.-T., "Label-efficient weakly supervised semantic segmentation for airborne LiDAR point clouds," *Journal of infrared and millimeter waves*, Vol. 44, No. 6, 2025, pp. 16-16.
- [53] Lei, X., Guan, H., Ma, L., Liu, J., Yu, Y., Wang, L., Dong, Z., Ni, H., and Li, J., "DAAL-WS: A weakly-supervised method integrated with data augmentation and active learning strategies for MLS point cloud semantic segmentation," *International Journal of Applied Earth Observation and Geoinformation*, Vol. 131, 2024, p. 103970.
- [54] Liu, G., van Kaick, O., Huang, H., and Hu, R., "Active self-training for weakly supervised 3D scene semantic segmentation," *Computational Visual Media*, Vol. 10, No. 3, 2024, pp. 425-438.
- [55] Xu, Z., Yuan, B., Zhao, S., Zhang, Q., and Gao, X., "Hierarchical point-based active learning for semi-supervised point cloud semantic segmentation," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 18098-18108.
- [56] Lee, D.-H., et al., "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," *Workshop on challenges in representation learning, ICML*, Vol. 3, Atlanta, 2013, p. 896.
- [57] Luo, Z., Zeng, T., Jiang, J., Cai, Z., Wu, W., Xie, Z., and Xu, Y., "P3CL: Pseudo-Label Confidence-Calibrated Curriculum Learning for Weakly Supervised Urban Airborne Laser Scanning Point Cloud Classification," *Remote Sensing*, Vol. 18, No. 4, 2026, p. 552.
- [58] Shen, L., Cao, Y., Wei, X., and Li, S., "Weakly supervised point cloud semantic segmentation using pseudo-label reliability and consistency regularization," *Neurocomputing*, Vol. 639, 2025, p. 130241.
- [59] Niu, Y., and Yin, J., "PA-Net: trustworthy weakly supervised point cloud semantic segmentation with primary-auxiliary structure," *Computers and Electrical Engineering*, Vol. 119, 2024, p. 109555.
- [60] Zhou, C., and Ling, Q., "Weakly supervised large-scale point cloud semantic segmentation based on dual consistency constraints and uncertainty-aware fusion," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 63, 2024, pp. 1-13.
- [61] Shi, H., Liu, F., Wu, Z., Xu, Y., and Lin, G., "Weakly Supervised Segmentation on Outdoor 4D Point Clouds With Progressive 4D Grouping," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 47, No. 5, 2025, pp. 3487-3499.

- [62] Xuan, W., Qi, H., and Xiao, A., "TSG-Seg: Temporal-selective guidance for semi-supervised semantic segmentation of 3D LiDAR point clouds," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 216, 2024, pp. 217-228.
- [63] Kong, L., Ren, J., Pan, L., and Liu, Z., "Lasermix for semi-supervised lidar semantic segmentation," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 21705-21715.
- [64] Unal, O., Sakaridis, C., and Van Gool, L., "Bayesian Self-training for Semi-supervised 3D Segmentation," *European Conference on Computer Vision*, Springer, 2024, pp. 89-107.
- [65] Kwon, D., Park, T., and Kwak, S., "RePL: Pseudo-label Refinement for Semi-supervised LiDAR Semantic Segmentation," *arXiv preprint arXiv:2604.06825*, 2026.
- [66] Li, J., Yan, F., Liu, Y., and Zhuang, Y., "Weakly Supervised Semantic Segmentation for Point Clouds with Self-Purification on Pseudo Labels," *2023 13th International Conference on Information Science and Technology (ICIST)*, IEEE, 2023, pp. 100-107.
- [67] Chen, K., Guan, H., Wang, L., Yu, Y., Zang, Y., Qin, N., Liu, J., and Li, J., "A feature perturbation weakly supervised learning network for airborne multispectral LiDAR pointcloud classification," *International Journal of Applied Earth Observation and Geoinformation*, Vol. 127, 2024, p. 103683.
- [68] Liu, J., Guan, H., Lei, X., and Yu, Y., "Weakly supervised semantic segmentation of mobile laser scanning point clouds via category balanced random annotation and deep consistency-guided self-distillation mechanism," *The Photogrammetric Record*, Vol. 38, No. 184, 2023, pp. 581-602.
- [69] Lee, S., Lee, H., and Shim, H., "Learning from Spatio-temporal Correlation for Semi-Supervised LiDAR Semantic Segmentation," *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2024, pp. 14095-14102.
- [70] Liu, Y., Chen, Y., Wang, H., Belagiannis, V., Reid, I., and Carneiro, G., "Ittakestwo: Leveraging peer representations for semi-supervised lidar semantic segmentation," *European Conference on Computer Vision*, Springer, 2024, pp. 81-99.
- [71] Liu, C., Weng, X., Jiang, S., Li, P., Yu, L., and Xia, G.-S., "Exploring Scene Affinity for Semi-Supervised LiDAR Semantic Segmentation," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025, pp. 27380-27389.
- [72] Zhang, Y., Qu, Y., Xie, Y., Li, Z., Zheng, S., and Li, C., "Perturbed self-distillation: Weakly supervised large-scale point cloud semantic segmentation," *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 15520-15528.
- [73] Wu, Z., Wu, Y., Lin, G., Cai, J., and Qian, C., "Dual adaptive transformations for weakly supervised point cloud segmentation," *European conference on computer vision*, Springer, 2022, pp. 78-96.
- [74] Wu, Y., Yan, Z., Cai, S., Li, G., Han, X., and Cui, S., "Pointmatch: A consistency training framework for weakly supervised semantic segmentation of 3d point clouds," *Computers & Graphics*, Vol. 116, 2023, pp. 427-436.
- [75] Zhang, Y., Yao, J., Zhang, R., Wang, X., Chen, S., and Fu, H., "Havana: Hard negative sample-aware self-supervised contrastive learning for airborne laser scanning point cloud semantic segmentation," *Remote Sensing*, Vol. 16, No. 3, 2024, p. 485.
- [76] Luo, Z., Zeng, T., Jiang, X., Peng, Q., Ma, Y., Xie, Z., and Pan, X., "Dense Supervised Dual-Aware Contrastive Learning for Airborne Laser Scanning Weakly Supervised Semantic Segmentation," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 63, 2025, pp. 1-15.
- [77] Dai, M., Xing, S., Xu, Q., Li, P., Pan, J., Zhang, G., and Wang, H., "Multiprototype relational network for few-shot ALS point cloud semantic segmentation by transferring knowledge from photogrammetric point clouds," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 62, 2024, pp. 1-17.
- [78] Li, R., Cao, A.-Q., and de Charette, R., "Coarse3d: Class-prototypes for contrastive learning in weakly-supervised 3d point cloud segmentation," *arXiv preprint arXiv:2210.01784*, 2022.
- [79] Liu, M., Zhou, Y., Qi, C. R., Gong, B., Su, H., and Anguelov, D., "Less: Label-efficient semantic segmentation for lidar point clouds," *European conference on computer vision*, Springer, 2022, pp. 70-89.
- [80] Jiang, L., Shi, S., Tian, Z., Lai, X., Liu, S., Fu, C.-W., and Jia, J., "Guided point contrastive learning for semi-supervised point cloud semantic segmentation," *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 6423-6432.
- [81] Du, A., Zhou, T., Pang, S., Wu, Q., and Zhang, J., "Pcl: Point contrast and labeling for weakly supervised point cloud semantic segmentation," *IEEE Transactions on Multimedia*, Vol. 26, 2024, pp. 8902-8914.
- [82] Wang, J., He, J., Liu, Y., Chen, C., Zhang, M., and Tan, H., "Multi-scale classification and contrastive regularization: Weakly supervised large-scale 3D point cloud semantic segmentation," *Remote Sensing*, Vol. 16, No. 17, 2024, p. 3319.
- [83] Liu, K., Zhao, Y., Nie, Q., Gao, Z., and Chen, B. M., "Weakly supervised 3d scene segmentation with region-level boundary awareness and instance discrimination," *European conference on computer vision*, Springer, 2022, pp. 37-55.

- [84] Zeng, T., Luo, F., Guo, T., Gong, X., Shu, W., and Zhao, Y., "HSCN: Semi-supervised ALS point cloud semantic segmentation via hybrid structure constraint network," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 62, 2024, pp. 1-13.
- [85] Li, M., Xie, Y., Shen, Y., Ke, B., Qiao, R., Ren, B., Lin, S., and Ma, L., "Hybridcr: Weakly-supervised 3d point cloud semantic segmentation via hybrid contrastive regularization," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 14930-14939.
- [86] Su, Y., Xu, X., and Jia, K., "Weakly supervised 3D point cloud segmentation via multi-prototype learning," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 33, No. 12, 2023, pp. 7723-7736.
- [87] Zhang, Y., Li, Z., Xie, Y., Qu, Y., Li, C., and Mei, T., "Weakly supervised semantic segmentation for large-scale point cloud," *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, 2021, pp. 3421-3429.
- [88] Cheng, M., Hui, L., Xie, J., and Yang, J., "Sspc-net: Semi-supervised semantic 3d point cloud segmentation network," *Proceedings of the AAAI conference on artificial intelligence*, Vol. 35, 2021, pp. 1140-1147.
- [89] Deng, S., Dong, Q., Liu, B., and Hu, Z., "Superpoint-guided semi-supervised semantic segmentation of 3D point clouds," *2022 International conference on robotics and automation (ICRA)*, IEEE, 2022, pp. 9214-9220.
- [90] Hu, Q., Yang, B., Fang, G., Guo, Y., Leonardis, A., Trigoni, N., and Markham, A., "SqN: Weakly-supervised semantic segmentation of large-scale 3d point clouds," *European conference on computer vision*, Springer, 2022, pp. 600-619.
- [91] Wei, J., Lin, G., Yap, K.-H., Liu, F., and Hung, T.-Y., "Dense supervision propagation for weakly supervised semantic segmentation on 3d point clouds," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 34, No. 6, 2023, pp. 4367-4377.
- [92] Jiang, Z., Xiang, P., and Peng, C., "Enhancing Weakly Supervised Point Cloud Semantic Segmentation via Dynamic Label Propagation and Pseudo-Label Optimization," *2024 International Symposium on Intelligent Robotics and Systems (ISoIRS)*, IEEE, 2024, pp. 138-142.
- [93] Chen, J., Zhang, Y., Li, J., Zeng, T., Li, R., Qi, Y., Sun, Z., and Huang, K., "DPLG-Net: Deep pseudo-label-guided network for weakly-supervised semantic segmentation of ALS point clouds," *IEEE Transactions on Geoscience and Remote Sensing*, 2026.
- [94] Lei, X., Guan, H., Ma, L., Yu, Y., Dong, Z., Gao, K., Delavar, M. R., and Li, J., "WSPointNet: A multi-branch weakly supervised learning network for semantic segmentation of large-scale mobile laser scanning point clouds," *International journal of applied earth observation and geoinformation*, Vol. 115, 2022, p. 103129.
- [95] Gao, Y., Xia, S., Wang, C., Xi, X., Yang, B., and Xie, C., "Semantic segmentation of airborne lidar point clouds with noisy labels," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 62, 2024, pp. 1-14.
- [96] Nong, X., Bai, W., and Liu, G., "Airborne LiDAR point cloud classification using PointNet++ network with full neighborhood features," *Plos one*, Vol. 18, No. 2, 2023, p. e0280346.
- [97] Widyaningrum, E., Bai, Q., Fajari, M. K., and Lindenbergh, R. C., "Airborne laser scanning point cloud classification using the DGCNN deep learning method," *Remote Sensing*, Vol. 13, No. 5, 2021, p. 859.
- [98] Pan, J., Cao, J., Xing, S., Dai, M., Liu, J., Wang, X., Zhang, Y., and Huang, G., "An aerial point cloud classification using point transformer via multi-feature fusion," *Scientific Reports*, Vol. 15, No. 1, 2025, p. 22924.
- [99] Zhang, L., and Bi, Y., "Weakly-supervised point cloud semantic segmentation based on dilated region," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 62, 2024, pp. 1-20.
- [100] Yuan, J., Ma, H., Zhang, L., Deng, J., Luo, W., Liu, K., and Cai, Z., "Transfer Learning Based on Multi-Branch Architecture Feature Extractor for Airborne LiDAR Point Cloud Semantic Segmentation with Few Samples," *Remote Sensing*, Vol. 17, No. 15, 2025, p. 2618.
- [101] Kölle, M., Walter, V., and Sörgel, U., "Building a fully-automatized active learning framework for the semantic segmentation of geospatial 3D point clouds," *PFG-Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, Vol. 92, No. 2, 2024, pp. 131-161.
- [102] Wang, P., Yao, W., Shao, J., and He, Z., "Test-time adaptation for geospatial point cloud semantic segmentation with distinct domain shifts," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 229, 2025, pp. 422-435.
- [103] Gao, Y., Xia, S., Nie, S., Wang, C., Xi, X., and Yang, B., "APCoTTA: Continual Test-Time Adaptation for semantic segmentation of airborne LiDAR point clouds," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 237, 2026, pp. 339-354.
- [104] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., et al., "Segment anything," *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 4015-4026.
- [105] Yang, X., and Gong, X., "Foundation model assisted weakly supervised semantic segmentation," *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2024, pp. 523-532.

- [106] Ošep, A., Meinhardt, T., Ferroni, F., Peri, N., Ramanan, D., and Leal-Taixé, L., "Better call sal: Towards learning to segment anything in lidar," *European Conference on Computer Vision*, Springer, 2024, pp. 71-90.
- [107] Zhou, Y., Gu, J., Chiang, T. Y., Xiang, F., and Su, H., "Point-sam: Promptable 3d segmentation model for point clouds," *arXiv preprint arXiv:2406.17741*, 2024.
- [108] Yu, Q., Du, H., Liu, C., and Yu, X., "When 3d bounding-box meets sam: Point cloud instance segmentation with weak-and-noisy supervision," *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024, pp. 3719-3728.
- [109] Li, P., Chen, Z., Long, C., Zhang, H., Lv, Y., Huang, R., Dong, Z., and Yang, B., "A novel projection map driven multimodal fusion framework for ALS point cloud semantic segmentation," *International Journal of Applied Earth Observation and Geoinformation*, Vol. 144, 2025, p. 104907.
- [110] Cao, Z., Mi, X., Qiu, B., Cao, Z., Long, C., Yan, X., Zheng, C., Dong, Z., and Yang, B., "Cross-modal semantic transfer for point cloud semantic segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 221, 2025, pp. 265-279.
- [111] Zhou, Z.-H., "A brief introduction to weakly supervised learning," *National Science Review*, Vol. 5, No. 1, 2018, pp. 44-53.
- [112] Gao, Y., Xia, S., Wang, P., Xi, X., Nie, S., and Wang, C., "LiDAR remote sensing meets weak supervision: Concepts, methods, and perspectives," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 235, 2026, pp. 72-104.

AUTHORS



Ziwei Luo received a B.S. degree in spatial information and digital technology and a Ph.D. degree in surveying science and technology from China University of Geosciences (Wuhan), Wuhan, China, in 2018 and 2024, respectively. She is currently an assistant professor at the School of Computer and Artificial Intelligence, Wuhan Textile University, Wuhan. Her research interests include 3D vision and point cloud processing.



Jun Jiang received the B.S. degree in software engineering from Hubei University, China. He is currently pursuing an M.S. degree in artificial intelligence at the School of Computer and Artificial Intelligence, Wuhan Textile University. His research includes point cloud processing and deep learning.



Yuan Wu received a Ph.D. degree in engineering from the School of Computer Science, Wuhan University, China, in June 2022. He is currently an associate professor with the School of Computer Science and Artificial Intelligence, Wuhan Textile University, Wuhan. His research interests include deep learning, artificial intelligence and machine interaction.



Xiaoyun Yan received a Ph.D. degree in control science and engineering from Huazhong University of Science and Technology, Wuhan, China, in 2019. He is currently a lecturer with the School of Computer Science and Artificial Intelligence, Wuhan Textile University, Wuhan. His research interests include machine learning, computer vision, image retrieval, and intelligent multimedia content analysis.



Yangjun Ou received a B.S. degree in photogrammetry and remote sensing from Wuhan University in 2016 and a Ph.D. degree in photogrammetry and remote sensing from Wuhan University in 2021. She is currently a lecturer with Wuhan Textile University. Her research interests include computer vision and video understanding.



Zhaojing Wang received a Ph.D. degree in control science and engineering from Huazhong University of Science and Technology, Wuhan, China, in 2021. She joined the School of Computer Science and Artificial Intelligence, Wuhan Textile University, as a lecturer in 2022, and currently serves as an associate professor. Her research interests include process monitoring, fault diagnosis, soft sensor modeling and equipment health management.