



Delft University of Technology

Document Version

Final published version

Licence

CC BY

Citation (APA)

Parvaz, S., Teferle, F. N., Nurunnabi, A., Lindenberg, R., & Leiva, L. A. (2026). Airborne Point Cloud Fusion with Local Plane Constraints for Advanced Semantic Consistency. *Remote Sensing*, 18(15), Article 2598.
<https://doi.org/10.3390/rs18152598>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Article

Airborne Point Cloud Fusion with Local Plane Constraints for Advanced Semantic Consistency

Shahoriar Parvaz ^{1,*}, Felicia N. Teferle ¹, Abdul Nurunnabi ¹, Roderik Lindenbergh ² and Luis A. Leiva ³

¹ Geodesy and Geospatial Engineering, Faculty of Science, Technology and Medicine, University of Luxembourg, 1359 Kirchberg, Luxembourg; rebecca.teferle@uni.lu (F.N.T.); pyal1471@yahoo.com (A.N.)

² Department of Geoscience and Remote Sensing, Delft University of Technology, 2628 CN Delft, The Netherlands; r.c.lindenbergh@tudelft.nl

³ Department of Computer Science, Faculty of Science, Technology and Medicine, University of Luxembourg, 4364 Esch-sur-Alzette, Luxembourg; luis.leiva@uni.lu

* Correspondence: shahoriar.parvaz@uni.lu

Highlights

This study presents an airborne cross-source point cloud fusion framework that leverages local plane constraints and point-to-plane alignment to improve geometric consistency between LiDAR and photogrammetric data, particularly Dense Image Matching (DIM) point clouds. Unlike conventional workflows for cross-source fusion, which end with rigid registration or georeferencing datasets into a common reference frame, this research goes further by extracting PCA-derived local planar features and incorporating them into the fusion process to enhance surface preservation and reduce structural inconsistencies arising from density variations, noise, and acquisition differences. Experimental results demonstrate improved semantic consistency and a more reliable representation of urban features, particularly for buildings and ground surfaces, compared with conventional fusion strategies.

What are the main findings?

- The proposed point-to-plane sequential refinement for the cross-source fusion strategy significantly improved geometric alignment and semantic consistency between airborne LiDAR and DIM point clouds in complex urban environments.
- Incorporating PCA-based local plane constraints improved surface preservation and reduced registration errors caused by differences in sensor accuracy, noise, and other cross-source discrepancies.

What are the implications of the main findings?

- Improved fusion quality can support more accurate automated 3D city modeling, urban mapping, and semantic segmentation applications using heterogeneous airborne datasets.
- The proposed framework demonstrates that the local plane constraint in airborne cross-source point cloud fusion provides a robust alternative to conventional fusion methods for large-scale dataset integration and reconstruction.



Academic Editor: Sheng Nie

Received: 14 May 2026

Revised: 21 July 2026

Accepted: 23 July 2026

Published: 5 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Abstract

Point cloud fusion is crucial in geospatial analysis, combining data from multiple sources (e.g., LiDAR and photogrammetry) to provide a more complete and accurate environmental representation. However, integrating airborne hybrid sensors or cross-source point clouds remains challenging due to variations in geometric accuracy, data precision, gaps, and sensor attributes. Despite recent advancements, these challenges remain and are among the most demanding aspects in geospatial data processing for remote sensing applications. We

propose a new point cloud fusion algorithm that leverages local plane constraints to achieve advanced semantic consistency. The proposed method dynamically fits local planes to the target point clouds, enabling robust alignment of source points to these planes. Evaluation on two real-world datasets demonstrates significant gains in accuracy and preservation of geometric details. Our algorithm also improves the accuracy of downstream tasks such as semantic segmentation. In our experiment, the overall accuracy for the Dodelange dataset increases from 48.5% to 80.1%, and that for the Dublin dataset increases from 72.9% to 88.0%. While challenges persist with sparse and noisy datasets, experimental results highlight the effectiveness of the proposed method, offering valuable insights for maximizing the potential of cross-source point cloud data.

Keywords: ALS; airborne sensor; photogrammetry; geospatial data; feature extraction; cross-source; multi-source data; classification; city modelling

1. Introduction

Point clouds, consisting of dense collections of three-dimensional (3D) coordinates, have emerged as a cornerstone for spatial data representation in various domains. They provide a precise and detailed depiction of real-world objects and terrains, making them indispensable for high-spatial-accuracy applications. The importance of point clouds lies in their ability to offer a geometric basis for advanced analyses, such as classification, segmentation, and object recognition [1–3], enabling decision-making processes across industries such as topographic modeling [4,5], forest monitoring [6], building information modeling [7], environmental analysis [8], and 3D city modeling [9,10].

The process of acquiring airborne point clouds mainly relies on two methods: laser-based and photogrammetry (Figure 1). Aerial laser scanning (ALS), also known as Light Detection and Ranging (LiDAR), excels at capturing precise geometric data across large regions, including detailed terrain models and vegetation mapping [4,11]. On the other hand, photogrammetry, particularly Dense Image Matching (DIM), reconstructs 3D points by analyzing overlapping aerial images, resulting in high-resolution data enriched with texture. A significant advancement in this field is oblique photogrammetry (Figure 1-DIM), which takes images from tilted angles and has further enhanced the ability to model complex urban environments [5,12]. This approach significantly improves downstream research by revealing important details of building facades and intricate structures that may be hidden in standard vertical (nadir) photogrammetry.

Despite the individual strengths of LiDAR and DIM, a single data source is often insufficient to capture the full complexity of urban scenarios. Cross-source point clouds, derived by combining LiDAR and DIM, address this limitation by integrating the complementary strengths of both technologies [10,13]. LiDAR provides precise structural and terrain data, while DIM enriches the dataset with high-resolution textures and finer details. This integration also improves semantic segmentation or scene understanding, which is crucial for many applications, such as disaster management, infrastructure monitoring, and 3D city modeling [6,9,10].

However, the fusion of cross-source point clouds presents significant challenges, beginning with the registration process [6,14]. Point cloud registration, which involves aligning datasets into a common coordinate system, is a critical and necessary first step for any subsequent processing. Registration methods primarily focus on global alignment [15,16]. While they ensure initial alignment, they may not adequately handle local geometric variations. Precise point cloud fusion, which builds upon accurate registration, is essential for

ensuring consistency, reducing noise, and eliminating redundancies in the merged dataset. Without precise fusion, inaccuracies and inconsistencies in registered point clouds can propagate into later stages.

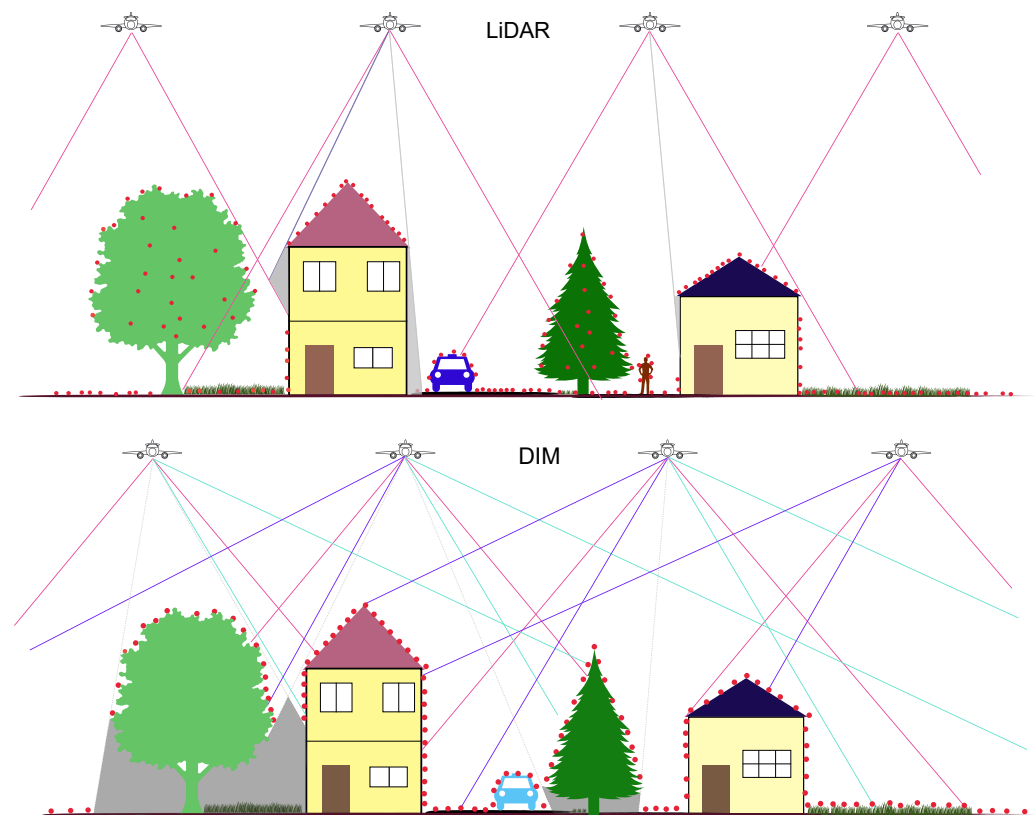


Figure 1. Data acquisition layout with airborne LiDAR and DIM from nadir and oblique imagery (Reproduced from Parvaz et al. [17]).

Recent studies have sought to address the challenges associated with point cloud fusion [14,17,18]. Conventional methods, particularly rigid transformations, employ iterative optimization techniques to achieve alignment between point clouds [8,19,20]. While these methods can be effective under controlled conditions, they often encounter limitations when applied to cross-source datasets that exhibit irregular densities, noise, and varying resolutions. These inconsistencies are particularly relevant for airborne DIM point clouds, where vertical accuracy is generally lower than that of LiDAR and the reconstructed geometry can vary depending on the photogrammetric processing method, making spatially varying residual errors difficult to correct using a single global transformation. Furthermore, recent non-rigid alignment methods also require the establishment of point correspondences between the point clouds [21,22], a task that becomes increasingly challenging in large-scale airborne cross-source datasets.

Given these limitations, this study proposes a single-pass local refinement and fusion method for globally pre-registered airborne LiDAR and DIM point clouds. Rather than introducing PCA for registration, plane fitting, or point-to-plane projection as independent algorithms, the contribution lies in combining PCA-derived target planes with source-normal-guided point-wise adjustments to reduce local cross-source geometric inconsistencies before fusion. The proposed framework therefore addresses the methodological gap between global registration and local cross-source geometric refinement. Beyond geometric refinement, the study also investigates whether local plane-constrained cross-source alignment improves downstream semantic consistency on two complementary urban datasets and compares it with conventional and modern fine-registration-based approaches. The

remainder of the paper reviews related work, presents the mathematical formulation and methodology, and describes the experimental datasets and evaluation procedure.

2. Related Literature

Research on LiDAR and DIM point clouds has primarily focused on registration, whereas fewer studies have addressed their subsequent local refinement and fusion. Since the proposed method assumes that global registration is already available, this section briefly reviews only the registration approaches relevant to the selected baselines, followed by cross-source fusion methods and their principal challenges.

Conventional methods like Iterative Closest Point (ICP) remain foundational [23,24]. ICP iteratively computes correspondences between closest points and estimates a transformation to align the datasets. Although effective for pre-aligned datasets [15,16], ICP is sensitive to initialization, outliers, and noise. Beyond ICP, feature-based [25,26], graph-based [21,27], and learning-based [28–30] methods have been developed to improve registration under noise, occlusion, and large initial misalignment. However, their application to heterogeneous airborne LiDAR and DIM data remains challenging due to differences in density, coverage, sensor accuracy, and the need for cross-source local alignment [16,31]. As the present study focuses on local refinement after global registration, these approaches are reviewed only as contextual background.

2.1. Point Cloud Fusion

Point cloud fusion aims to integrate multiple point clouds into a single, coherent 3D representation [4,17,20,32]. The effective fusion of LiDAR with photogrammetry-derived DIM point clouds requires precise registration to resolve the geometric inconsistencies that arise from their differing characteristics [6,13,33]. Conventionally, after registration or georeferencing within the same coordinate system, point cloud datasets are fused by simply combining them into a unified representation (Equation (1)). This usually works with the same-source or terrestrial data, but for cross-source and airborne data, it creates two layers of data when no fusion operation is performed [17]. Effective fusion is particularly valuable in multi-sensor environments, where it can combine the strengths of different technologies to overcome individual limitations. Our contribution in this research is to find this fusion operator Φ .

$$\mathcal{F} = \mathcal{P}_s \cup \mathcal{P}_t \quad (1)$$

$$\mathcal{F} = \Phi(\mathcal{P}_s, \mathcal{P}_t), \quad (2)$$

where $\mathcal{P}_s$ is the source point cloud, $\mathcal{P}_t$ is the target point cloud, Φ denotes the fusion operation and $\mathcal{F}$ represents the resulting fused point cloud.

Fusion between datasets acquired by different sensors is of particular interest [6,13]. LiDAR provides precise geometric information and can penetrate vegetation, while DIM offers high-density data and detailed color information. By integrating these complementary modalities, fusion yields high-quality 3D reconstructions.

Typical workflows involve coarse georeferencing of DIM data using GNSS, IMU information or ground control points, followed by fine registration using algorithms such as ICP or VGICP [6,14,20]. Scene segmentation (e.g., by street networks or object boundaries) is often employed to enable localized alignment. Methods such as Line-Based fusion have also been explored to combine structural features across modalities [17,33,34], and hybrid adjustment frameworks have emerged that jointly process LiDAR and DIM data to enhance overall georeferencing [8,19]. The cross-source point cloud fusion work of Parvaz et al. [17] is notable for addressing these challenges using a slice-to-slice adjustment method. Their approach breaks down point clouds into smaller slices and performs local alignments,

mitigating the impact of cross-source variations on overall fusion accuracy. A comparison of selected representative registration and fusion studies, which may serve as relevant benchmarks for evaluating new cross-source registration and fusion methods, is provided in Table 1.

Table 1. Summary of representative point cloud registration and fusion studies relevant to cross-source airborne data.

Study	Year	Data Sources	Main Task	Core Method	Main Limitation
Besl and McKay [23]	1992	Point clouds	Rigid registration	ICP	Sensitive to initialization and outliers
Huang et al. [26]	2020	Cross-source point clouds	Global registration	Feature matching	Dependent on reliable cross-source descriptors
Toschi et al. [13]	2021	Airborne LiDAR and DIM	Data integration	Geometric quality assessment	Does not perform point-wise local refinement
Koide et al. [24]	2021	Point clouds	Rigid registration	ICP	Primarily estimates a global transformation
Parvaz et al. [17]	2024	Airborne LiDAR and DIM	Cross-source fusion	Slice-to-slice adjustment	Depends on predefined slicing and structural features
Proposed method	–	Airborne LiDAR and DIM	Cross-source fusion	Local plane-constrained point-wise adjustment	Requires locally planar neighborhoods

2.2. Challenges in Fusion

Despite notable advances, cross-source point cloud registration and fusion remain affected by several challenges. These difficulties primarily stem from the inherent variability of data collected by different sensors [14]. The main challenges are summarized as follows.

Heterogeneous data characteristics: Point clouds captured by different sensors often vary significantly in scale, density, noise level, and completeness.

- **Scale variation:** Different sensors may operate using varying physical metrics, introducing relative scale discrepancies [12,35]. Accurate registration under unknown scaling conditions requires robust normalization or scale-invariant strategies.
- **Density differences:** LiDAR point density depends on range distance and incidence angle. The incidence angle is notably varying in urban environments due to differences in surface orientation; e.g., roofs and facades. In contrast, DIM may exhibit non-uniform density in a different way, as it is influenced by image overlap or surface texture [27,36]. Such disparities complicate correspondence matching.
- **Noise and outliers:** Sensor-specific limitations and environmental conditions contribute to varying noise characteristics. For example, LiDAR data might be affected by atmospheric refraction and ghost points, artificially placed in between scattering surfaces [37], whereas DIM can introduce artifacts from mismatched image features. Robust statistical approaches proposed by Nurunnabi et al. [38,39] have proven effective for reliable outlier detection.
- **Missing data and occlusions:** Each sensor has its own limitations regarding surface visibility and occlusion [27], leading to missing regions in the reconstructed scenes.

- Smoothing effects: LiDAR data typically has crisp but irregular edges, while DIM often shows rounded building edges.

Correspondence establishment: Accurately matching features across cross-source is challenging due to the lack of consistent correspondence and geometric characteristics [21,30,33]. Conventional feature descriptors may fail across modalities, motivating the development of cross-modal and learning-based feature extraction methods.

Fusion accuracy and robustness: Algorithms must cope with large initial misalignments, inconsistent densities, and high outlier ratios. Ensuring robust and accurate alignment under these conditions, especially for large-scale scenes, remains a difficult task [31]. Furthermore, the generalization of learning-based methods across datasets with diverse sensor characteristics is an active research question.

Computational efficiency: Applications such as real-time 3D object detection and autonomous driving require high-speed processing of large-scale point clouds [16,40]. Fusion through conventional non-rigid registration methods are computationally demanding, depend on reliable dense correspondences, and are evaluated mainly on relatively small deformable-object datasets [21,22,41]. Their scalability and transferability to partially overlapping, density-imbalanced airborne LiDAR–DIM clouds containing millions of urban points therefore remain insufficiently demonstrated.

Lack of standardized benchmarks: There is a shortage of publicly available and standardized datasets specifically curated for evaluating cross-source fusion methods [16,17,31]. As emphasized by Huang et al. [31], establishing well-defined benchmarks is crucial for enabling fair comparisons and driving further development in the field.

Addressing these challenges is fundamental for unlocking the full potential of cross-source point cloud fusion. Recent research emphasizes that beyond accurate alignment, semantic consistency across modalities is critical for reliable scene interpretation [1,6,42]. As such, efforts have increasingly shifted toward evaluating how well fused data supports downstream tasks like classification and segmentation. Benchmark datasets such as H3D [43] expose the need for algorithms that can operate simultaneously on ultra-dense UAV–LiDAR and multi-view-stereo meshes, encouraging research on cross-source label transfer and evaluation. Various methods are employed for cross-source point cloud segmentation, including conventional machine learning algorithms utilizing geometric and color features [44] and, increasingly, deep learning approaches [2,42]. Probabilistic frustum-based frameworks integrate 2D image detections with LiDAR data into volumetric 3D grids, enabling robust object detection under occlusion and sparsity [40]. Multimodal attention networks [42] enhance results by fusing features within and across modalities, improving small-object detection and class balance. However, generalization across sensor types and environments remains a key challenge. Future research should prioritize the development of generalizable, scalable, and efficient solutions that can handle the diversity inherent in multi-sensor 3D data.

3. Mathematical Definition

The problem of cross-source point cloud fusion can be defined as the process of aligning and merging two point clouds, a source point cloud (e.g., DIM) $\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$ and a target point cloud (e.g., LiDAR) $\mathbf{T} = \{\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_m\}$, into a unified coordinate system such that the geometric inconsistencies between the datasets are minimized. Unlike rigid alignment, where a single transformation matrix is applied to the entire source point cloud [15,29], the cross-source point cloud fusion requires to operate on a per-point basis, enabling non-rigid alignment.

For each point $\mathbf{s}_i \in \mathbf{S}$, the objective is to find a local transformation vector $\mathbf{d}_{si} \in \mathbb{R}^3$ that minimizes the distance to its corresponding plane in the target point cloud $\mathbf{T}$. This can be expressed mathematically as:

$$\mathbf{F}_i = \arg \min_{\mathbf{d}_{si}} \|\mathbf{n}_{s(i)} \cdot (\mathbf{s}_i + \mathbf{d}_{si} - \mathbf{p}_j)\|^2 \quad (3)$$

where $\mathbf{s}_i$ is the individual source point under the local transformation $\mathbf{F}_i$, $\mathbf{p}_j$ is the orthogonal projection point on the locally fitted plane in $\mathbf{T}$, and $\mathbf{n}_{s(i)}$ is the normal vector of that source point.

This process ensures that each point in the source point cloud is aligned to its locally corresponding geometry in the target point cloud. As described later, by employing Principal Component Analysis (PCA) to derive the plane and the point $\mathbf{p}_j$ on it, the algorithm adapts to local variations in density and structure, making it suitable for handling hybrid datasets. Because a separate plane is estimated for each target neighborhood, the method can represent spatial variations in local surface orientation, although its reliability depends on sufficient point density and local planarity. Importantly, the update is computed in closed form for each source point without any inner iterative optimization. The point-to-plane distance is evaluated once, and if it is below the prescribed threshold, the point is projected onto the plane. Otherwise, its original position is preserved. This leads to an overall computational complexity that scales linearly with the number of source points.

This point-to-plane approach effectively functions as a non-rigid alignment technique, addressing the limitations of conventional global methods and enabling a more precise and reliable fusion of cross-source point clouds.

4. Methodology

The proposed method for point cloud fusion relies on PCA-based point-to-plane alignment (Figure 2) to emphasize local geometric adaptation to overcome the challenges associated with cross-source datasets. The method (Figure 3) is broken down into the following key steps.

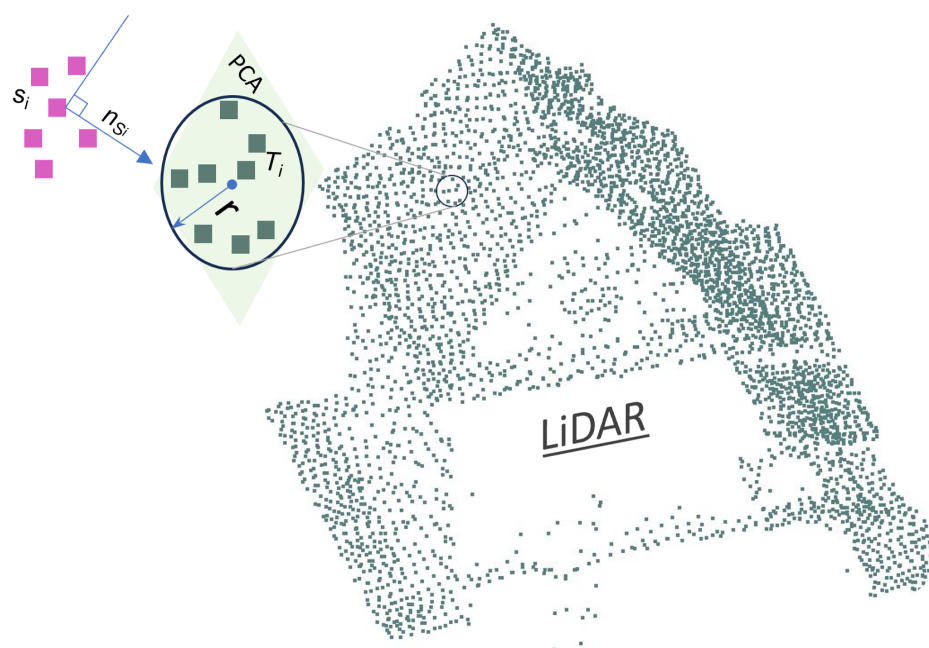


Figure 2. Illustration of the fusion process using local plane constraints.

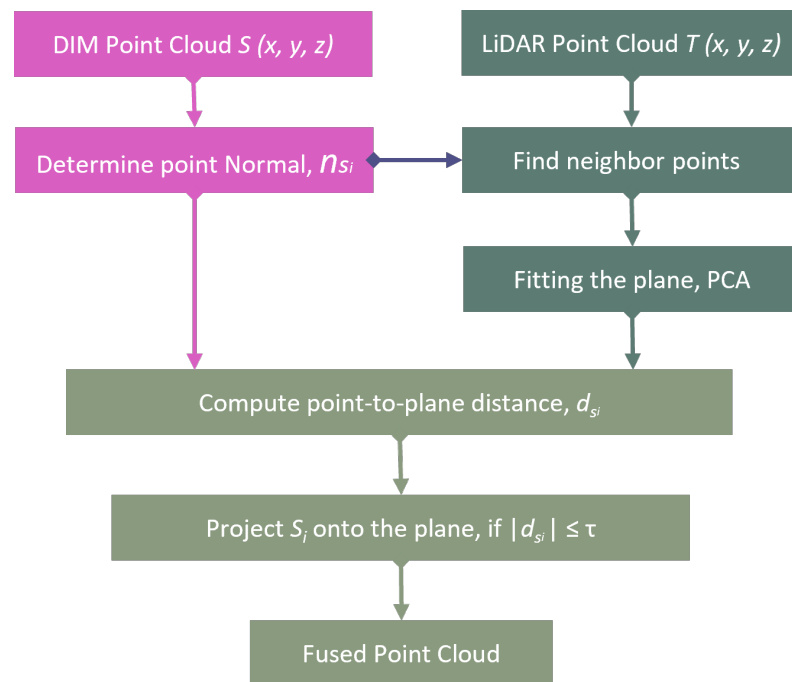


Figure 3. Fusion steps for cross-source point cloud.

1. **Preprocessing:** Prior to fusion, both the DIM and LiDAR point clouds undergo pre-processing to enhance data quality. This includes noise and outlier removal through statistical filters to eliminate spurious points as well as coordinate system alignment to ensure consistent spatial references. Additionally, source point's normal $\mathbf{n}_s$ is either determined during the process or can be imported as pre-calculated normals.
2. **Neighborhood selection:** For each point $\mathbf{s}_i$ in the source point cloud, a corresponding neighborhood $N_{t_{nn(i)}}$ in the target point cloud is identified. This is achieved using a fixed-radius nearest neighbor search, ensuring that the selected neighborhood captures the local geometry around the nearest point $\mathbf{t}_i$ in the target cloud.
3. **Local plane fitting:** Using the identified neighborhood $N_{t_{nn(i)}}$, a local plane is fitted by PCA. The eigenvectors corresponding to the first two eigenvalues define the plane, while the centroid of the neighborhood serves as a representative point $\mathbf{p}_j$ on the plane.
4. **Point-to-plane projection:** Each source point $\mathbf{s}_i$ is projected toward the fitted local plane of the target point cloud along the direction of its surface normal to preserve local surface orientation. This critical decision was adopted because the facade points from airborne LiDAR near the ground and roof are sparse, and the PCA-fitted plane is mainly governed by horizontal surface points. The facade points from the source should remain fused with the LiDAR facade points, and this is only possible if they move along the source normal within the threshold τ . The projection distance d_{si} is computed as

$$d_{si} = \mathbf{n}_{s(i)} \cdot (\mathbf{s}_i - \mathbf{p}_j), \quad (4)$$

where $\mathbf{n}_{s(i)}$ represents the normal vector of the source point $\mathbf{s}_i$.

5. **Local transformation:** Instead of estimating a full rigid or non-rigid transformation matrix, the proposed method performs a constrained local point update along the source point normal direction. Each source point $\mathbf{s}_i$ is adjusted according to the computed projection distance d_{si} , enabling local geometric refinement while preserving the original surface characteristics.
6. **Sequential refinement:** The source point cloud is refined sequentially by processing each source point individually. For every source point $\mathbf{s}_i$, a local neighborhood is determined

in the target point cloud, followed by local plane fitting and point-to-plane projection. After fusing one source point, the algorithm proceeds to the next until all source points have been processed. The proposed method therefore performs a single-pass local refinement rather than repeated iterative optimization.

7. Fusion output: After all source points have been refined, the updated source point cloud is merged with the target point cloud to generate the final fused dataset. Following fusion, nearest-neighbor interpolation of RGB and intensity values is applied to improve visual and semantic consistency between the two point clouds.

This proposed fusion Algorithm 1 ensures that the source point cloud is adaptively fused to the target point cloud, addressing local geometric discrepancies while preserving global consistency. In sum, our single-pass local refinement method provides a robust framework for handling the challenges of cross-source airborne point clouds.

Algorithm 1 Fusion via Point-to-Plane Alignment

Input: Source point cloud $S = \{s_1, s_2, \dots, s_n\}$,
 target point cloud $T = \{t_1, t_2, \dots, t_m\}$,
 neighborhood radius r ,
 distance threshold τ

Output: Aligned source point cloud S'

```

1: for each point  $s_i \in S$  do
2:   Find neighborhood  $N_{t_{nn(i)}}$  within radius  $r$ 
3:   if  $|N_{t_{nn(i)}}| \geq 3$  then
4:     Fit a plane to  $N_{t_{nn(i)}}$  using PCA
5:     Project  $s_i$  onto the plane to obtain  $\mathbf{p}_j$ 
6:     Compute vector:  $\mathbf{v}_{s_i} = \mathbf{s}_i - \mathbf{p}_j$ 
7:     Extract normal vector  $\mathbf{n}_{s(i)}$ 
8:     Compute point-to-plane distance:  $d_{s_i} = \mathbf{v}_{s_i} \cdot \mathbf{n}_{s(i)}$ 
9:     if  $|d_{s_i}| \leq \tau$  then
10:      Adjust point:  $s'_i = s_i + d_{s_i} \mathbf{n}_{s(i)}$ 
11:    else
12:      Retain original point:  $s'_i = s_i$ 
13:    end if
14:  else
15:    Skip point
16:  end if
17: end for
18: return  $S' = \{s'_1, s'_2, \dots, s'_n\}$ 

```

5. Experiment Results

Two experiments were conducted to evaluate the effectiveness of the proposed fusion algorithm in terms of alignment accuracy and semantic segmentation performance. Alignment metrics assess how well the LiDAR and DIM point clouds are integrated, while semantic segmentation results highlight the benefits of improved fusion for downstream tasks such as surface interpretation and object classification.

To objectively evaluate the effectiveness of the proposed fusion algorithm, several alignment metrics are considered: fitness metric, overlap ratio, correspondence count and root mean square error (RMSE). These metrics provide a comprehensive evaluation of the accuracy, structural consistency, and robustness of the fused point clouds [15,17], and are calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{c=1}^N \|S_c - T_c\|^2} \quad (5)$$

$$\text{Fitness} = \frac{|S_c|}{|S|} \quad (6)$$

$$\text{Overlap ratio} = \frac{|T_c|}{|T|} \quad (7)$$

$$\text{Correspondence} = d(S_i, T_j) \leq \max 0.02 \text{ m} \quad (8)$$

where S_c and T_c are corresponding points in the source S and target T point clouds, respectively, and N is the total number of correspondences. The double vertical bars $\|\cdot\|$ denote the Euclidean norm, while single vertical bars $|\cdot|$ indicate the total count of points in a set and $d(S_i, T_j)$ is the Euclidean distance between a source point S_i and a target point T_j . Only source–target point pairs that are considered to have a valid correspondence are included in the evaluation, where validity is defined by a maximum distance of 0.02 m.

These metrics serve complementary roles in assessing the fusion performance. Firstly, RMSE quantifies the average error between corresponding points in the registered datasets. Secondly, the Fitness Metric measures the fraction of points in the source cloud that have a valid corresponding point in the target cloud, reflecting the degree of alignment between the two datasets. Thirdly, the Overlap Ratio represents the percentage of spatial regions in the source point cloud that successfully overlaps with the target cloud after transformation, providing insight into how well the fusion process preserves geometric structures. Lastly, the Correspondence Count indicates the number of successfully matched points between the two datasets, serving as an indicator of how well different regions of the point clouds are integrated. These metrics collectively ensure that the fusion process is evaluated not only in terms of global alignment accuracy but also in terms of local consistency and geometric completeness.

Semantic segmentation, which is the process of assigning a label to each point or pixel [1,3], is a prerequisite for further analysis and enhancement of point clouds [8]. To evaluate the practical impact of the proposed fusion algorithm, four primary classes (Ground, Building, Trees, and Unclassified) are segmented using widely accessible semantic segmentation tools for airborne point clouds. The ground-truth point clouds for both study areas were manually labelled through visual interpretation of the point cloud geometry and RGB information. Points that could not be assigned confidently to Ground, Building, or Trees were retained as Unclassified. The same manually labelled reference data and identical segmentation and evaluation procedures were used for all compared point cloud inputs to ensure a consistent assessment of the effect of the proposed fusion method. We employ several standard semantic segmentation performance metrics derived from the normalized confusion matrix.

$$\text{OA} = \frac{\sum_{i=1}^N TP_i}{\sum_{i=1}^N (TP_i + FP_i + TN_i + FN_i)} \quad (9)$$

$$\text{IoU}_i = \frac{TP_i}{TP_i + FP_i + FN_i} \quad (10)$$

$$\text{F1}_i = \frac{2 \cdot TP_i}{2 \cdot TP_i + FP_i + FN_i} \quad (11)$$

where FN_i , FP_i , and TP_i denote the false negatives, false positives, and true positives for class i , respectively. Overall Accuracy (OA) measures the proportion of correctly classified points over the entire dataset. Mean Intersection-over-Union (mIoU) evaluates the average

overlap between the predicted and actual class regions, effectively penalizing both under- and over-segmentation. Finally, the mean F1 score balances precision and recall, providing a harmonic mean of these two metrics to summarize per-class performance.

To evaluate the proposed fusion algorithm, we used two real-world urban datasets, representing diverse urban environments. The first dataset corresponds to Dudelange (Luxembourg) (<https://data.public.lu/en/datasets/lidar-2019-releve-3d-du-territoire-luxembourgeois/>, accessed on 1 September 2022), a medium-sized urban area with a mix of isolated residential buildings, open spaces, and vegetation. The second dataset represents Dublin (Ireland) [45], characterized by highly structured urban landscapes with dense building clusters, narrow streets, less vegetation and large-scale infrastructure. Both datasets contain LiDAR points with corresponding DIM data. The LiDAR point density is roughly 15 points/m² for the Dudelange dataset and 250 points/m² for the Dublin dataset. These differences contribute to variations in geometric details, occlusions, and data completeness, making them suitable for evaluating cross-source point cloud fusion methods.

5.1. Experiment 1

The first dataset was acquired from an altitude of 3000 m, which is considered high. The accuracy of airborne point clouds tends to decrease with elevation [11]. To account for this, a 30 cm distance threshold was applied to the sum of both data accuracies (5 cm for LiDAR and 25 cm for DIM) during the fusion process. The neighbourhood search radius is selected based on this distance threshold, and the average point density of the LiDAR data was set such that a sufficient number of points are available within the neighbourhood while avoiding overly large regions that would blur the local geometry. Accordingly, due to the lower density of LiDAR points in this dataset and the requirement of at least three points for plane fitting, a 40 cm neighbourhood search radius was selected. This ensured that sufficient neighbouring points were available to construct reliable local plane approximations.

While a few other research studies discuss the fusion of heterogeneous point clouds at a conceptual level, they do not provide a directly comparable local and non-rigid fusion algorithm with a publicly available implementation. For this reason, we use well-established baselines, ICP [23], VGICP [24] and Line-Based fusion [17], for quantitative comparison. All methods were evaluated on the same workstation equipped with 13th Gen Intel(R) Core(TM) i9-13950HX, 64 GB of memory and 24 physical cores/32 logical processors.

Table 2 compares the runtime and peak memory consumption of the evaluated methods using the same source (924,571 points) and target (440,921 points) clouds pre-aligned with a reference coordinate frame. ICP and VGICP require substantially less computation because they estimate a single global rigid transformation, whereas the Line-Based and proposed Plane-Based methods perform local pointwise refinement. Among the local methods, the proposed approach reduced the runtime from 135.40 s to 113.46 s and the peak memory consumption from 9.12 GB to 7.53 GB.

Table 2. Computational performance of the evaluated methods.

Method	Runtime (s)	Peak Memory (GB)	Alignment Type
ICP (point-to-plane)	1.77	0.38	Global rigid
VGICP	1.00	0.16	Global rigid
Line-Based	135.40	9.12	Local point-wise
Plane-Based (proposed)	113.46	7.53	Local point-wise

5.1.1. Alignment Metrics-Based Evaluation

The results presented in Table 3 demonstrate the performance differences among the three fusion methods employed in this study. ICP and VGICP exhibit the highest RMSE,

lowest fitness metric, and minimal overlap ratio, indicating poor alignment between point clouds. With only 261 and 501 correspondences established, both methods struggle to handle cross-source fusion effectively due to their reliance on rigid transformations and point-to-point distances, which do not fully account for variations in point cloud density and noise. In contrast, the Line-Based fusion method achieves a significantly lower RMSE (0.0025), suggesting a more precise alignment. The fitness metric and overlap ratio indicate moderate improvements in data fusion, largely due to its ability to incorporate linear structures within the point clouds. However, despite a considerable increase in correspondences (35,195 against 261 or 501), Line-Based fusion still lacks the adaptability required for heterogeneous datasets. Our proposed Plane-Based fusion approach outperforms the other methods in terms of correspondence count, fitness score, and overlap ratio, indicating a higher level of integration between the source and target point clouds. Although its RMSE is slightly higher than that of the Line-Based method due to being calculated on more correspondences, its superior fitness metric highlights its robustness in capturing geometric relationships beyond simple linear structures, making it more effective for complex urban datasets.

Table 3. Comparison of algorithm accuracy based on alignment metrics (Dataset-1).

Method	RMSE	Fitness	Overlap	Correspondences
ICP	0.0145	0.0004	0.0001	261
VGICP	0.0156	0.0008	0.0001	501
Line-Based	0.0025	0.0551	0.0781	35,195
Plane-Based (Ours)	0.0090	0.2360	0.2761	151,374

To further assess the effectiveness of the proposed point cloud fusion algorithm, we extracted three representative profiles from different urban environments (Figure 4). Each profile was taken from three datasets (LiDAR, DIM, and the fused DIM) allowing for a comparative analysis of geometric consistency and structural alignment. In the building facade profile, the alignment between LiDAR and DIM before fusion or with global transformation shows noticeable discrepancies, particularly in facade details. After fusion, the fused DIM point cloud closely aligns with the LiDAR data, demonstrating improved structural coherence. Similarly, in the narrow road profile, the fused point cloud reduces the deviations observed in the original DIM dataset, capturing road edges and surface variations with higher accuracy. The open ground profile, where elevation differences are minimal, further confirms that the fusion process effectively maintains terrain continuity without introducing artifacts. These profiles illustrate the improvement in geometric accuracy and structural consistency achieved through the proposed fusion algorithm, demonstrating its applicability to complex urban environments.

5.1.2. Semantic Segmentation-Based Evaluation

To further validate the effectiveness of our point cloud fusion approach, we evaluate semantic segmentation accuracy across four scenarios: LiDAR-only, DIM-only, fusion via standard ICP (VGICP is not evaluated, as it produces alignment results similar to ICP using a global transformation matrix), and fusion using our proposed method. As shown in Figure 5, the confusion matrices reveal significant differences in class-wise performance. The DIM point cloud suffers from considerable misclassification, especially between “Ground” and “Unclassified” labels, due to inherent noise and lower geometric fidelity. ICP-based fusion improves certain classes but still exhibits confusion between similar surfaces.

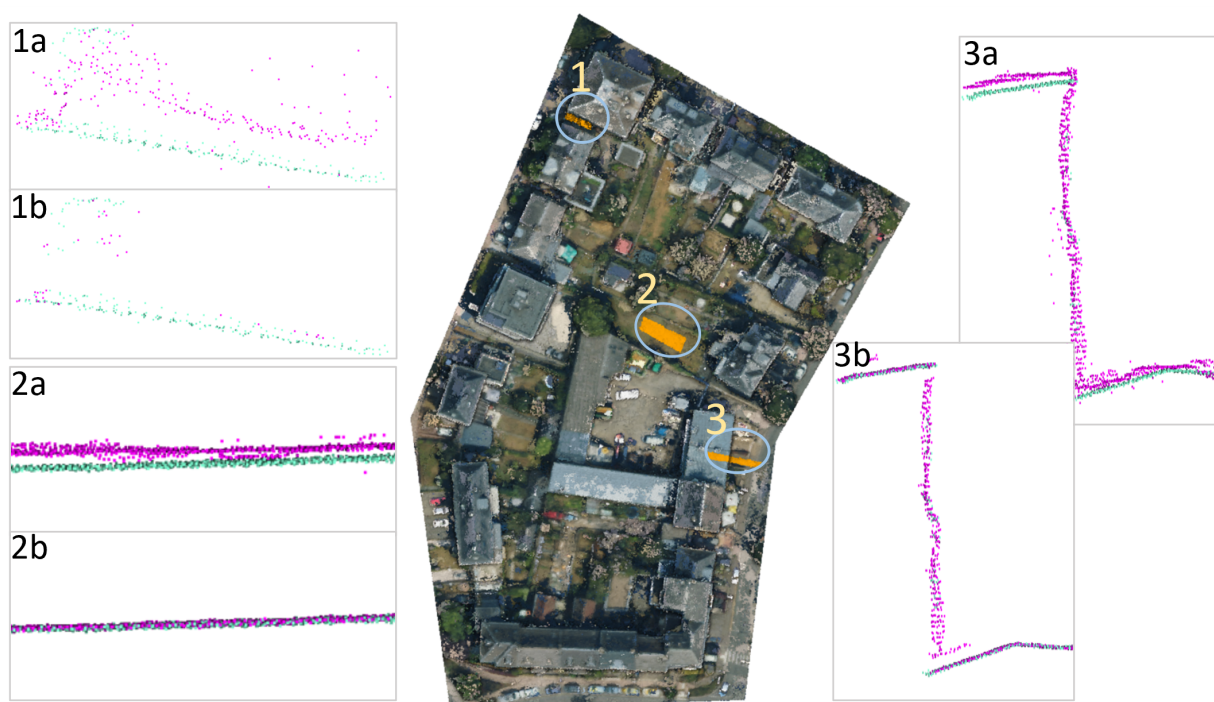


Figure 4. Graphical evaluation across different urban scenarios in Dataset-1. Profiles extracted from the LiDAR (green) and DIM (purple) point clouds are shown before fusion in (1a, 2a, and 3a) and after fusion in (1b, 2b, and 3b), corresponding to (1) a narrow road, (2) an open ground area, and (3) a building facade.

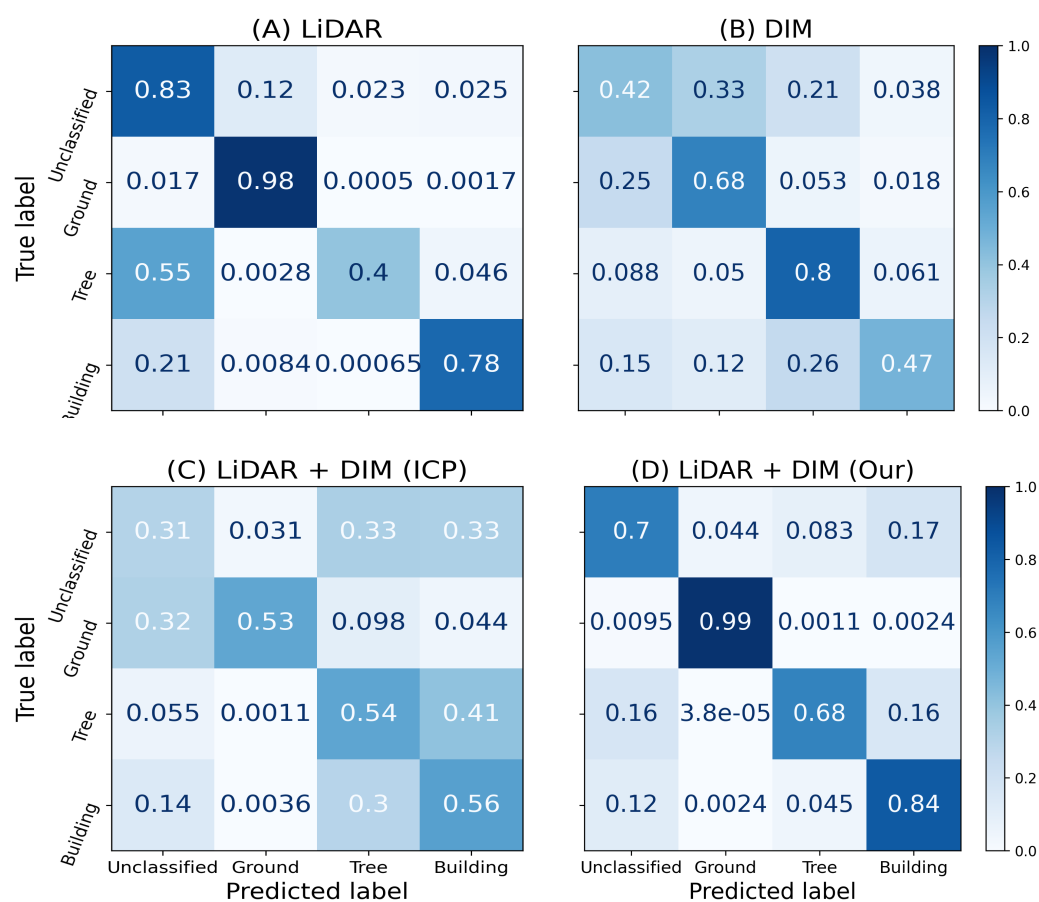


Figure 5. Normalized confusion matrices for semantic segmentation (Dataset-1).

Our proposed fusion method achieves the highest segmentation accuracy overall, with particularly strong performance in the “Ground” and “Building” classes, achieving 98–99% and 84% accuracy, respectively. Moreover, it substantially reduces misclassification in ambiguous classes like “Tree” and “Unclassified.” In addition to the confusion matrix analysis, we include a comparative summary of semantic segmentation metrics (Table 4). The fused point cloud generated by our method outperforms all other inputs in each metric, achieving an OA of 80.1%, mIoU of 68%, and mean F1 score of 80%. These gains highlight the impact of our Plane-Based alignment strategy on enhancing semantic coherence across class boundaries.

To complement the quantitative evaluation, Figure 6 presents semantic segmentation results overlaid on the point clouds for visual inspection. The proposed fusion method results in cleaner object boundaries, reduced class confusion (e.g., between trees and buildings), and more complete surface reconstructions, especially in occluded or low-texture regions. These visual and numerical validations confirm that our fusion method not only improves alignment but also preserves and enhances the semantic integrity of urban environments.

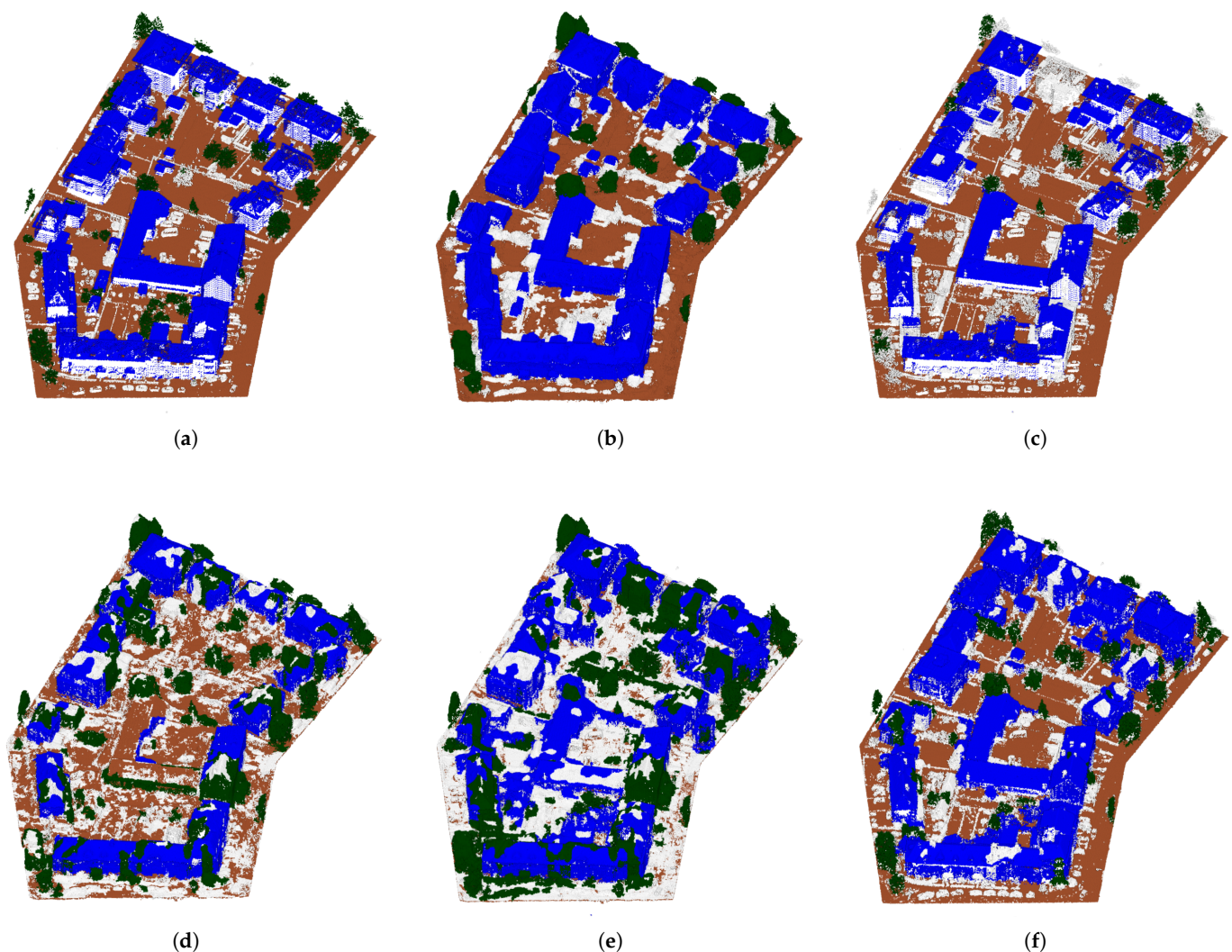


Figure 6. Semantic segmentation results for Dataset-1: (a,b) Ground truth for LiDAR and DIM; (c,d) Predictions for LiDAR and DIM; (e) ICP-based fusion; (f) Proposed fusion. Colors represent the semantic classes: Ground (brown), Building (blue), Trees (green), and Unclassified (white).

Table 4. Overall semantic segmentation accuracy (Dataset-1).

Input	OA (%)	mIoU (%)	mF1 (%)
LiDAR	74.8	61	74
DIM	59.3	42	58
(LiDAR + DIM) ICP	48.5	33	49
(LiDAR + DIM) Ours	80.1	68	80

5.2. Experiment 2

The second experiment dataset, captured at 300 m altitude using a UAV, required a 25 cm distance threshold to compensate for reduced accuracy at higher elevations [17]. Due to the higher LiDAR point density than in the experiment 1 dataset, a 25 cm neighborhood search radius was used to ensure reliable local plane approximation.

5.2.1. Alignment Metrics-Based Evaluation

The algorithm results are summarized in Table 5, which further validate the advantages of Plane-Based fusion. While ICP and VGICP maintain high RMSE and struggle with low fitness metric and overlap ratio, the Line-Based fusion method performs well, achieving a lower RMSE (0.0038), a significant increase in fitness metric, and a much higher overlap ratio. The most notable improvement, however, is observed in the proposed Plane-Based fusion approach, which achieves an even higher fitness metric (0.479) and overlap ratio (0.508), while also establishing over 3.2 million correspondences. These results suggest that Plane-Based fusion is effective in datasets with increased urban density, maintaining robust alignment and high correspondence accuracy across large-scale urban point clouds.

Table 5. Comparison of algorithm accuracy based on alignment metrics (Dataset-2).

Method	RMSE	Fitness	Overlap	Correspondences
ICP	0.0155	0.0014	0.0002	7480
VGICP	0.0154	0.0086	0.0016	54,579
Line-Based	0.0038	0.2293	0.3308	1,537,927
Plane-Based (Our)	0.0073	0.4794	0.5076	3,216,567

For the second dataset, the same profile-based evaluation was conducted as for Experiment 1, focusing on building facades, narrow roads, and open ground areas (Figure 7). Similar trends were observed, with the fused DIM data aligning more accurately with the LiDAR data compared to the original DIM. Structural discrepancies in the building facades were significantly reduced, while road edges and surface continuity improved in the narrow road and open ground profiles. These results further validate the robustness of the proposed fusion approach across different urban environments.

5.2.2. Semantic Segmentation-Based Evaluation

In the second experiment, we repeated the semantic segmentation-based validation to further assess generalizability. As shown in the normalized confusion matrices in Figure 8, our method again achieved superior class separation and reduced misclassifications compared to DIM and ICP-based fusion. Specifically, our fusion output improved tree (92%) and ground (99%) segmentation accuracy, while maintaining strong performance on building and unclassified classes.

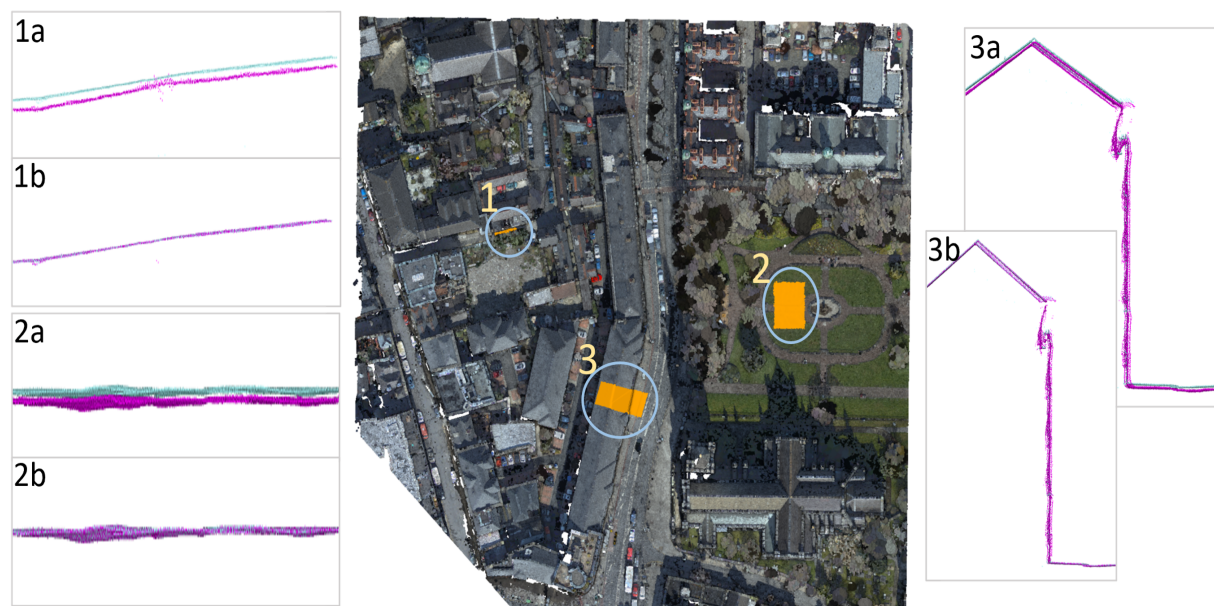


Figure 7. Graphical evaluation across different urban scenarios in Dataset-2. Profiles extracted from the LiDAR (green) and DIM (purple) point clouds are shown before fusion in (1a, 2a, and 3a) and after fusion in (1b, 2b, and 3b), corresponding to (1) a narrow road, (2) an open ground area, and (3) a building facade.

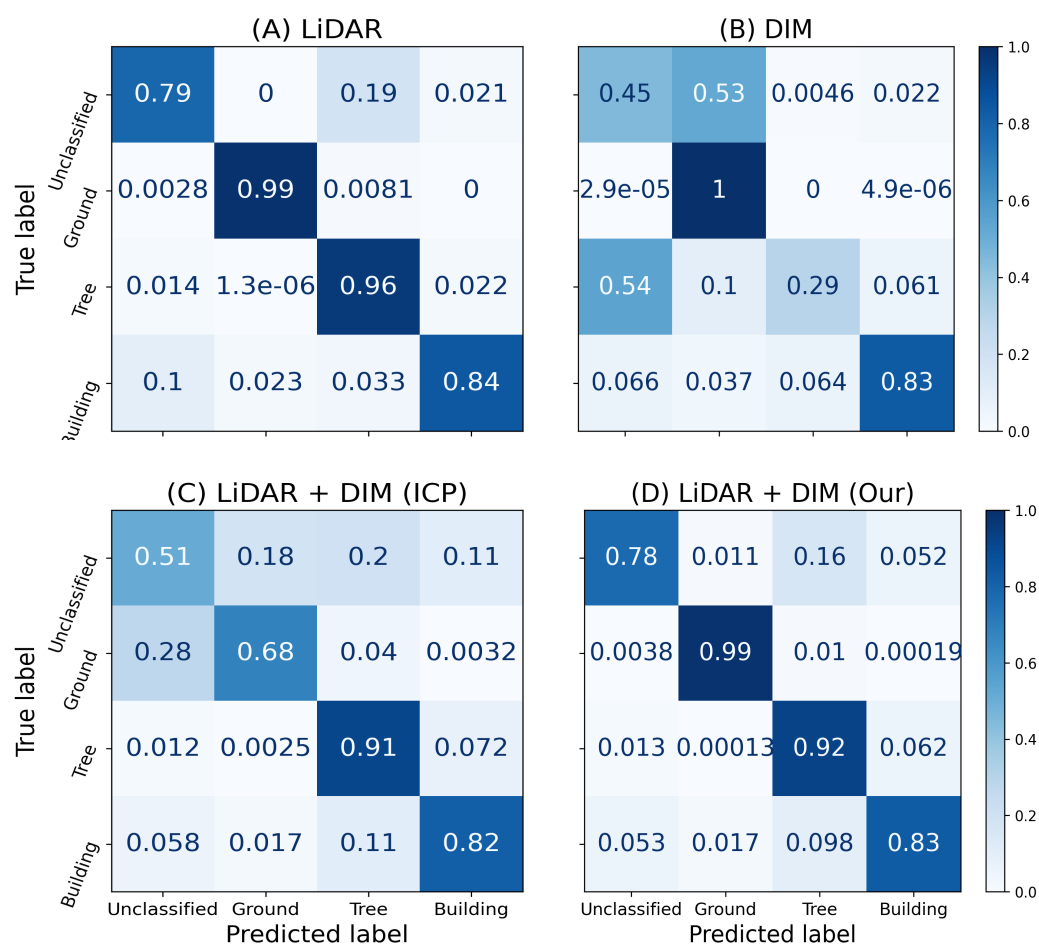


Figure 8. Normalized confusion matrices for semantic segmentation (Dataset-2).

Quantitative metrics reported in Table 6 further reinforce these results. Notably, the proposed method nearly matches LiDAR’s performance ($OA = 88\%$, $mIoU = 79\%$), demonstrating that the fusion strategy effectively combines the strengths of both inputs while mitigating their individual weaknesses. Visual inspection (Figure 9) of semantic overlays confirms fewer fragmented regions and clearer object boundaries, especially in occluded zones and areas with weak texture. These findings indicate that the proposed point-to-plane alignment approach consistently enhances semantic structure and robustness across diverse urban scenarios.

Table 6. Overall semantic segmentation accuracy (Dataset-2).

Input	OA (%)	mIoU (%)	mF1 (%)
LiDAR	89.6	82	89
DIM	64.3	48	62
(LiDAR + DIM) ICP	72.9	58	72
(LiDAR + DIM) Ours	88.0	79	88

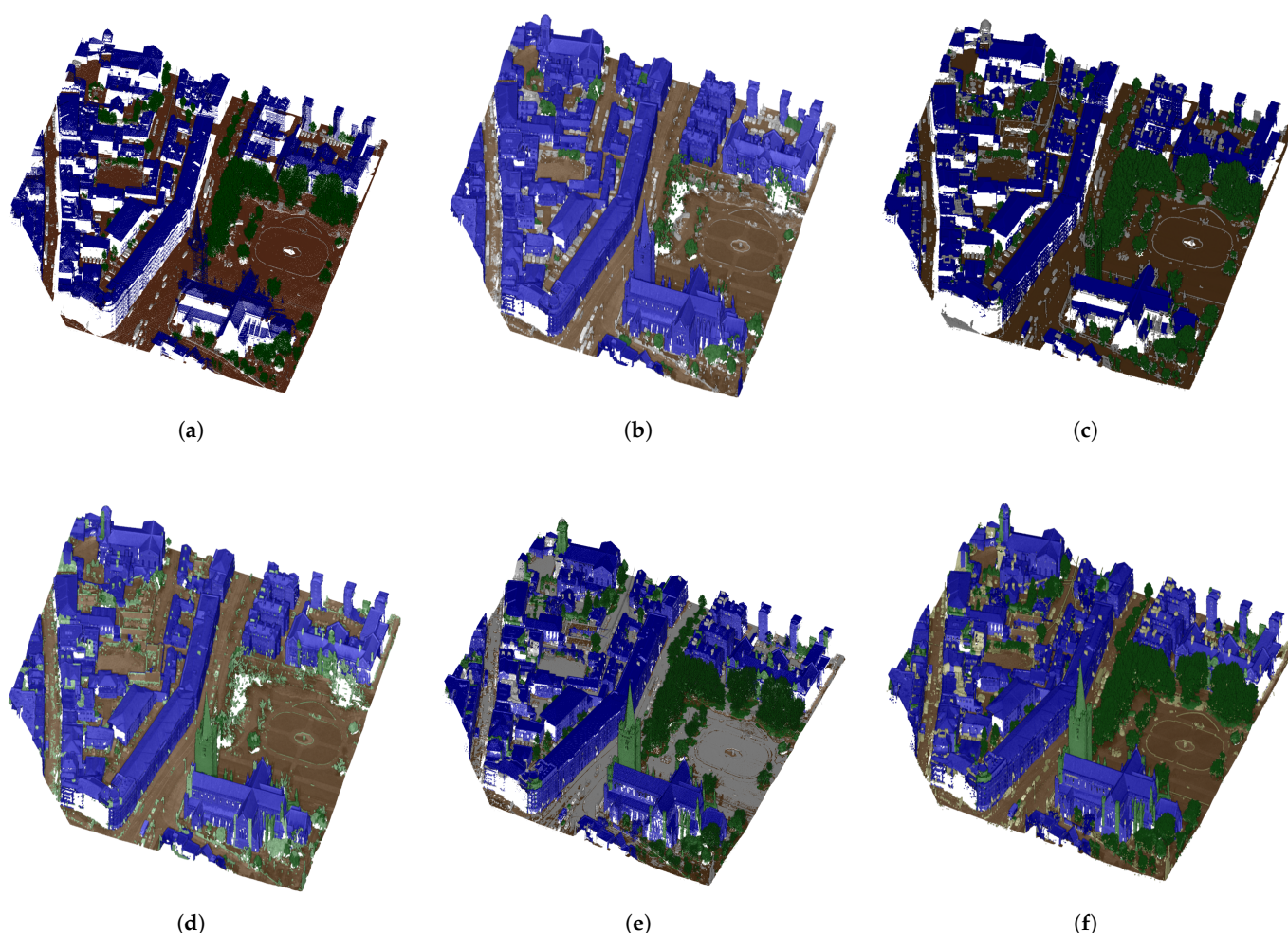


Figure 9. Semantic segmentation results for Dataset-2: (a,b) Ground truth for LiDAR and DIM; (c,d) Predictions for LiDAR and DIM; (e) ICP-based fusion; (f) Proposed fusion. Colors represent the semantic classes: Ground (brown), Building (blue), Trees (green), and Unclassified (gray).

6. Discussion

The comparative results indicate that global rigid registration methods have limited ability to correct spatially varying discrepancies between heterogeneous LiDAR and DIM point clouds. Although the Line-Based method improves alignment by exploiting linear

structures, its effectiveness decreases in irregular regions. In contrast, the proposed method performs point-wise local refinement using PCA-derived target planes, which explains the higher fitness, overlap ratio, and correspondence count observed for both datasets. The profile comparisons further show that the local adjustment improves the geometric consistency of building roofs and façades. This underscores the importance of adopting advanced point-to-plane alignment strategies. The higher computational requirements of the local alignment-based fusion methods stem from neighborhood searching and geometric estimation for each source point; nevertheless, the proposed method was more efficient than the Line-Based baseline on the tested dataset.

The semantic results suggest that improved local geometric consistency contributes to clearer object boundaries and reduces confusion among buildings, trees, and ground surfaces. The improvement is most evident for the building and ground classes, where local geometric refinement produces more stable surface boundaries and reduces class confusion. This gain appears to be driven primarily by improved geometric consistency rather than point density alone, since the denser and more complete DIM-only point cloud does not outperform the LiDAR-only input. The two datasets also reveal complementary behavior: the first dataset presents relatively isolated buildings with more vegetation, and the second presents denser arrangements of attached and multistorey buildings with less vegetation. The consistent improvements across these different urban configurations indicate that the approach can accommodate variations in urban morphology and point cloud density. However, the LiDAR-only input retains the highest overall accuracy in Dataset 2, indicating that fusion does not necessarily outperform the reference LiDAR on every evaluation metric, although it substantially improves the original DIM input in the merged dataset.

The method is most suitable for structured urban surfaces where sufficiently planar target neighborhoods and reliable source normals are available. Its performance may decrease in vegetation, thin structures, sparse neighborhoods, or strongly non-planar regions, where local plane fitting and normal estimation become less reliable. It also assumes adequate global pre-registration and may provide limited improvement when the initial point clouds are already locally consistent. From a practical perspective, the method requires repeated neighborhood searches and local geometric estimation for individual source points, which increases computational cost and memory demand for very large datasets. From a theoretical perspective, the method relies on the assumption that the local target neighborhood can be represented sufficiently well by a plane and that the estimated source normal provides a meaningful direction for local adjustment.

7. Conclusions

We addressed the challenge of airborne cross-source point cloud fusion by proposing local plane constraints and the Plane-Based alignment method, which improves geometric accuracy and thereby enhances semantic consistency across datasets. This approach effectively mitigates alignment errors caused by varying point densities and structural inconsistencies between LiDAR and DIM point clouds. The mathematical foundation of our method ensures that each source point is independently aligned via single-pass local refinement, enabling a non-rigid transformation that adapts to local geometric variations.

Through comprehensive experiments on urban datasets from Luxembourg and Ireland, we have demonstrated the effectiveness of our proposed local plane-constrained fusion strategy in consistently improving point cloud completeness and semantic consistency across datasets. The results showed that the proposed approach effectively mitigates the limitations of DIM point clouds, increasing segmentation OA from 48.5% to 80.1% on the Dudelage dataset and from 72.9% to 88.0% on the Dublin dataset, indicating that

the fused point cloud achieved a level of precision comparable to LiDAR sensors with complete (vertical) surface coverage. In comparison to ICP and Line-Based approaches, our Plane-Based algorithm achieved significantly higher overlap ratios and fitness metrics while maintaining a lower RMSE. The results from different urban environments (including buildings, narrow roads, and open spaces) further validated the adaptability of our approach to complex urban landscapes.

In summary, our Plane-Based fusion method offers a more reliable addition for cross-source point cloud alignment, particularly in urban environments where variations in density and occlusions pose significant challenges. Although the approach assumes locally planar neighborhoods for PCA-based plane estimation, strongly curved or vegetated regions do not introduce severe artifacts in our airborne setting because DIM primarily captures the upper canopy surface, and facade regions are largely unaffected by the algorithm as LiDAR is sparse or absent on facades. The results indicate that leveraging local geometric properties enhances fusion quality, paving the way for more accurate 3D city modeling, digital twin applications, and geospatial data integration. Future research will focus on optimizing computational efficiency and extending the method to the city scale, while incorporating advanced learning-based methods for further refinement.

Author Contributions: S.P.: Conceptualization, Formal Analysis, Investigation, Methodology, Software and Writing—Original Draft. F.N.T.: Data curation, Supervision, Writing—Review and Editing. A.N.: Conceptualization, Writing—Review and Editing. R.L.: Writing—Review and Editing. L.A.L.: Writing—Review and Editing. All authors have read and agreed to the published version of the manuscript.

Funding: This document is one of the results of the PhD research project “Improved Airborne Data Fusion for Advancing Automated 3D City Modelling (DF4CM)”, project No 17042266, supported by the Luxembourg National Research Fund (FNR) under the *Aides à la Formation-Recherche (AFR)* PhD funding scheme.

Data Availability Statement: The datasets used in this research are openly available in [Dublin (<https://archive.nyu.edu/handle/2451/38684>, accessed on 1 September 2022) and Luxembourg (<https://data.public.lu/en/datasets/lidar-2019-releve-3d-du-territoire-luxembourgeois/>, accessed on 10 October 2023)] spatial data repositories. The experimental part will be available upon request. The original contributions on the development of an algorithm presented in this study are included in the article.

Acknowledgments: The authors gratefully acknowledge the Luxembourg National Research Fund (FNR) for supporting this research. The authors also thank the Administration du Cadastre et de la Topographie (ACT) for their contribution to making the airborne imagery and LiDAR datasets available to society.

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

Abbreviations

The following abbreviations are used in this manuscript:

2D	Two-Dimensional
3D	Three-Dimensional
DIM	Dense Image Matching
FPFH	Fast Point Feature Histograms
GNSS	Global Navigation Satellite System
ICP	Iterative Closest Point
IMU	Inertial Measurement Unit
KD-tree	K-Dimensional Tree
LiDAR	Light Detection and Ranging

mF1	Mean F1-Score
mIoU	Mean Intersection over Union
OA	Overall Accuracy
PCA	Principal Component Analysis
RANSAC	Random Sample Consensus
RMSE	Root Mean Square Error
UAV	Unmanned Aerial Vehicle
VGICP	Voxelized Generalized Iterative Closest Point

References

1. Zhang, R.; Candra, S.A.; Vetter, K.; Zakhori, A. Sensor fusion for semantic segmentation of urban scenes. In Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26–30 May 2015; pp. 1850–1857. [\[CrossRef\]](#)
2. Nurunnabi, A.; Teferle, F.N.; Li, J.; Lindenbergh, R.C.; Parvaz, S. Investigation of PointNet for semantic segmentation of large-scale outdoor point clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2021**, *XLVI-4/W5-2021*, 397–404. [\[CrossRef\]](#)
3. Charles, R.Q.; Su, H.; Kaichun, M.; Guibas, L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 77–85. [\[CrossRef\]](#)
4. Mandlbürger, G.; Wenzel, K.; Spitzer, A.; Haala, N.; Glira, P.; Pfeifer, N. Improved topographic models via concurrent airborne LiDAR and dense image matching. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *IV-2/W4*, 259–266. [\[CrossRef\]](#)
5. Moe, K.; Toschi, I.; Poli, D.; Lago, F.; Schreiner, C.; Legat, K.; Remondino, F. Changing the production pipeline—Use of oblique aerial cameras for mapping purposes. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B4*, 631–637. [\[CrossRef\]](#)
6. Yang, W.; Liu, Y.; He, H.; Lin, H.; Qiu, G.; Guo, L. Airborne LiDAR and photogrammetric point cloud fusion for extraction of urban tree metrics according to street network segmentation. *IEEE Access* **2021**, *9*, 97834–97842. [\[CrossRef\]](#)
7. Liu, Y.; Obukhov, A.; Wegner, J.D.; Schindler, K. Point2Building: Reconstructing buildings from airborne LiDAR point clouds. *ISPRS J. Photogramm. Remote Sens.* **2024**, *215*, 351–368. [\[CrossRef\]](#)
8. Haala, N.; Kölle, M.; Cramer, M.; Laupheimer, D.; Mandlbürger, G.; Glira, P. Hybrid georeferencing, enhancement and classification of ultra-high resolution UAV LiDAR and image point cloud for monitoring applications. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *V-2-2020*, 727–734. [\[CrossRef\]](#)
9. Zhou, K. Combining LiDAR and Photogrammetry to Generate Up-to-Date 3D City Models. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 2020. [\[CrossRef\]](#)
10. Parvaz, S.; Teferle, F.N.; Nurunnabi, A. Towards Precise Building Models: LOD Generation from Airborne Multi-Source Point Clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2025**, *XLVIII-1/W6-2025*, 183–189. [\[CrossRef\]](#)
11. Zhang, Z.; Vosselman, G.; Gerke, M.; Persello, C.; Tuia, D.; Yang, M.Y. Detecting building changes between airborne laser scanning and photogrammetric data. *Remote Sens.* **2019**, *11*, 2417. [\[CrossRef\]](#)
12. Toschi, I. Airborne oblique imaging: Towards the hybrid era. *Arch. Photogramm. Cartogr. Remote Sens.* **2019**, *31*, 21–28. [\[CrossRef\]](#)
13. Toschi, I.; Farella, E.; Welpöner, M.; Remondino, F. Quality-based registration refinement of airborne LiDAR and photogrammetric point clouds. *ISPRS J. Photogramm. Remote Sens.* **2021**, *172*, 160–170. [\[CrossRef\]](#)
14. Bracci, F.; Drauschke, M.; Kühne, S.; Márton, Z.C. Challenges in fusion of heterogeneous point clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-2*, 155–162. [\[CrossRef\]](#)
15. Li, P.; Wang, R.; Wang, Y.; Tao, W. Evaluation of the ICP algorithm in 3D point cloud registration. *IEEE Access* **2020**, *8*, 68030–68048. [\[CrossRef\]](#)
16. Huang, X.; Mei, G.; Zhang, J.; Abbas, R. A comprehensive survey on point cloud registration. *arXiv* **2021**, arXiv:2103.02690. [\[CrossRef\]](#)
17. Parvaz, S.; Teferle, F.; Nurunnabi, A. Airborne Cross-Source Point Clouds Fusion by Slice-to-Slice Adjustment. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2024**, *X-4/W4-2024*, 161–168. [\[CrossRef\]](#)
18. Li, Z.; Jin, F.; Wang, J.; Zhang, Z.; Zhu, L.; Sun, W.; Chen, X. Adaptive fusion of different platform point cloud with improved particle swarm optimization and supervoxels. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *130*, 103934. [\[CrossRef\]](#)
19. Glira, P.; Pfeifer, N.; Mandlbürger, G. Hybrid orientation of airborne LiDAR point cloud and aerial images. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *IV-2/W5*, 567–574. [\[CrossRef\]](#)
20. Wang, Y.; Bu, S.; Chen, L.; Dong, Y.; Li, K.; Cao, X.; Li, K.; Jin, J. HybridFusion: LiDAR and Vision Cross-Source Point Cloud Fusion. *IEEE Robot. Autom. Lett.* **2024**, *9*, 1139–1146. [\[CrossRef\]](#)
21. Chang, S.; Ahn, C.; Lee, M.; Oh, S. Graph-matching-based correspondence search for nonrigid point cloud registration. *Comput. Vis. Image Underst.* **2020**, *192*, 102899. [\[CrossRef\]](#)

22. Luo, J.; Yuan, M.; Fu, K.; Wang, M.; Zhang, C. Deep Graph Matching Based Dense Correspondence Learning Between Non-Rigid Point Clouds. *IEEE Robot. Autom. Lett.* **2022**, *7*, 5842–5849. [\[CrossRef\]](#)
23. Besl, P.; McKay, N.D. A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 239–256. [\[CrossRef\]](#)
24. Koide, K.; Yokozuka, M.; Oishi, S.; Banno, A. Voxelized GICP for fast and accurate 3D point cloud registration. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 30 May–5 June 2021; pp. 11054–11059. [\[CrossRef\]](#)
25. Rusu, R.B.; Blodow, N.; Beetz, M. Fast Point Feature Histograms (FPFH) for 3D registration. In Proceedings of the 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, 12–17 May 2009; pp. 3212–3217. [\[CrossRef\]](#)
26. Huang, X.; Mei, G.; Zhang, J. Feature-Metric Registration: A Fast Semi-Supervised Approach for Robust Point Cloud Registration Without Correspondences. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos, CA, USA, 13–19 June 2020; pp. 11363–11371. [\[CrossRef\]](#)
27. Huang, X.; Zhang, J.; Fan, L.; Wu, Q.; Yuan, C. A systematic approach for cross-Source point cloud registration by preserving macro and micro structures. *IEEE Trans. Image Process.* **2017**, *26*, 3261–3276. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Aoki, Y.; Goforth, H.; Srivatsan, R.A.; Lucey, S. PointNetLK: Robust & Efficient Point Cloud Registration Using PointNet. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos, CA, USA, 15–20 June 2019; pp. 7156–7165. [\[CrossRef\]](#)
29. Zhang, Z.; Dai, Y.; Sun, J. Deep learning based point cloud registration: An overview. *Virtual Real. Intell. Hardw.* **2020**, *2*, 222–246. [\[CrossRef\]](#)
30. Yu, H.; Li, F.; Saleh, M.; Busam, B.; Ilic, S. CoFiNet: Reliable Coarse-to-fine Correspondences for Robust Point Cloud Registration. *arXiv* **2021**, arXiv:2110.14076. [\[CrossRef\]](#)
31. Huang, X.; Mei, G.; Zhang, J. Cross-source point cloud registration: Challenges, progress and prospects. *Neurocomputing* **2023**, *548*, 126383. [\[CrossRef\]](#)
32. Abdelazeem, M.; Elamin, A.; Afifi, A.; El-Rabbany, A. Multi-sensor point cloud data fusion for precise 3D mapping. *Egypt. J. Remote Sens. Space Sci.* **2021**, *24*, 835–844. [\[CrossRef\]](#)
33. Peng, S.; Ma, H.; Zhang, L. Automatic registration of optical images with airborne LiDAR point cloud in urban scenes based on line-point similarity invariant and extended collinearity equations. *Sensors* **2019**, *19*, 1086. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Li, S.; Ge, X.; Li, S.; Xu, B.; Wang, Z. Linear-Based Incremental Co-Registration of MLS and Photogrammetric Point Clouds. *Remote Sens.* **2021**, *13*, 2195. [\[CrossRef\]](#)
35. Li, L.; Wang, R.; Zhang, X. A Tutorial Review on Point Cloud Registrations: Principle, Classification, Comparison, and Technology Challenges. *Math. Probl. Eng.* **2021**, *2021*, 9953910. [\[CrossRef\]](#)
36. Remondino, F.; Spera, M.G.; Nocerino, E.; Menna, F.; Nex, F. State of the art in high density image matching. *Photogramm. Rec.* **2014**, *29*, 144–166. [\[CrossRef\]](#)
37. Zhou, K.; Lindenbergh, R.; Gorte, B.; Zlatanova, S. LiDAR-guided dense matching for detecting changes and updating of buildings in Airborne LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 200–213. [\[CrossRef\]](#)
38. Nurunnabi, A.; West, G.; Belton, D. Outlier detection and robust normal-curvature estimation in mobile laser scanning 3D point cloud data. *Pattern Recognit.* **2015**, *48*, 1404–1419. [\[CrossRef\]](#)
39. Nurunnabi, A.; Belton, D.; West, G. Robust statistical approaches for local planar surface fitting in 3D laser scanning data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 106–122. [\[CrossRef\]](#)
40. Gong, Z.; Lin, H.; Zhang, D.; Luo, Z.; Zelek, J.; Chen, Y.; Nurunnabi, A.; Wang, C.; Li, J. A Frustum-based probabilistic framework for 3D object detection by fusion of LiDAR and camera data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 90–100. [\[CrossRef\]](#)
41. Yuan, X.; Maharjan, A. Non-rigid point set registration: Recent trends and challenges. *Artif. Intell. Rev.* **2023**, *56*, 4859–4891. [\[CrossRef\]](#)
42. Sun, Y.; Fu, Z.; Sun, C.; Hu, Y.; Zhang, S. Deep Multimodal Fusion Network for Semantic Segmentation Using Remote Sensing Image and LiDAR Data. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5404418. [\[CrossRef\]](#)
43. Kölle, M.; Laupheimer, D.; Schmohl, S.; Haala, N.; Rottensteiner, F.; Wegner, J.D.; 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 J. Photogramm. Remote Sens.* **2021**, *1*, 100001. [\[CrossRef\]](#)
44. Duran, Z.; Ozcan, K.; Atik, M.E. Classification of Photogrammetric and Airborne LiDAR Point Clouds Using Machine Learning Algorithms. *Drones* **2021**, *5*, 104. [\[CrossRef\]](#)
45. Laefer, D.F.; Abuwarda, S.; Vo, A.V.; Truong-Hong, L.; Gharibi, H. *2015 Aerial Laser and Photogrammetry Datasets for Dublin, Ireland's City Center*; New York University: New York, NY, USA, 2017. [\[CrossRef\]](#)

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