Framework for the Georeferencing and Processing of BikePack LiDAR Data for Urban Tree Mapping

Chunxi Zhao[1], Sungwoong Hyung[1], Raja Manish[1], Sang-Yeop Shin[1], Sangyoon Park[1], Songlin Fei[2], and Ayman Habib[1]

[1] Lyles School of Civil and Construction Engineering, Purdue University, West Lafayette, USA – (zhao1259, hyungs, rmanish, shin337, park1858, ahabib)@purdue.edu
[2] Department of Forestry and Natural Resources, Purdue University, West Lafayette, USA – sfei@purdue.edu

Keywords: BikePack LiDAR, Trajectory Enhancement, Super Point Transformer, Urban Trees

Abstract

Wearable backpack Light Detection and Ranging (LiDAR) systems have been widely used for high-resolution data collection in urban environments. However, they are often limited by the operator's mobility and time required for complete coverage of the area of interest. This paper introduces a BikePack LiDAR system, which uses a bicycle for efficient urban data acquisition. Despite its advantages, the system face challenges of intermittent Global Navigation Satellite System (GNSS) signal availability due to dense, tall buildings and other objects in urban environments. The study proposes a framework to enhance the trajectory and mapping results for the BikePack LiDAR in GNSS-challenging urban areas. The proposed framework offers an option to incorporate airborne LiDAR data to improve the absolute georeferencing accuracy of the derived point cloud, enabling the integration of the two sources – terrestrial and airborne, to produce a comprehensive 3D map of the urban environment. Following the enhancement, this study also demonstrates a learning strategy for isolating vegetation from other man-made and natural objects. Based on the well-aligned terrestrial BikePack and airborne Geiger-mode LiDAR data, a deep learning strategy is applied to the latter, with derived semantic segmentation results transferred to the BikePack point clouds through a cross-labelling process. The experimental results show that the proposed trajectory enhancement strategy can significantly improve the relative and absolute accuracy of the BikePack point cloud with the assistance of Geiger-mode airborne LiDAR data, achieving planimetric and vertical trajectory adjustments of 0.36 m and 0.27 m, respectively. Furthermore, the semantic segmentation results show that the proposed cross-labelling strategy outperforms other methods, improving overall accuracy by approximately 16%, and increasing the mean Intersection over Union (IoU) and Cohen's Kappa score by 0.17 and 0.24, respectively.

1. Introduction

Urban trees provide essential ecosystem services such as air pollution mitigation and carbon sequestration, supporting the well-being of urban population. A recent study estimates that urban land in the US covers more than 68 million acres and is growing at a rate of nearly 10 million acres per decade (Nowak et al., 2018). To maintain the desired tree density, an estimated 31 million trees need to be planted annually in urban areas (American Forests, 2025). Hence, urban tree mapping is essential for sustainable city planning and environmental monitoring (Wallace et al., 2021). Current approaches to map urban trees are mainly manual based, which are expensive and time consuming. Applications of proximal mapping systems, such as terrestrial laser scanners and wearable backpack LiDAR systems, have been explored to automate the inventory process. However, these proximal systems are often limited by the operator's mobility and time required for complete coverage of the area of interest. Vehicle-based mapping systems enable rapid data acquisition but cannot access narrow streets or pedestrian pathways, limiting their effectiveness in dense urban environments. This paper introduces a BikePack system comprising a LiDAR sensor directly georeferenced by an integrated Global Navigation Satellite System/Inertial Navigation System (GNSS/INS) unit. The portability of the system allows for its deployment as a backpack pedestrian platform or onboard a manual or powered bicycle for efficient urban data acquisition. Despite the advantages of mobility and data acquisition efficiency, the system faces significant processing challenges. For example, GNSS signal obstructions due to buildings and dense vegetation can impact the quality of the trajectory, resulting in inaccurate point clouds and mapping results (Nagai et al., 2020).

To improve trajectory and mapping results in urban environments, Simultaneous Localization and Mapping (SLAM) offers an option for accurate mapping in such environments. Liu et al. (2019) improved localization accuracy using semantic feature extraction for urban environments, where the features include ground, road-curbs, edges, and surfaces. Li et al. (2021) proposed a semantic point cloud registration and integrated a semantic graph place recognition method within a loop closure detection module to improve localization accuracy. Du et al. (2021) removed dynamic objects from LiDAR scans, then extracted and incorporated semantic information into odometry and mapping threads to improve localization and mapping results. However, these strategies have difficulties in extracting sufficient semantic features in complex urban scenarios, resulting in compromised mapping results.

To overcome this limitation, Rufus et al. (2020) proposed a real-time odometry and mapping approach. They calculate approximate pose parameters, which are used for a point-to-plane Iterative Closest Point (ICP) to refine the matching and transformation parameters among successive scans. Shan et al. (2020) proposed a framework for a tightly-coupled LiDAR inertial odometry and mapping. This strategy combines LiDAR and inertial data using a factor graph for real-time state estimation and mapping, reducing drift and improving localization accuracy.

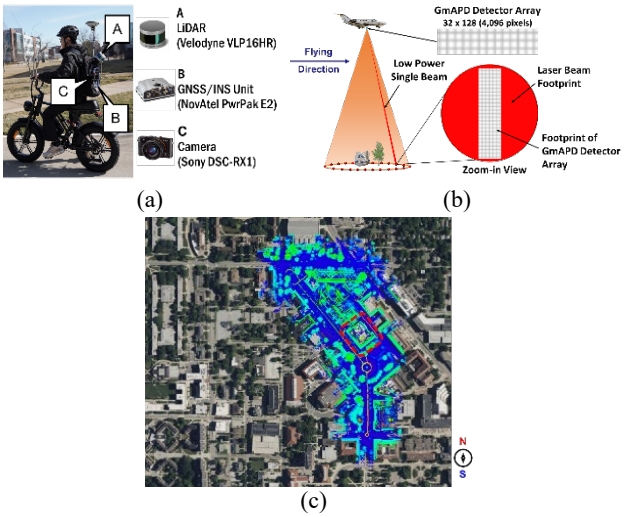
Ramezani et al. (2022) introduced a LiDAR-inertial SLAM to extract and match surface elements from sliding-window scans. The authors reported that their proposed method can efficiently map environments through pose graph optimization by incorporating Inertial Measurement Unit (IMU) and LiDAR data. Despite these advances, none of the methods can incorporate other geospatial data, which prevents the integration of multi-platform/multi-temporal LiDAR point clouds.

Beyond improving the quality of derived point clouds, 3D data classification and segmentation are needed to extract meaningful information for urban tree mapping. Recent advancements in deep-learning approaches have introduced novel architectures that significantly enhance classification and segmentation performance, enabling robust and scalable processing of 3D data (Sarker et al., 2024). Landrieu and Simonovsky (2018) proposed Superpoint Graph (SPG), which segments point clouds into geometric superpoints and uses graph convolutions to capture contextual relationships. This approach enhances segmentation accuracy, particularly in large-scale outdoor LiDAR data, by modelling long-range dependencies. Robert et al. (2023) introduced Superpoint Transformer (SPT), a resource-efficient model that partitions point clouds into hierarchical superpoints and employs sparse self-attention for improved segmentation. SPT achieves superior accuracy while significantly reducing model size and training time, making it well-suited for scalable LiDAR processing. Wu et al. (2024) introduced Point Transformer V3 (PTv3), which improves point cloud segmentation by replacing the k-Nearest Neighbours (KNN) with a serialized neighbourhood mapping, expanding the receptive field while reducing computational complexity. PTv3 achieves state-of-the-art performance across 3D perception tasks, making it highly relevant for LiDAR applications. These learning strategies rely heavily on training data, making them sensitive to variations in dataset characteristics. For point cloud segmentation, if the dataset used to train the model has different characteristics from the dataset being processed, prediction quality may suffer, often requiring retraining from scratch (Saltori et al., 2023). Transfer learning can help mitigate this limitation by adapting a pretrained model to a new dataset with different characteristics, reducing the need for retraining from scratch. However, it does not fully resolve the issue, as differences in point cloud properties – such as density, noise levels, or sensor types – may still lead to suboptimal segmentation performance (Williams et al., 2020). In such cases, further fine-tuning or additional domain-specific training data may be required to achieve high accuracy.

To address the aforementioned limitations, the presented study proposes a General Trajectory Enhancement and Mapping (G-TEAM) framework to improve the trajectory and mapping results using an in-house developed BikePack LiDAR system operating in GNSS-challenging urban areas. The proposed framework offers an option to incorporate airborne LiDAR and legacy elevation data to improve the absolute georeferencing accuracy of derived point clouds, enabling the integration of multi-platform LiDAR data and comprehensive mapping of urban environments. Following the enhancement, this study also demonstrates a point cloud segmentation method (referred to as cross-labelling) for isolating vegetation from other objects such as buildings, streetscapes, vehicles, and utility poles. The remainder of this paper is structured as follows: Section 2 introduces data acquisition systems and datasets used in this study; Section 3 covers the background of point cloud reconstruction followed by a detailed explanation of the proposed G-TEAM and semantic segmentation framework; Section 4 presents results from the proposed framework and assesses its performance; and lastly, Section 5 summarizes the findings of the research and provides recommendations for future work.

2. Data Acquisition Systems and Datasets Description

BikePack and Geiger-mode LiDAR systems have been used in this study, as shown in Figure 1. The BikePack system comprises a Velodyne VLP-16 Hi-Res LiDAR, a Sony DSC-RX1 camera, and a NovAtel PwrPak E2 GNSS/INS unit. A BikePack dataset was collected at the Purdue University campus in West Lafayette, IN, USA in January 2024. This dataset covered an area of 700 × 400 m, visualized in Figure 1c, at a speed of 5 m/s. The Geiger-mode LiDAR data was acquired in December 2021 over the same area. BikePack data captures detailed information about the lower parts of objects in urban environments but encounters misalignment issues due to GNSS signal outages. However, Geiger-mode data offers high georeferencing accuracy but struggles to capture detailed object structures.



(a) (b)

(c)

Figure 1. Mapping systems used in this study and the location surveyed: (a) BikePack LiDAR, (b) Geiger-mode airborne LiDAR, and (c) study site covered by LiDAR data (region of interest is highlighted by a red box).

3. Methodology

LiDAR point cloud reconstruction is based on the point positioning equation, as described in Equation (1) with a graphical illustration in Figure 2. Laser ranges and beam orientation are used to derive the position ($r_I^{lu(t)}$) of the laser beam footprint at an object point I captured at time t relative to the laser unit frame (lu). The pose parameters of the laser unit frame at the time t, comprising its position and orientation ($r_{lu(t)}^m/R_{lu(t)}^m$), are used to evaluate the 3D coordinates of the object point I in the mapping frame (m), represented by r_I^m. Based on the equation, accurate pose parameters ($r_{lu(t)}^m/R_{lu(t)}^m$) directly derived from the GNSS/INS trajectory are crucial to derive accurate point cloud coordinates in the mapping frame (r_I^m). In scenarios with GNSS signal outages, it may not be possible to achieve the desired level of accuracy through direct georeferencing. The proposed framework aims to improve trajectory and mapping results for the BikePack system, and subsequently conduct an efficient semantic segmentation for urban tree mapping. Figure 3 outlines the workflows of the proposed approaches, which comprises two key steps: G-TEAM and SPT segmentation. The following subsections discuss these steps in detail.

$$r_I^m = r_{lu(t)}^m + R_{lu(t)}^m r_I^{lu(t)} \quad (1)$$

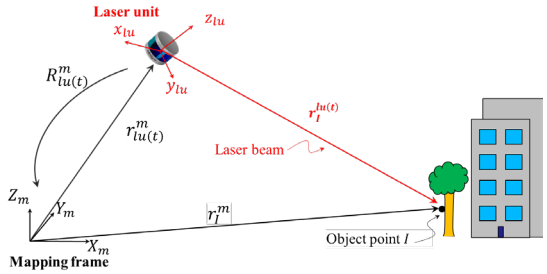


Figure 2. Parameters and coordinate systems for LiDAR point cloud reconstruction.

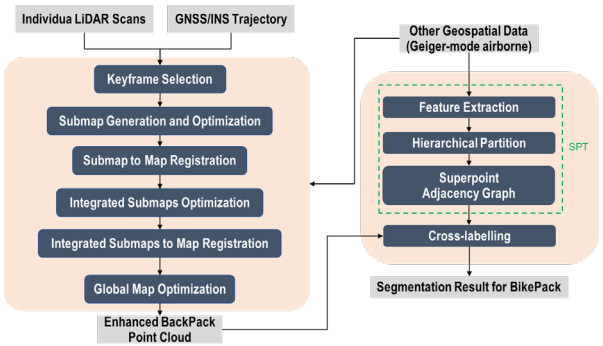


Figure 3. Workflow of the proposed approaches: G-TEAM and SPT segmentation.

3.1 G-TEAM

Using available LiDAR scans and GNSS/INS trajectory information, the G-TEAM framework starts with keyframe selection. Within a sequence of successive scans, keyframes are scans where there is a significant pose change exceeding a certain threshold. In challenging environments with slow data collection, the pose change threshold is lowered to include more keyframes, ensuring stable and accurate mapping. The keyframe selection is followed by several layers of map optimization that use surface elements (surfels) as primitives. Map optimization starts at a submap level by combining multiple successive keyframes. The first optimized submap is used to create an initial map to which subsequent submaps are sequentially registered. Similar to the submap optimization, each integrated submap, combining multiple submaps, is optimized and registered with a global map, which consists of optimized and registered submaps. Lastly, a global map optimization further improves point cloud alignment and mapping quality. As an option, existing geospatial LiDAR data can be incorporated into the framework for multi-platform data integration.

3.1.1 Surfel-based Optimization

The surfel-based optimization is performed within the submap, integrated submap, and global map optimization layers. The workflow of the optimization is shown in Figure 4. Surfels are generated by voxelizing point clouds and defining points within each voxel as a surfel. The planarity measure of a surfel is derived using eigenvalue analysis of a variance-covariance matrix describing the 3D spread of points within the corresponding voxel (Demantké et al., 2011). Surfels with a planarity value less than a threshold (e.g., 0.3) are removed. A non-linear least squares adjustment (LSA) is conducted to minimize the sum of normal distances between points and the corresponding surfel in each voxel. Pose and surfel parameters are refined during the optimization. The refined pose parameters are then used to update the point cloud. The updated point cloud is then used as input to iteratively conduct the surfel-based optimization. The voxel size gradually decreases within the iterative process, and the process stops when it falls below a threshold (e.g., 0.2 m). The pose parameters and point cloud are optimized throughout the iterative procedure. A sample tree before and after applying the surfel-based optimization for a submap, is shown in Figure 5.

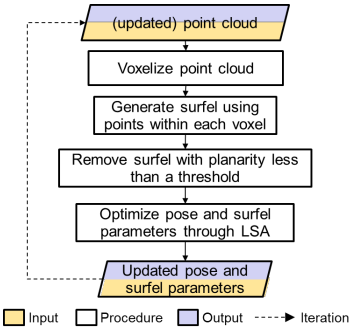
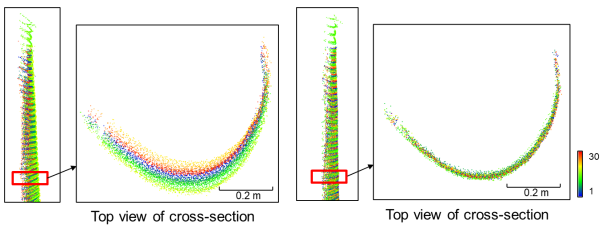


Figure 4. Workflow of the surfel-based optimization.



(a) (b)

Figure 5. A sample tree (colored by keyframe ID) from a submap (a) before and (b) after applying surfel-based optimization.

3.1.2 Map Registration

As shown in Figure 3, submaps and integrated submaps are registered with the global map through the respective optimization steps. The workflow of the registration is shown in Figure 6. The reference and target maps are denoted as M_{ref} and M_{target}. In this study, M_{target} can be a submap or an integrated submap, M_{ref} is a subset of the global map. Points from the global map within the 2D bounding box of M_{target} are established as M_{ref}. First, the reference and target point clouds are used to generate corresponding surfel maps SM_{ref} and SM_{target}. Next, surfel correspondences are identified based on the distance between the two surfel maps. Point-to-surfel constraints are established to optimize the pose parameters of the target point cloud through a non-linear LSA. The target surfel map is updated using the optimized pose parameters. The process iterates until a prespecified number of iterations has been reached. The pose parameters of the target map are changed through the iterative process.

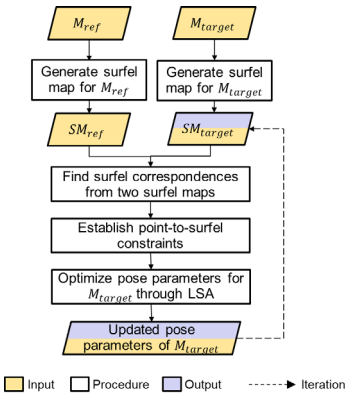
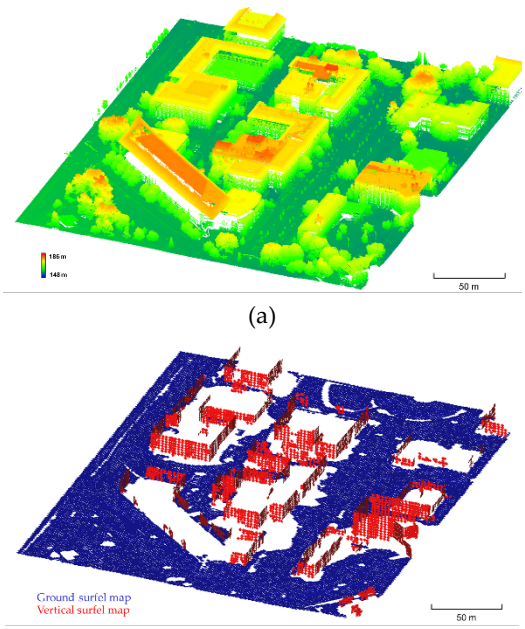


Figure 6. Workflow of map registration between submap or integrated submap and global map.

3.1.3 Incorporation of Existing Geospatial Data

The proposed G-TEAM provides an option to incorporate other geospatial data to improve the georeferencing accuracy through the integration of multi-platform LiDAR data. First, bare earth (BE) and above-ground (AG) points are identified in the supplemental geospatial data (e.g., Geiger-mode data) using a modified cloth simulation algorithm (Lin et al., 2021). Both the BE and AG point clouds are voxelized and surfel maps (S_{BE} and S_{AG}) are generated using points in each voxel. Surfels with normal vectors that are approximately perpendicular to the XY plane are identified as ground surfels, S_{BE}; whereas those with normal vectors parallel to the XY plane are classified as vertical surfels, S_{AG}. The ground and vertical surfel maps for a sample Geiger-mode LiDAR data are shown in Figure 7. Ground and vertical surfels are also extracted from the BackPack point cloud and matched with corresponding surfel maps from Geiger-mode data. Then, point-to-surfel constraints are established and incorporated into surfel-based optimization within the different optimization layers to improve the georeferencing accuracy of BikePack data.



(a)

(b)

Figure 7. Sample (a) Geiger-mode point cloud (colored by height) and (b) corresponding ground (in blue) and vertical (in red) surfel maps.

3.2 Point Cloud Segmentation

3.2.1 Superpoint Transformer

The SPT is a superpoint-based architecture specifically designed for efficient semantic segmentation of large-scale 3D point clouds. The method begins by partitioning the input point cloud using a hierarchical superpoint structure, where geometrically homogeneous superpoints are created across multiple scales. This hierarchical partitioning is achieved through a computationally efficient energy minimization approach, leveraging adjacency graphs to encode relationships among superpoints. Inspired by U-Net, the SPT architecture comprises an encoder-decoder structure (Ronneberger et al., 2015), where the encoder aggregates features across hierarchical superpoints and the decoder refines these features for classification. At each stage, the model employs self-attention mechanisms to propagate context between superpoints within the same hierarchical level, capturing both local and long-range interactions. The model outputs semantic labels for superpoints instead of individual points, significantly improving computational efficiency. Thanks to these characteristics, SPT is highly practical for large-scale datasets as it efficiently processes massive point clouds by grouping points into hierarchical superpoints, reducing both memory usage and computational costs.

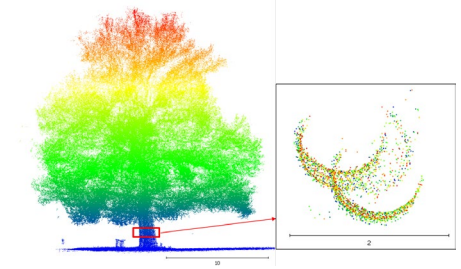
3.2.2 Cross-labelling Strategy for Segmentation

In this study, a cross-labelling strategy is proposed. Given that the Geiger-mode LiDAR and BikePack datasets are spatially well-aligned, it is possible to leverage the segmented Geiger-mode data as a reference for cross-labelling the BikePack point cloud. This approach helps in segmenting the BackPack point cloud by transferring reliable class labels from the Geiger-mode dataset. The cross-labelling procedure is consisted of the following steps. First, a pretrained SPT is applied to the Geiger-mode dataset, producing a set of labelled points. Then, cross-labelling using the KNN algorithm is performed to transfer labels from the Geiger-mode dataset to BikePack point cloud. Taking advantage of the well-aligned Geiger-mode and BikePack datasets, neighbouring points in the BikePack point cloud are identified for each point in the Geiger-mode data and assigned with the same label.

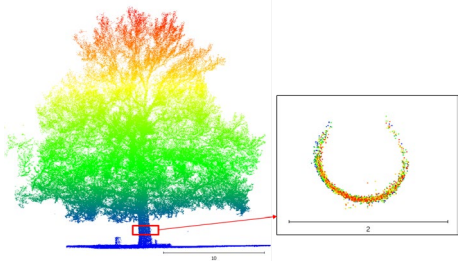
4. Experimental Results

To evaluate the proposed framework, the BikePack and Geiger-mode datasets collected at the study site were processed using G-TEAM. Figure 8 presents sample tree trunk cross sections taken at an elevation of 1.4 m from ground before and after applying the G-TEAM algorithm. The quality of point cloud alignment is greatly improved when comparing the enhanced point cloud with that reconstructed using the initial GNSS/INS trajectory. Figure 9a depicts the alignment of Geiger-mode point cloud with that from the BackPack derived solely by the G-TEAM showing a 0.7 m elevation discrepancy and a 0.4 m planimetric difference for the sample building. On the other hand, Figure 9b highlights a noticeable improvement in alignment achieved by incorporating the Geiger-mode data into G-TEAM. Another sample of the Geiger-mode-assisted G-TEAM, Geiger-mode, and combined point clouds are shown in Figure 10a, 10b and 10c, respectively. Through a comparison of these datasets, one can notice that the combined point cloud takes advantage of ground and airborne LiDAR data, providing comprehensive details of existing structures based on well-aligned datasets. The statistics of trajectory adjustments using Geiger-mode assisted G-TEAM are

presented in Table 1, where the mean, Standard Deviation (STD), and Root Mean Squared (RMS) of adjustment values for the trajectory comprising 4,030 sets of pose parameters are reported. The Geiger-mode assisted G-TEAM results in an adjustment of around 0.36 m for X and Y directions, along with a vertical change of approximately 0.27 m for the trajectory. For the orientation parameters of the trajectory, an adjustment of around 0.1 degrees is applied for roll and pitch angles, while the heading angle shows a change of 0.2 degrees.

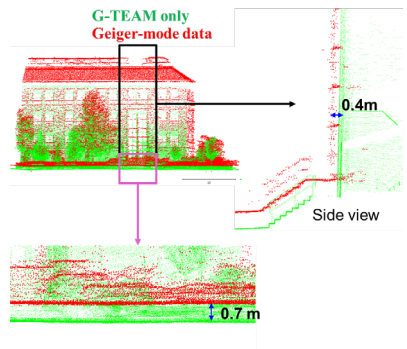


(a)

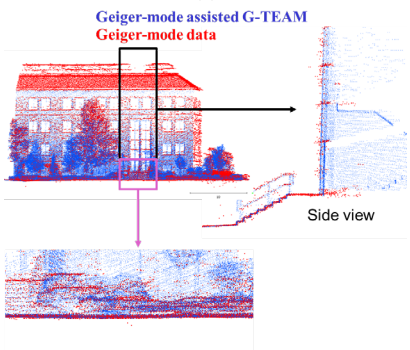


(b)

Figure 8. Sample point cloud (colored by height) for a single tree reconstructed using the (a) original GNSS/INS trajectory and (b) enhanced trajectory from G-TEAM.

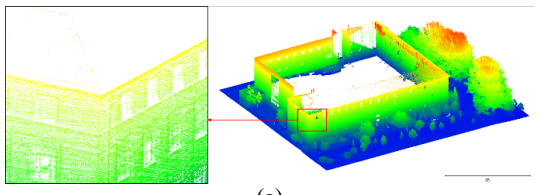


(a)

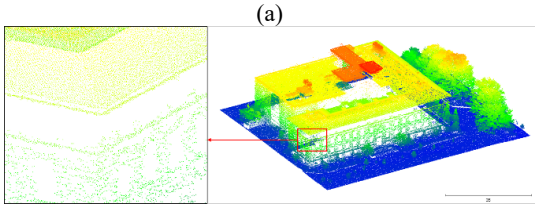


(b)

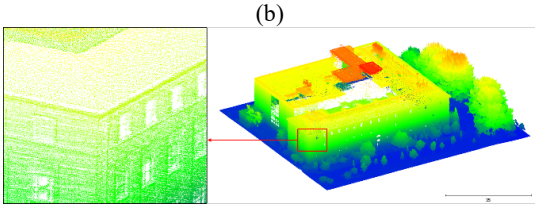
Figure 9. Sample illustration of Geiger-mode point cloud (in red) alignment with BackPack point clouds from (a) G-TEAM only (in green) and (b) Geiger-mode-assisted G-TEAM (in blue).



(a)



(b)



(c)

Figure 10. Sample point clouds from BikePack and Geiger-mode datasets: (a) BikePack after applying the Geiger-mode assisted G-TEAM, (b) Geiger-mode point cloud, and (c) combined Geiger-mode and BackPack data (zoom-in views of the areas within the red rectangle are illustrated on the left).

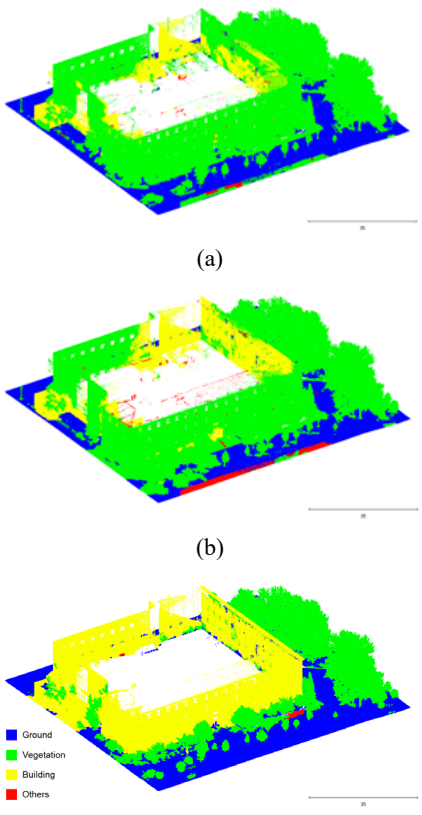
	Mean	STD	RMS
X_{dif} (m)	-0.192	0.300	0.356
Y_{dif} (m)	0.142	0.334	0.363
Z_{dif} (m)	0.261	0.059	0.267
ω_{dif} (°)	0.007	0.120	0.120
φ_{dif} (°)	0.011	0.094	0.095
κ_{dif} (°)	0.060	0.200	0.209

Table 1. Statistics of differences between initial and refined trajectory parameters for the BikePack dataset using the Geiger-mode assisted GTEAM (number of adjusted sets of pose parameters: 4,030).

To evaluate the performance of the proposed cross-labelling strategy for semantic segmentation of the BikePack dataset, the results are compared with another two segmentation approaches. The first approach uses a pretrained SPT model on the DALES airborne LiDAR dataset (Varney et al., 2020). The second approach utilizes a refined SPT model through transfer learning using the labelled Toronto-3D dataset (Tan et al., 2020), which was collected using a Mobile Mapping System (MMS). The Toronto data was used for the transfer learning as it has similar characteristics to that derived from the BackPack. Figure 11 presents the segmentation results from three different segmentation strategies. A significant number of building points are incorrectly segmented as vegetation for the pretrained and refined SPT models, as shown in Figure 11a and 11b, respectively. In contrast, Figure 11c shows the results after applying cross-labelling, demonstrating a significant improvement with most building points correctly segmented. Figure 12 shows samples of segmented trees.

In terms of quantitative analysis, a manually labelled BikePack dataset is used as a reference. Based on this reference, accuracy, IoU, and Cohen’s Kappa score are reported to assess segmentation performance, as shown in Table 2. The point cloud

is segmented into ground, vegetation, buildings, and others (i.e., streetscapes, vehicles, bushes, and utility poles). The difference in the number of points between the segmentation results is primarily due to the application of the KNN algorithm in the cross-labelling approach. Among the three semantic segmentation strategies, cross-labelling yields the best results. For vegetation segmentation, the proposed method achieves the highest user’s accuracy of 0.801. This accuracy is further supported by around 30% improvement in the IoU for vegetation. Compared to the refined SPT model, the proposed approach shows a more substantial improvement. Building segmentation from cross-labelling method shows the best results with the producer’s accuracy and IoU increasing by approximately 50% compared to those from the pretrained model. Similarly, ground segmentation benefits from the proposed approach, with producer’s accuracy and IoU increasing by approximately 10%. Even though the cross-labelling strategy is scoring lowest accuracy and IoU for others’ class, it still achieves the highest overall accuracy. Since the others’ class accounts for only 3% of the dataset, its lower segmentation accuracy has minimal impact on the overall accuracy and Cohen’s Kappa score. In contrast, ground points make up approximately 50% of the dataset, reinforcing the significance of accurate segmentation for major classes. Overall, the cross-labelling approach outperforms both the pretrained and refined SPT models, as reflected in its superior accuracy, mean IoU, and Cohen’s Kappa score. These performance metrics highlight the effectiveness of leveraging the well-aligned Geiger-mode and BackPack LiDAR data for label transfer, significantly improving segmentation quality across all classes.



(a)

(b)

(c)

Figure 11. Sample segmentation results of BikePack dataset from (a) pretrained and (b) refined SPT models together with the (c) proposed cross-labelling strategy.

Figure 12. Sample segmented trees from the BikePack dataset.

5. Conclusions and Recommendations for Future Work

This study introduced a framework to integrate airborne and BikePack LiDAR data for trajectory enhancement and mapping, together with a learning-based point cloud segmentation for urban tree mapping. The proposed G-TEAM applies surfel features within multi-layer mapping optimization to improve trajectory and mapping results in urban environments. It also offers an option to incorporate other geospatial data to improve georeferencing accuracy for BikePack data, enabling integration of multi-platform LiDAR point clouds. Based on experimental results, the relative and absolute accuracy of BikePack point cloud after applying the airborne-assisted G-TEAM are significantly improved, achieving seamless integration with Geiger-mode data. Additionally, the proposed cross-labelling method transfers segmentation results between well-aligned datasets. Experimental results demonstrate a significant improvement over both the pretrained and refined-SPT models. The proposed cross-labelling method effectively utilizes well-georeferenced data, enabling segmentation without the need for additional training datasets. This makes it a practical and efficient solution for addressing training data scarcity in traditional learning-based strategies while also providing useful segmentation information for urban tree mapping.

For the limitations of the proposed framework, surfels used for map optimization in G-TEAM may struggle to present fine structural details in complex urban environments. Additionally, the presence of dynamic objects can also impact mapping accuracy. Future work will explore the incorporation of a mesh-based model to enhance the representation of fine structures. More accurate representation of objects in urban environments within a trajectory enhancement and mapping strategy can result in better point clouds. Moreover, dynamic object segmentation could be employed to mitigate the impact of moving elements. Another limitation is the dependency of the proposed cross-labelling strategy on the segmented Geiger-mode dataset. Segmentation errors in the Geiger-mode data will be directly transferred to the BikePack dataset. Future work could focus on enhancing the cross-labelling method by implementing error-filtering mechanisms or correction strategies before transferring labels between datasets.

Methods		Pretrained Model			Refined SPT Model			Cross-Labelling		
Number of points		200,637,510			200,637,510			193,222,322		
Metrics		Producer's Accuracy	User's Accuracy	IoU	Producer's Accuracy	User's Accuracy	IoU	Producer's Accuracy	User's Accuracy	IoU
Labels	Ground	0.865	0.994	0.860	0.699	0.998	0.698	0.948	0.943	0.897
	Vegetation	0.997	0.486	0.485	0.992	0.493	0.491	0.966	0.801	0.780
	Building	0.234	0.980	0.232	0.578	0.944	0.512	0.732	0.912	0.684
	Others	0.134	0.716	0.128	0.511	0.557	0.364	0.022	0.081	0.018
Overall Accuracy		0.733			0.727			0.890		
Mean IoU		0.426			0.516			0.594		
Cohen's Kappa Score		0.588			0.596			0.823		

Table 2. Quantitative evaluation of segmentation results for the BikePack dataset.

Acknowledgment

This work is partially supported by the USDA NRCS (award #NR233A750004G044) and USDA NIFA (award #2023-68012-38992 and #2024-67021-42879). The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government, NRCS, or NIFA. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

References

American Forests, 2025. For Tree Equity and Climate Change, How Many Urban Trees Do We Need? https://https://www.americanforests.org/tools-research-reports-and-guides/research-reports/climate-change-urban-forests/ (1 February 2025).

Demantké, J., Mallet, C., David, N., Vallet, B., 2011. Dimensionality based scale selection in 3D lidar point clouds. In Laserscanning.

Du, S., Li, Y., Li, X., Wu, M., 2021. LiDAR odometry and mapping based on semantic information for outdoor environment. Remote Sensing, 13(15), 2864.

Landrieu, L., Simonovsky, M., 2018. Large-scale point cloud semantic segmentation with superpoint graphs. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4558-4567).

Li, L., Kong, X., Zhao, X., Li, W., Wen, F., Zhang, H., Liu, Y., 2021. SA-LOAM: Semantic-aided LiDAR SLAM with loop closure. 2021 IEEE International Conference on Robotics and Automation (ICRA).

Lin, Y. C., Manish, R., Bullock, D., Habib, A., 2021. Comparative analysis of different mobile LiDAR mapping systems for ditch line characterization. Remote Sensing, 13(13).

Liu, H., Ye, Q., Wang, H., Chen, L., Yang, J., 2019. A precise and robust segmentation-based lidar localization system for automated urban driving. Remote Sensing, 11(11), 1348.

Nagai, K., Fasoro, T., Spenko, M., Henderson, R., Pervan, B., 2020. Evaluating GNSS navigation availability in 3-D mapped urban environments. 2020 IEEE/ION Position, Location and Navigation Symposium (PLANS),

Nowak, D. J., Greenfield, E. J., 2018. US urban forest statistics, values, and projections. Journal of Forestry, 116(2), 164-177.

Ramezani, M., Khosoussi, K., Catt, G., Moghadam, P., Williams, J., Borges, P., Pauling, F., Kottege, N., 2022. Wildcat: Online Continuous-Time 3D Lidar-Inertial SLAM. arXiv preprint arXiv:2205.12595.

Robert, D., Raguet, H., Landrieu, L., 2023. Efficient 3D semantic segmentation with superpoint transformer. Proceedings of the IEEE/CVF International Conference on Computer Vision,

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing,

Rufus, N., Nair, U. K. R., Kumar, A. S. B., Madiraju, V., Krishna, K. M., 2020. SROM: Simple real-time odometry and mapping using LiDAR data for autonomous vehicles. 2020 IEEE Intelligent Vehicles Symposium (IV),

Saltori, C., Galasso, F., Fiameni, G., Sebe, N., Poiesi, F., Ricci, E., 2023. Compositional semantic mix for domain adaptation in point cloud segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Sarker, S., Sarker, P., Stone, G., Gorman, R., Tavakkoli, A., Bebis, G., Sattarvand, J., 2024. A comprehensive overview of deep learning techniques for 3D point cloud classification and semantic segmentation. Machine Vision and Applications, 35(4), 67.

Shan, T., Englot, B., Meyers, D., Wang, W., Ratti, C., Rus, D., 2020. Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping. *In 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS),* pp. 5135-5142.

Tan, W., Qin, N., Ma, L., Li, Y., Du, J., Cai, G., Yang, K., Li, J., 2020. Toronto-3D: A large-scale mobile LiDAR dataset for semantic segmentation of urban roadways. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (pp. 202-203).

Varney, N., Asari, V. K., Graehling, Q., 2020. DALES: A large-scale aerial LiDAR data set for semantic segmentation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops,

Wallace, L., Sun, Q. C., Hally, B., Hillman, S., Both, A., Hurley, J., Saldias, D. S. M., 2021. Linking urban tree inventories to

remote sensing data for individual tree mapping. Urban Forestry & Urban Greening, 61, 127106.

Williams, J., Tadesse, A., Sam, T., Sun, H., Montañez, G. D., 2020. Limits of transfer learning. In Machine Learning, Optimization, and Data Science: 6th International Conference, LOD 2020, Siena, Italy, July 19–23, 2020, Revised Selected Papers, Part II 6 (pp. 382-393).

Wu, X., Jiang, L., Wang, P. S., Liu, Z., Liu, X., Qiao, Y., Ouyang, W., He, T., Zhao, H., 2024. Point Transformer V3: Simpler Faster Stronger. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4840-4851).