HeLiPR: Heterogeneous LiDAR Dataset for inter-LiDAR Place Recognition under Spatiotemporal Variations

Minwoo Jung[1], Wooseong Yang[1], Dongjae Lee[1], Hyeonjae Gil[1], Giseop Kim[2], Ayoung Kim[1]

Abstract

Place recognition is crucial for robot localization and loop closure in simultaneous localization and mapping (SLAM). Light Detection and Ranging (LiDAR), known for its robust sensing capabilities and measurement consistency even in varying illumination conditions, has become pivotal in various fields, surpassing traditional imaging sensors in certain applications. Among various types of LiDAR, spinning LiDARs are widely used, while non-repetitive scanning patterns have recently been utilized in robotics applications. Some LiDARs provide additional measurements such as reflectivity, Near Infrared (NIR), and velocity from Frequency modulated continuous wave (FMCW) LiDARs. Despite these advances, there is a lack of comprehensive datasets reflecting the broad spectrum of LiDAR configurations for place recognition. To tackle this issue, our paper proposes the HeLiPR dataset, curated especially for place recognition with heterogeneous LiDARs, embodying spatiotemporal variations. To the best of our knowledge, the HeLiPR dataset is the first heterogeneous LiDAR dataset supporting inter-LiDAR place recognition with both non-repetitive and spinning LiDARs, accommodating different field of view (FOV)s and varying numbers of rays. The dataset covers diverse environments, from urban cityscapes to high-dynamic freeways, over a month, enhancing adaptability and robustness across scenarios. Notably, HeLiPR includes trajectories parallel to MulRan sequences, making it valuable for research in heterogeneous LiDAR place recognition and long-term studies. The dataset is accessible at https://sites.google.com/view/heliprdataset.

1 Introduction

Place recognition is an essential task in robotics, involving the ability to identify whether a place has been visited before or not. The significance of this task stems from its role as an initial step towards localization and its contribution to enabling loop closure in SLAM. Traditionally, it has been accomplished by searching a query image within a database using image sensors (Zhang et al. 2010; Arandjelovic et al. 2016; Lee and Kim 2021). However, recent advancements have facilitated the adoption of LiDAR for place recognition, attributable to its enhanced geometric sensing capabilities. LiDAR-based place recognition has been gaining attraction thanks to its capacity to measure the range precisely, and distinct from image sensors, LiDAR has the advantage of capturing geometric structures with illumination invariance. Conventionally, LiDAR descriptors (Kim et al. 2021; Xu et al. 2022b; Luo et al. 2021) are generated from the scan and subsequently used to ascertain the presence or absence of a place through comparison with a comprehensive set of descriptors. While place recognition can be replaced using global positioning system (GPS), it has limitations in environments where signals are weak. LiDAR overcomes these challenges with high-resolution spatial data, enabling accurate place recognition even in GPS-denied areas, underscoring its importance in complex navigation tasks.

With the advancement of place recognition, the hardware capabilities of LiDAR have also evolved significantly. For instance, specific LiDARs deploy non-repetitive scanning patterns to achieve dense mapping, thus deviating from traditional spinning LiDARs. Additionally, some LiDARs feature a more significant number of rays, surpassing the conventional 16 or 32-ray configurations, and incorporate additional channels such as reflectivity and NIR. More recently, the advent of FMCW LiDAR has made it possible to measure relative velocity along the radial direction utilizing the Doppler effect, commonly called velocity measurement. Considering these developments in LiDARs, place recognition with non-repetitive scanning pattern LiDARs (Yuan et al. 2023, 2024) has also been pursued. Furthermore, studies (Wang et al. 2020; Shan et al. 2021; Chen et al. 2020) that leverage the information offered by the additional channels in LiDAR have also emerged.

Nevertheless, despite these advancements, there currently exists a scarcity of datasets incorporating diverse combinations of LiDARs for place recognition. This shortfall highlights a gap in the availability of benchmark datasets for validating place recognition operating with heterogeneous LiDARs. Several datasets (Kim et al. 2020; Knights et al. 2023; Geiger et al. 2012) are conducive for tasks involving

[1]Dept. of Mechanical Engineering, SNU, Seoul, S. Korea
[2]NAVER LABS, Gyeonggi-do, S. Korea

Corresponding author:
Ayoung Kim, Dept. of Mechanical Engineering, SNU, Seoul, S. Korea
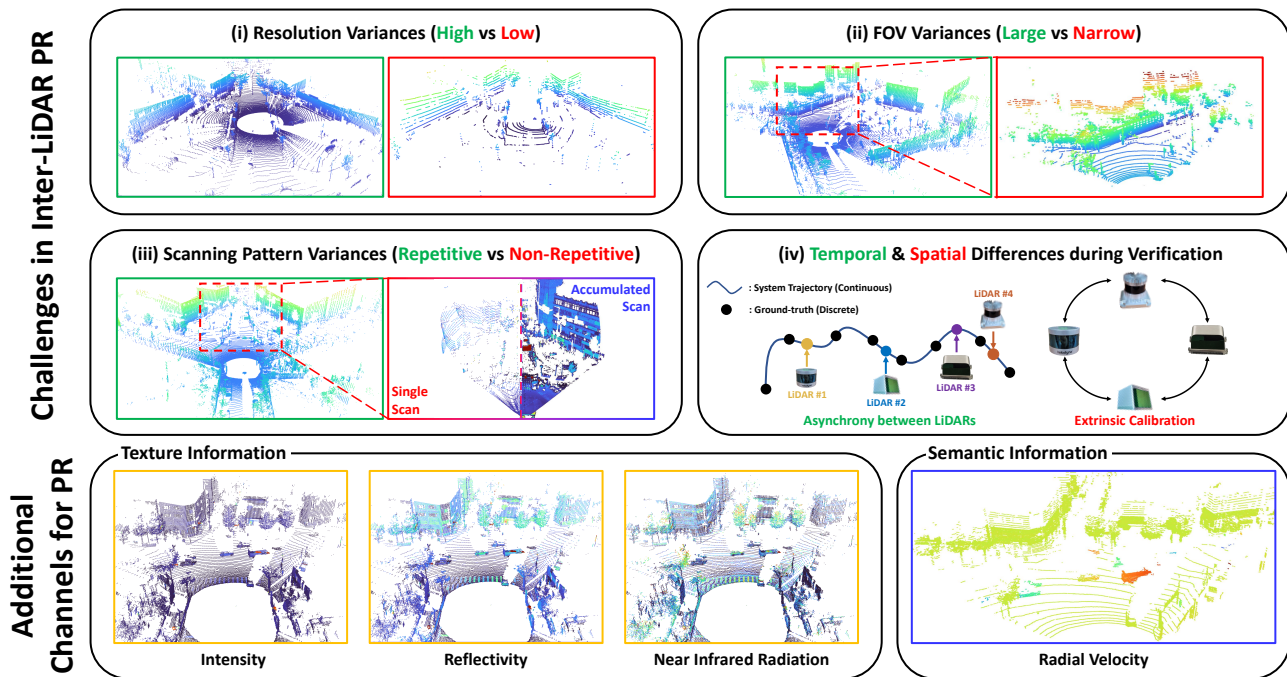Email: ayoungk@snu.ac.kr

Figure 1. (Top row) LiDAR place recognition challenges (i) Variance in resolution between high and low ray count LiDARs affects sensing abilities. (ii) While some LiDARs perform 360-degree scans, others have limited FOV due to occlusion or sensor limitations. (iii) Most LiDARs scan the repetitive area, whereas non-repetitive LiDAR densely scan by stacking individual scans. However, each scan tends to be sparse, as depicted in the left red box. (iv) Ground truth, crucial for executing LiDAR place recognition, is challenging to determine due to varying LiDAR coordinates and scan acquisition times. (Bottom row) HeLiPR dataset provides heterogeneous LiDARs and additional channels, thereby granting opportunities to utilize texture information from LiDAR.

place recognition, although they are equipped solely with a spinning LiDAR. On the other hand, while there are datasets inclusive of multiple LiDARs, these predominantly feature spinning LiDARs (Jeong et al. 2019; Barnes et al. 2020; Agarwal et al. 2020; Hsu et al. 2021), or they comprise heterogeneous LiDARs that are ill-suited for place recognition (Qingqing et al. 2022; Helmberger et al. 2021; Jung et al. 2023).

This paper introduces the HeLiPR dataset, a unique heterogeneous LiDAR dataset for place recognition, encapsulating spatiotemporal variations. Regarding environmental diversity, our dataset was acquired over 37 days, with data collection occurring 3 to 4 times. This acquisition provides a variety of environments encompassing a narrow residential area, urban cityscape, and environments with high dynamic change. One of the key aspects of our dataset is the focus on inter-LiDAR place recognition, which refers to the challenge of using multiple and diverse LiDARs both within individual sessions, known as intra-session, and across different sessions, referred to as inter-session. The complex challenges associated with this form of place recognition are comprehensively represented in Figure. 1. The HeLiPR dataset thoroughly encompasses each highlighted challenge, demonstrating its applicability in the field. In addition, the HeLiPR dataset includes trajectories similar to sequences acquired from MulRan (Kim et al. 2020), enabling another heterogeneous LiDAR and long-term place recognition with a term of four years. The salient contributions of the HeLiPR dataset are as follows:

1. The HeLiPR dataset includes heterogeneous LiDARs, with OS2-128, VLP-16, Livox Avia, and Aeries II, while most of the existing dataset involves only spinning LiDARs. This configuration can underscore

the impact of disparities in resolution and scanning patterns. Their additional channels, such as NIR, reflectivity, and radial velocity, pave the way for novel strategies in place recognition.

2. The HeLiPR dataset tackles heterogeneous LiDAR place recognition. Based on our benchmark results, the HeLiPR dataset underscores the growing need for dedicated research in heterogeneous inter-LiDAR place recognition. Furthermore, this dataset plays a significant role in facilitating and guiding essential research explorations in this field.

3. The HeLiPR dataset captures diverse environments monthly, from residential to dynamic urban areas. Moreover, trajectories akin to those in MulRan enable heterogeneous LiDAR place recognition and support long-term research spanning four years. This broad spectrum of data acquisition positions HeLiPR as a pivotal tool for generalizing place recognition across varied scenarios.

4. The HeLiPR dataset provides individual LiDAR ground truth corresponding to the acquisition time of each LiDAR. This accurate ground truth, which also considers spatial relationships, facilitates more accessible validation and improves the reliability of place recognition.

2   Related works

This section presents an overview of LiDAR datasets pertinent to our research. A summary is provided in Table. 1.

The KITTI dataset (Geiger et al. 2012), gathered using a carlike vehicle, represents a mid-sized cityscape. While

Table 1. Dataset comparison for LiDAR-based place recognition. The number of channels refers to the additional channels, excluding intensity. When the dataset contains non-identical paths that partially overlap with other sessions, it has inter-session difference. Furthermore, the ★ represents the spatial scale based on sequence length.

Name	LiDAR			Loop Closure				Spatial Scale	Total Distance
	# Spinning	# Solid State	# Channels	Intra-session	Inter-session	Temporal Diversity	Inter-session Differnce		
KITTI Geiger et al. (2012)	1	✗	✗	✓	✗	✗	✗	★★	44 km
Complex Urban Jeong et al. (2019)	2	✗	✗	✓	✓	15 months	✓	★★★	190 km
Oxford Rdar Barnes et al. (2020)	2	✗	✗	✓	✓	✗	✗	★★★	280 km
Ford Multi AV Agarwal et al. (2020)	4	✗	✗	✓	✓	4 months	✗	★★★	198 km
MulRan Kim et al. (2020)	1	✗	✗	✓	✓	2 months	✓	★★★	123 km
NTU VIRAL Nguyen et al. (2022)	2	✗	2	✗	✗	✗	✗	★	1.9 km
Boreas Burnett et al. (2023)	1	✗	✗	✓	✓	12 months	✗	★★★	350 km
Pohang Canal Chung et al. (2023)	3	✗	2	✗	✓	1 month	✗	★★	45 km
Wild Place Knights et al. (2023)	1	✗	✗	✓	✓	14 months	✓	★★	33 km
Hilti 2021 Helmberger et al. (2021)	1	1	2	✗	✗	✗	✗	★	2.1 km
Tiers Qingqing et al. (2022)	3	3	2	✗	✗	✗	✗	★	2 km
HeLiPR	**2**	**2**	**3**	**✓**	**✓**	**1 + 53 months**	**✓**	**★★★**	**164 km**

Table 2. The heterogeneous LiDARs utilized in Our Dataset

Sensor	Manufacture	Model	Channel	FOV (H× V)	Range
Spinning	Ouster	OS2-128	128	360° × 22.5°	200 m
Spinning	Velodyne	VLP-16	16	360° × 30°	100 m
Solid state	Livox	Avia	6	70° × 77°	450 m
Solid state	Aeva	Aeries II	64	120° × 19.2°	150 m

it facilitates intra-session place recognition, the dataset falls short in supporting inter-session place recognition, with data acquisition solely reliant on a single HDL-64E. On the other hand, the Oxford Robotcar Radar Dataset (Barnes et al. 2020), which shares a similar environment with KITTI, introduces the possibility for inter-session place recognition. However, even though multiple LiDARs are incorporated, all are of the spinning type. The Ford Multi-AV Dataset (Agarwal et al. 2020) stands out due to its extensive trajectory covering a range of environments from urban to vegetated, including tunnels, and showcasing seasonal changes. Similarly, Boreas (Burnett et al. 2023) meets the conditions necessary for intra and inter-session place recognition. However, each sequence from Boreas and the Ford Multi-AV dataset consists of similar paths, reducing the complexity in inter-session place recognition. The Complex Urban Dataset (Jeong et al. 2019) and UrbanNav Dataset (Hsu et al. 2021), both situated within urban environments, lean more towards intra-session place recognition, offering limited avenues for inter-session recognition. The Wild Places (Knights et al. 2023) stands apart by ensuring both intra-session and inter-session place recognition, factoring in temporal variations. Nevertheless, it focuses on unstructured terrains and employs a single spinning LiDAR. Unlike the previous dataset, the Pohang Canal dataset (Chung et al. 2023) utilizes multiple LiDARs. However, sessions for inter-session place recognition are not adequate as the trajectory of all sessions has an identical path. The NTU VIRAL dataset also exploits multiple LiDARs; however, its primary focus on unmanned aerial vehicle (UAV) localization, particularly in smaller areas for maintaining the tracking of the Leica laser system, tends to overshadow its application in place recognition.

Many of the datasets mentioned above primarily rely on spinning LiDARs. More recent datasets, such as Tiers (Qingqing et al. 2022), Hilti 2021 (Helmberger et al. 2021), and City (Jung et al. 2023) dataset, have begun to incorporate heterogeneous LiDARs. The Tiers and Hilti 2021 datasets feature short-term indoor and outdoor data collection using carlike vehicles and handheld systems. Similarly, the City dataset captures urban areas using a vehicle-based system. Although these datasets employ heterogeneous LiDARs, their primary focus is on SLAM. As a result, they tend to have relatively short paths, which means that revisits are either minimal or non-existent in sequences, rendering intra-session place recognition unfeasible. Additionally, the lack of overlap in their sequences means these datasets are unsuitable for validating inter-session place recognition.

The HeLiPR dataset distinguishes itself from others by showcasing diverse LiDARs, encompassing the OS2-128, VLP-16, Livox Avia, and Aeva Aeries II, each with unique attributes. These sensors capture data channels such as NIR, reflectivity, and radial velocity, ushering in new avenues for inventive place recognition. Significantly, the HeLiPR dataset captures each sequence over a month, supplying rich environments that support both intra-session and inter-session place recognition. Furthermore, the HeLiPR dataset trajectories resonate with sequences derived from MulRan, thus promoting research in heterogeneous LiDAR place recognition and offering an extended temporal perspective. Conclusively, every sequence in HeLiPR illustrates a vast environment with variations.

3 System Overview

3.1 System Configuration

Our system comprises four distinct LiDARs, as depicted in Table. 2 and Figure. 2. The spinning OS2-128 LiDAR is mounted at the center of the system, elevated to allow dense scanning without any occlusion. In contrast, the spinning LiDAR, VLP-16, experiences particular self-occlusion because of its proximity to the front box and surrounding
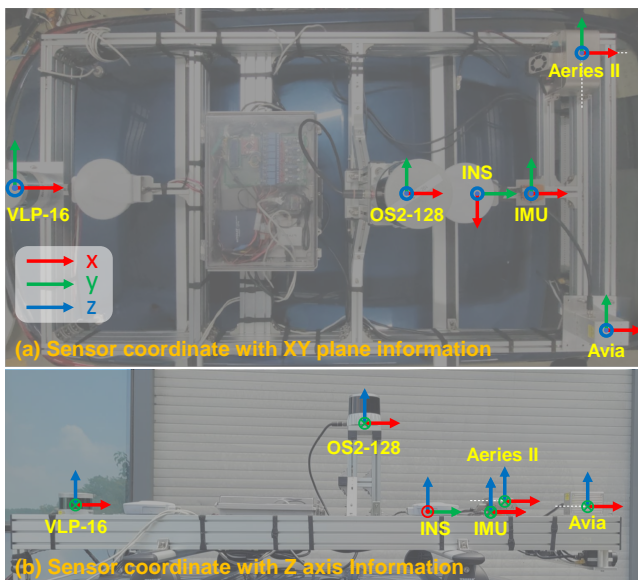
Figure 2. Sensors coordinate information between sensors. (a) and (b) represent the transformation with the xy-plane and z-axis. After the extrinsic calibration, the inter-sensor transformation can be found in the `Calibration` folder.

sensors. This self-occlusion leads to an inability to fully scan the front view, impacting the data collection and operational limitations. Also, due to inherent hardware constraints, it casts a significantly smaller ray than the OS2-128, enabling only a peripheral scan of its surroundings. The remaining LiDARs, Livox Avia and Aeva Aeries II are oriented to scan the front view of the vehicle, and each presents unique limitations. In the case of the Avia, its unconventional scanning patterns deviate from traditional spinning LiDARs; thus, direct comparison with them is challenging. However, Avia can construct dense maps with accumulating non-repetitive scans based on the relative transformation between scans. The Aeries II also presents a narrow horizontal FOV. Even if this LiDAR has the advantage of detecting radial velocities of points, FMCW technology introduces noise into the range measurements. Among several range configurations, we choose the configuration with maximum range of 150 m. This choice is made since a longer-range configuration leads to more noise, which could compromise accuracy. This combination of LiDARs, with their unique scanning patterns, allows for an intriguing exploration in place recognition, including dealing with occlusion scenarios and contrasting low versus high resolution as shown in Figure. 1. Furthermore, leveraging the characteristics of these LiDARs could significantly enhance place recognition in environments characterized by substantial dynamics or rich textures. The additional channels offered by each sensor can be found in Figure. 4, and we have identically configured all the LiDAR sensors to operate at a frequency of 10 Hz.

In addition to the LiDARs, our system incorporates two types of inertial sensors, the inertial measurement unit (IMU) and the inertial navigation system (INS). These devices provide a means to determine the temporal and spatial relationships within the asynchronous LiDAR system. We employ the Xsens MTi-300, which measures inertial information at 100 Hz. We use the SPAN-CPT7 coupled with a dual VEXXIS GNSS-501 antenna to establish a baseline for the vehicle system. All baselines are achieved at
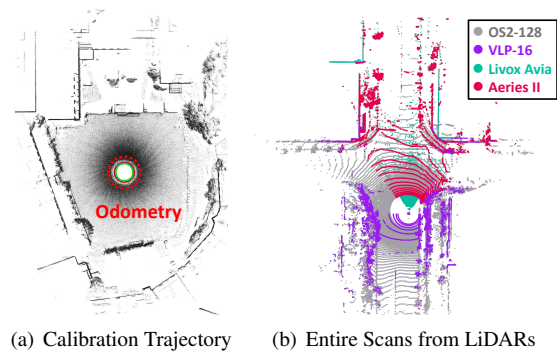

(a) Calibration Trajectory   (b) Entire Scans from LiDARs

Figure 3. (a) Extrinsic calibration trajectory: A circular path used for map construction and calibration purposes. (b) Post-calibration LiDAR alignment: A sky plot view illustrating the contours overlap between individual LiDAR scans.

a frequency of 50 Hz using RTK GPS and INS. Due to each sensor acquiring measurements in its own coordinate system, an extrinsic calibration process is necessary to integrate all the data into a standard coordinate system. This ensures consistency and accuracy across various measurements.

3.2   Sensor Calibration

For simplicity, we employ symbols to represent the coordinate systems: L corresponds to the LiDAR, N signifies the INS, I is used for the IMU, and W indicates the world system.

3.2.1   Multiple LiDAR Extrinsic Calibration To calculate the extrinsic calibration between LiDARs, we utilize the existing calibration method (Liu et al. 2022). In this method, the trajectories from each LiDAR are obtained and updated through batch optimization with scans from each LiDAR. After that, based on the updated trajectories and the initial extrinsic calibration parameters, batch optimization with multiple LiDARs is re-performed to calculate an accurate extrinsic calibration. To implement this method properly, a specified procedure is followed. The initial step involves moving the system a minimum distance to ensure trajectory accuracy. A complete 360-degree rotation follows this to facilitate the capture of loop closure for narrow FOV LiDARs. In the case of vehicles, given their inability to rotate in place, the system proceeded with a circular trajectory as shown in Figure. 3(a). Furthermore, considering the potential distortion of LiDAR that might occur during motion, we stop the movement for 10 seconds during a motion and acquire a total of 30 scans with stationary. Lastly, considering the sparsity of a single scan from the Livox Avia, LiDAR scans are accumulated in a stationary condition. The initial extrinsic calibration is established using the CAD model. Additionally, the odometry for each LiDAR is obtained using Direct LiDAR Odometry (Chen et al. 2022). As depicted in Figure. 3(b), it is clear that all the LiDARs are accurately aligned due to the precise extrinsic calibration.

3.2.2   IMU-LiDAR Extrinsic Calibration The extrinsic calibration between the IMU and the LiDAR commences with the CAD model serving as the initial estimation. The extrinsic calibration of the IMU-LiDAR ($\mathbf{T}_I^L$) is subsequently computed using LiDAR-Inertial Odometry (Xu et al. 2022a). It is updated while transforming the LiDAR scan to IMU coordinates and calculating the point-to-plane distance relative to the global map. However, achieving 6-DOF
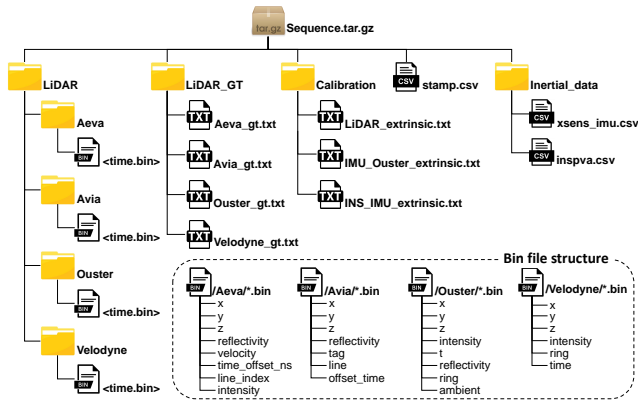
Figure 4. File structure of the HeLiPR dataset, illustrating the organization of LiDAR scans, ground truths, calibration, and inertial sensor measurements for each sequence

Figure 5. INS-based trajectories for sequences `01`, `02`, and `03`. The left shows trajectories on aerial images for `01`, while the right visualizes `02` (bottom) and `03` (top) with a color gradient. Notably, red indicates the start point, while blue designates the endpoint.

motion with a vehicle can prove challenging, which may adversely impact the accuracy of the extrinsic calibration between them. To mitigate this issue, the extrinsic calibration is updated with a minimal covariance, ensuring no significant deviation from the initial estimation. The entire process is executed based on the `Roundabout01` sequence, leading to the calculation of extrinsic calibration between the OS2-128 and the IMU. The reason for selecting these two sensors is that they are positioned colinearly, resulting in an almost zero distance between one axis. Additionally, as the two sensors share the same axis, the initial estimate of these parameters remains the most accurate among all the LiDARs.

3.2.3 INS-IMU Extrinsic Calibration The extrinsic calibration between the INS and IMU is conducted using MA-LIO (Jung et al. 2023). This method is particularly effective for asynchronous LiDARs. As the trajectory from MA-LIO is aligned to the IMU coordinate system, the subsequent hand-eye calibration between the INS and IMU can be executed. Specifically, the relative transformation, or $\Delta\mathbf{T}^N$, between $\mathbf{T}^N$ at two distinct timestamps t_i^N and t_j^N is determined. Similarly, the relative transformation, $\Delta\mathbf{T}^I$, between $\mathbf{T}^I$ at timestamps t_i^I and t_j^I is ascertained. Considering the non-coincident time, t^N, and t^I, we synchronize the acquisition time across both sensors to minimize time discrepancies. Then, the extrinsic calibration between the INS and IMU is achieved via the equation $\Delta\mathbf{T}^N \mathbf{T}_N^I = \mathbf{T}_N^I \Delta\mathbf{T}^I$. We employ 15000 transformations from the `Roundabout01` sequence to carry out the calibration, with the maximum time difference registering at approximately 1 ms$\cdot$ Moreover, it is worth noting that the fidelity of hand-eye calibration is inherently reliant on the precision of the MA-LIO. As such, we employ the CAD model specifically for the z-axis translation, which is most susceptible to errors in LIO. For all other components, we use the results from the hand-eye calibration.

4 Description of HeLiPR Dataset

4.1 Data Format

We offer sensor-specific data as individual files in diverse formats to optimize dataset management and facilitate access to each file and frame. Furthermore, we supply a file player based on Robot Operating System (ROS), tasked with
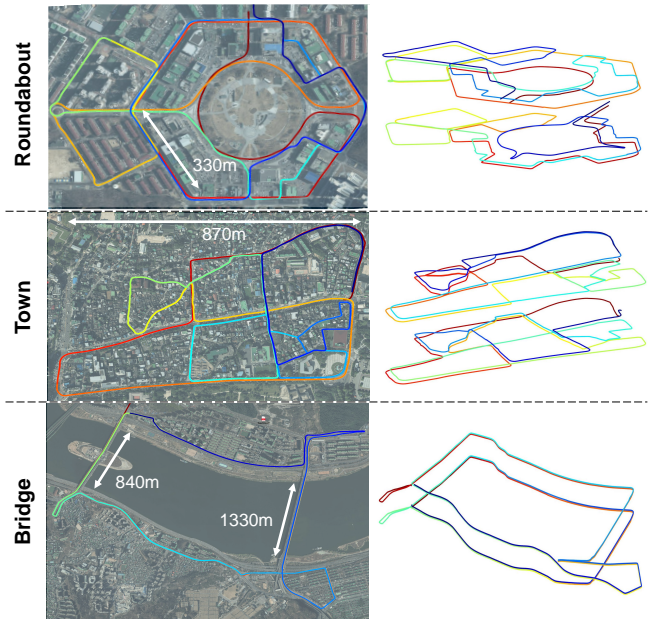
reading and publishing these files into ROS topics, ensuring seamless accessibility for pre-existing place recognition and tasks like SLAM. The file structure of the HeLiPR dataset is delineated in Figure. 4. The acquisition time of all measurements is stored in `stamp.csv`, and detailed descriptions of the data are presented subsequently.

4.1.1 Multiple LiDARs Data Individual LiDAR scans are stored as binary files in the `LiDAR/Sensor_name`. These files, identified as `<time.bin>`, encompass common channels such as (x, y, z), time offset, and ring (or line) index. We illustrate the array of unique channels for each LiDAR and the order of their storage in Figure. 4.

4.1.2 INS Data All INS data are stored in the `inspva.csv`. This file includes time, latitude, longitude, height, north_velocity, east_velocity, up_velocity, roll, pitch, azimuth, and data_status, organized in this order. Each value adheres to the East-North-Up (ENU) coordinate system, with the azimuth being determined by a left-handed rotation around the z-axis, in degrees, and clockwise from north.

4.1.3 IMU Data The complete set of IMU data is contained within the `xsens_imu.csv` file. Sequentially, this file encompasses time, quaternion (x, y, z, w), Euler angles (x, y, z), gyroscope (x, y, z), acceleration (x, y, z), and magnetic field (x, y, z).

4.1.4 Calibration and Ground truth Data The results of extrinsic calibration are saved in the `Calibration`. Additionally, we derive the individual LiDAR ground truth based on INS, LiDAR acquisition time and calibration parameters. Within the `LiDAR_GT`, the ground truth for each LiDAR is recorded, incorporating scan time, position (x, y, z), and quaternion (x, y, z, w). The procedure for generating this file is discussed in §4.3.

Table 3. The Description for Each Sequence.

Sequence Name	Characteristics	Sequence Index											
		01			02			03			04		
		Date	Duration	Distance	Date	Duration	Distance	Date	Duration	Distance	Date	Duration	Distance
Roundabout	Various rotation variations	2023-07-16	2730 s	9040 m	2023-08-01	2085 s	7447 m	2023-08-13	2515 s	9262 m	-	-	-
Town	FOV issue in narrow areas	2023-07-18	2414 s	7832 m	2023-07-31	2689 s	8203 m	2023-08-14	2528 s	8903 m	-	-	-
Bridge	Similar scenes and dynamic objects	2023-07-17	2144 s	23056 m	2023-07-31	2562 s	14615 m	2023-08-14	2009 s	19400 m	2023-08-21	3033 s	22958 m

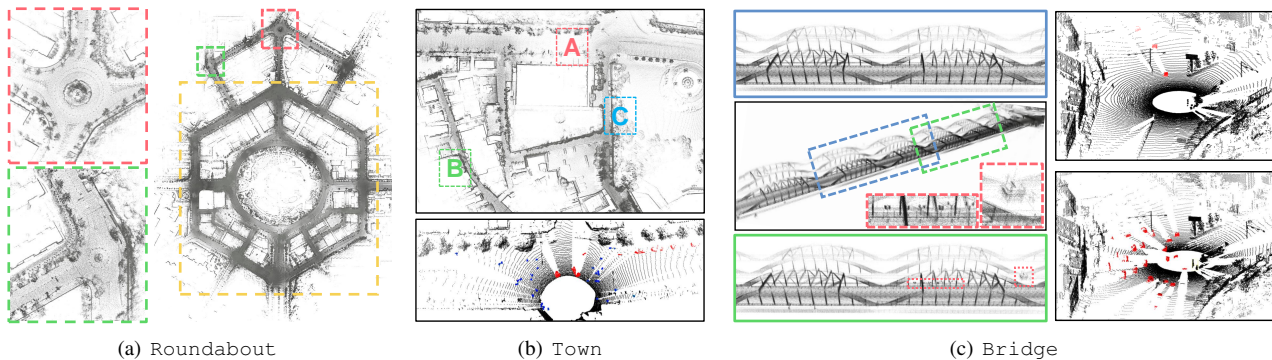(a) Roundabout   (b) Town   (c) Bridge

Figure 6. Characteristics of sequences. (a) Roundabout with three distinct roundabouts, each highlighted by a colored box. (b) Town showcasing a mix of narrow alleys and wider boulevards with width indications for areas A, B, and C, which are measured approximately 15 m, 3 m and 5 m in width. Dynamic entities, such as pedestrians (blue) and vehicles (Red), are marked. (c) Bridge emphasizing the challenge of scene similarities within the sequence. The portions of the bridge highlighted in blue and green boxes appear remarkably similar but display subtle differences, as indicated in the red box. Additionally, the presence of dynamic objects (red) and variations in numbers between Bridge01 (upper) and Bridge02 (lower) pose quite challenges for place recognition.

4.2 Sequence Explanation

In the HeLiPR dataset, we present three distinct places, namely Roundabout, Town, and Bridge. These places are meticulously acquired through three repetitions denoted as 01-03, with a two-week interval between each acquisition. The deliberate interval introduces temporal changes to enable inter-session place recognition. Furthermore, this temporal variation encompasses both night and day environments, leading to notable variations in the presence of dynamic objects throughout the sequences. It also allows for spatial variations, such as lane changes or reversing directions, when capturing data at the same location but on different paths. Detailed information, including acquisition time, duration, and distance, can be found in the Table. 3. Each sequence showcases unique environmental characteristics and introduces novel challenges in inter-LiDAR place recognition. We focus on enhancing both intra-session and inter-session loop closure candidates, with the primary objective of generating an abundant set of queries for place recognition. All sequences' trajectory and characteristics are represented in Figure. 5 and Figure. 6.

(*i*) **Roundabout01-03:** Roundabout stands out as the most formal environment for place recognition among all the sequences. Tall buildings and wide roads enrich the dataset with abundant features that aid in place recognition. As its name suggests, it consists of three roundabouts: one large and two of a comparatively smaller size, as shown in Figure. 6(a). The presence of a large roundabout and an outer hexagon design ensures easy revisiting of previously encountered locations. Moreover, the interconnected layout of roads and alleys within both the roundabout and hexagon facilitate seamless movement, enabling access to various exits from any entrance. These spatial features provide diverse candidates for place recognition and rotational variations not typically encountered in regular road scenarios.

(*ii*) **Town01-03:** Town presents a wide road environment in the center of the route, facilitating efficient scanning of buildings and structures. This characteristic shares similarities with the Roundabout sequence. However, in Town, the buildings are relatively short compared to those in Roundabout, and narrow alleys are more frequent. These alleys pose challenges in utilizing the wide sensing capabilities of LiDAR, creating a situation akin to indoor place recognition. These challenges can be found in Figure. 6(b). Furthermore, the presence of diverse dynamic objects in narrow alleys contributes to a lower proportion of static objects in the scene. Spinning LiDAR systems effectively address these limitations with their expansive FOV. In contrast, solid-state LiDAR systems inherently have a narrower FOV, potentially leading to a significantly reduced detection of static objects. These environment-related disparities add another layer of complexity to place recognition, demanding sophisticated approaches to handle such spatial variations effectively.

(*iii*) **Bridge01-04:** Bridge consists of a total of 2 laps, covering two bridges with lengths of 1.3 km and 0.8 km , respectively. It should be noted that Bridge01 and Bridge04 differ from Bridge02 and Bridge03, with the former being driven using a reverse trajectory. This place introduces a significant challenge in place recognition due to the consecutive appearance of similar scenes in most areas. The appearances with small differences represented in Figure. 6(c) can lead to result in many false positives. Moreover, numerous dynamic objects add complexity to place recognition, particularly depending on the sequence index. Notably, Bridge02 and Bridge04 display a relatively slow speed distribution attributed to the high density of dynamic objects. The Aeries II, which is a FMCW LiDAR, allows for measuring the radial velocity of a point, enabling the detection of certain dynamic objects. This capability opens up the potential for developing a novel place
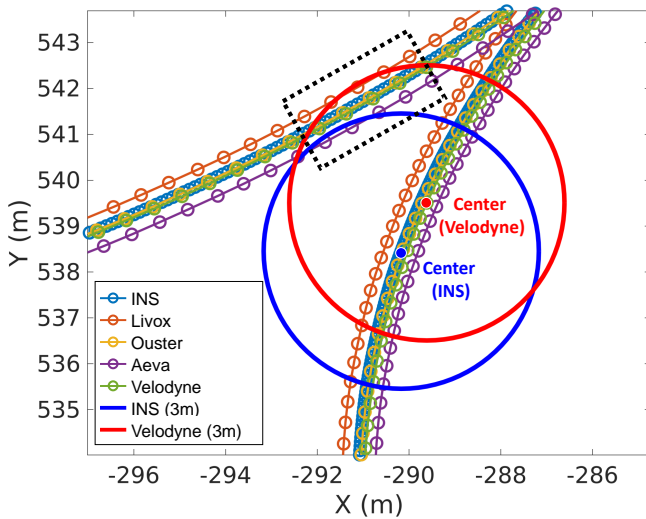
Figure 7. Partial individual ground truth for `Roundabout01`. Using individual ground truth with Velodyne as a reference, true positive place recognition candidates are observed for all LiDARs within a 3 m radius of other paths (black box). However, exact matches with the INS location (within a 5ms discrepancy) are only found in Aeva.

recognition based on this specific sequence and leveraging the unique features of the Aeries II.

4.3 Estimating Individual LiDAR Ground Truth with B-spline Interpolation

In place recognition, the positions of the query and the candidate must be determined, necessitating the precise positions of the LiDAR at each timestamp. Solutions such as GPS, INS, or SLAM can be employed to determine the trajectory of the system, with the resulting trajectory treated as the ground truth. However, considering that this ground truth represents the location of a reference point at specific timestamps, it is essential to account for spatiotemporal variances when expressing precise positions for all LiDARs. Extrinsic calibration, defining the spatial relationship, introduces differences between the LiDAR and the ground truth. Although all LiDARs capture scans under Coordinated Universal Time (UTC) via Precision Time Protocol (PTP), their individual acquisition times vary, leading to temporal discrepancies. As a result, timestamps from each LiDAR do not align with the ground truth time. The ground truths for each LiDAR and the single source are depicted in Figure 7, emphasizing the necessity of obtaining ground truths for all LiDARs by considering spatiotemporal variances rather than relying on a single ground truth. The figure also demonstrates that assuming Velodyne is at the same position as the ground truth from a single source results in only the Aeva set being identified as true positive. However, accounting for discrepancies allows different LiDAR sets to be recognized as true positives. This underscores the importance of considering these variances for more comprehensive and accurate place recognition. Therefore, incorporating such discrepancies is beneficial when leveraging multiple LiDARs, providing a more practical approach than relying on a single ground truth.

For the HeLiPR dataset, INS serves as ground truth. While INS operates at a frequency of 50 Hz, which is

five times faster than the LiDAR, timestamps of INS are not synchronous with timestamp of LiDARs. To mitigate this issue, we calculate the location of LiDARs at specific timestamp using B-spline interpolation (Mueggler et al. 2018) based on INS position, as relying solely on linear approximation may lead to imprecision. We choose B-spline interpolation for its compatibility with the inherently smooth characteristics of real vehicle trajectories. Given that the kth scan of a specific LiDAR, denoted as L_k, is acquired at time t_k, this position, $\mathbf{T}_W^N = (\mathbf{R}_W^N, \mathbf{t}_W^N)$ can be determined by leveraging four nearby INS measurements as control points. Nevertheless, our primary interest lies in determining the location of $\mathbf{T}_W^L = (\mathbf{R}_W^L, \mathbf{t}_W^L)$ in the world coordinates of L. It can be calculated as the multiplication with $\mathbf{T}_W^N$, IMU-LiDAR ($\mathbf{T}_I^L$) and INS-IMU ($\mathbf{T}_N^I$) extrinsic calibration. To achieve user convenience, we standardize $\mathbf{T}_W^L$ using the Universal Transverse Mercato (UTM) coordinate system, opting not to employ latitude or longitude. This decision simplifies plotting processes, and these coordinates streamline the direct comparison of trajectories with each other and the MulRan dataset.

In summary, our approach goes beyond simply utilizing multiple LiDARs for place recognition. By embracing the individual trajectories and specific ground truths of each LiDAR, we enable a more comprehensive evaluation of spatiotemporal variations. Our diverse ground truths significantly improve the accuracy and reliability of place recognition outcomes over relying solely on a single ground truth.

4.4 HeLiPR Dataset: Long-Term Place Recognition in Tandem with MulRan

The HeLiPR dataset not only encompasses the sequences of `Roundabout`, `Town`, and `Bridge`, but also integrates sequences from the MulRan dataset, including `KAIST`, `DCC`, and `Rivierside`, catering to long-term place recognition. Contrary to the MulRan dataset captured with OS1-64, the HeLiPR dataset employs the LiDARs as mentioned earlier. This introduces the potential for heterogeneous LiDAR place recognition. Furthermore, a temporal variance spanning approximately four years offers a novel challenge: long-term place recognition.

The Long-term place recognition challenge within the HeLiPR dataset includes sequences `04-06`. The `04` sequence, captured at midnight, exhibits an almost complete absence of dynamic objects. IMU measurements are not included in this sequence but are not crucial for place recognition. In contrast, the `05` sequence, captured during the daytime, features diverse dynamic objects resembling the existing MulRan sequences. Unlike `04` and `05`, `06` was captured four months later. This period, shorter than the four-year gap with MulRan, reduces its complexity as a long-term place recognition challenge. Nonetheless, the temporal differences between `04`, `05`, and `06` offer a distinct perspective, capturing scene changes over a more moderate term. A comprehensive overview of each sequence can be found in Table. 4 and Figure. 8. Thanks to GPS and INS, the trajectories from MulRan and HeLiPR align seamlessly despite being captured at distinct times, facilitating place recognition. As observed in Figure. 8, certain areas show

Table 4. The Description for Additional Sequences.

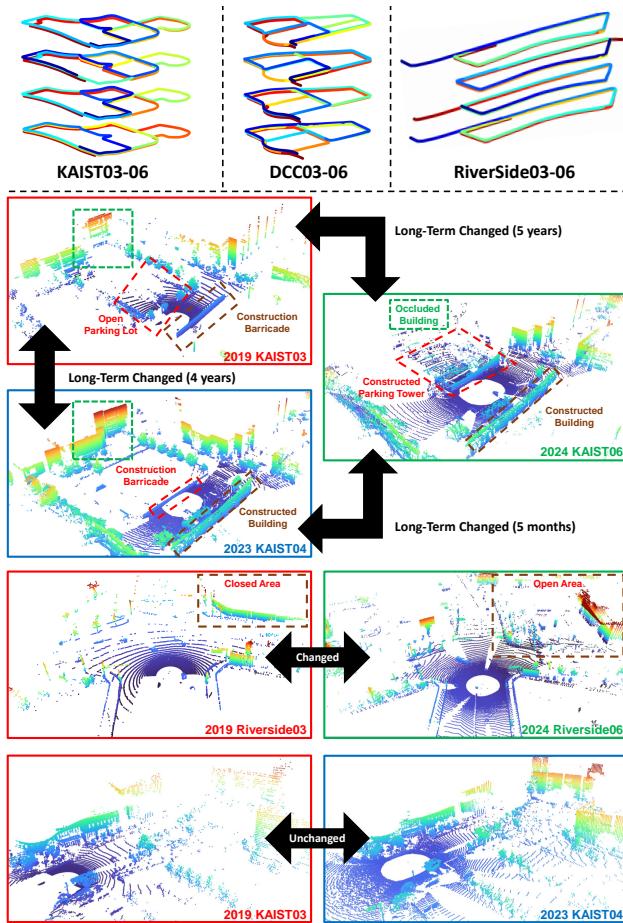| Sequence Name | Sequence Index | | | | | | | | | | | |
| | 04 | | | | 05 | | | | 06 | | | |
	Date	Time	Duration	Distance	Date	Time	Duration	Distance	Date	Time	Duration	Distance
KAIST	2023-08-31	Midnight	1261 s	6348 m	2023-08-31	Daytime	1248 s	6878 m	2024-01-16	Night	1215 s	6661 m
DCC	2023-08-31	Midnight	786 s	5506 m	2023-08-31	Daytime	1081 s	5309 m	2024-01-16	Night	1074 s	4648 m
Riverside	2023-08-31	Midnight	612 s	6523 m	2023-08-31	Daytime	855 s	6394 m	2024-01-16	Night	1195 s	7219 m

Figure 8. (Top row) Trajectories from GPS and INS data with three paths for each location, as Figure. 5. (Bottom row) shows environmental changes in the order `03`, `04`, and `06`. The `KAIST` depicts gradual scene changes at a similar location. `Riverside` highlights the distinct scan differences caused by different LiDARs and shows drastic changes between open and closed areas. The last row shows similar scans even after a long period.

partial modifications. From `KAIST03` to `KAIST06`, for instance, the construction of parking lots and buildings leads to significant scene changes. These new structures alter the landscape and occlude the LiDAR, causing scenes to appear different even in the same location. Such changes are crucial for both place recognition and change detection, as well as for map maintenance. Furthermore, in `Riverside03`, the scanning coverage is more limited compared to the current platform, with obstructions like a construction barricade blocking the upper right view from the vehicle. This limitation emphasizes the need to carefully compare similar areas and reduce reliance on regions with substantial scene changes for long-term place recognition in `Riverside06`. The challenge lies in accurately identifying revisits to specific locations, particularly in scenarios where parts of the environment have changed. In such cases, it becomes crucial not only to recognize these changes but also to possess the

capability to compare current scenes with previous ones by focusing on elements that remain unchanged. This situation also prompts further exploration into the extent of scene changes that can be accommodated in place recognition, questioning whether significantly altered locations can still be recognized as the same for revisiting purposes.

5 Benchmark Results with HeLiPR Dataset

This section presents an exhaustive analysis of state-of-the-art place recognition using the HeLiPR dataset. The comprehensive evaluation aims to identify the capabilities of state-of-the-art place recognition and emphasize the inherent need and importance of having datasets like HeLiPR to advance the field further. To assess the performance of the methodologies, which include Scan Context (SC) (Kim et al. 2021), RING++ (Xu et al. 2022b), BTC (Yuan et al. 2024) and LoGG3D-Net (Vidanapathirana et al. 2022), we employ three evaluation metrics: the Precision-Recall curve (PR-curve), the Area Under the Curve (AUC) score and R@1%.

Precision-Recall Curve (PR-Curve): This curve is an illustrative representation of a precision versus its recall. It provides a comprehensive visualization of the performance across different threshold levels. Mathematically, it can be expressed by the following equations:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1)$$

where TP denotes true positives, FP represents false positives, and FN signifies false negatives.

Area Under the Curve (AUC): This metric evaluates the overall performance of a place recognition. The AUC score provides a singular scalar value summarizing the entire PR curve. A perfect recognition method would achieve an AUC of 1.0, indicating flawless recognition, while a score closer to 0.5 might indicate a method performing no better than random guessing.
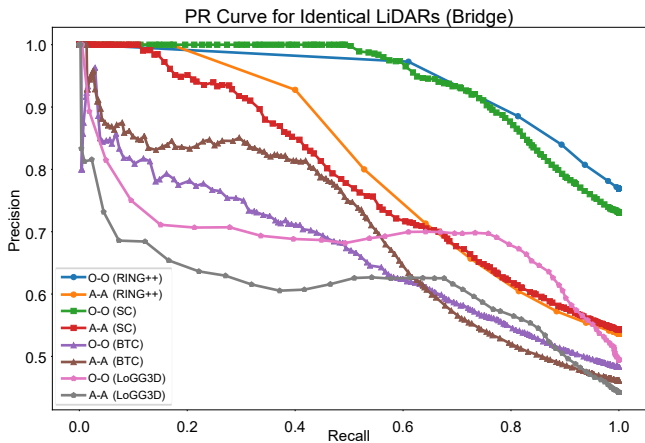
R@N(%): R@N(%) evaluates the recall within the top N or N percentage of results, which is calculated by determining if the true positive pairs are within the N or N percentage of the closest matches in the database. This metric provides insight into how effectively the algorithm identifies the most accurate matches from a larger pool of candidates, offering a more detailed perspective on its precision in high-relevance scenarios. For the benchmark results, we select $N = 1$ and represent both R@N and R@N%.

For the evaluation, the methodology entails sampling query scans at 10 m intervals and target scans at 5 m intervals. Successful place recognition is defined by identifying a candidate within a 7.5 m, termed a true positive. All scans have been undistorted and are configured with a maximum range of 100 m for descriptor extraction. For methods other than BTC, only the scans from Livox Avia are grouped into sets of 20 due to their sparse point
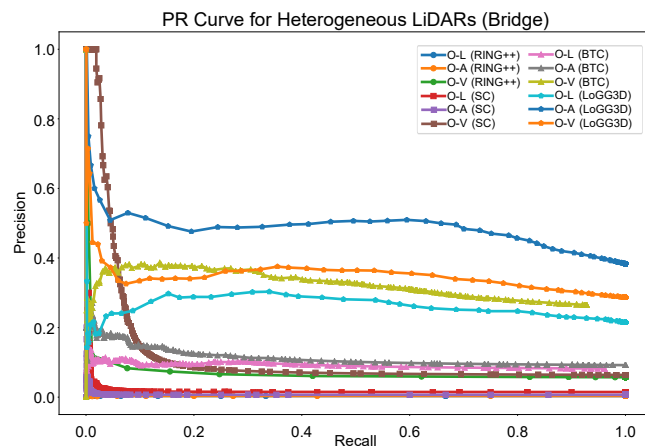
Table 5. AUC score and R@N for inter-LiDAR and inter-session place recognition

| Sequences | Method | Identical LiDARs | | | | | | Heterogeneous LiDARs | | | | | | | | |
| | | O-O | | | A-A | | | O-V | | | O-A | | | O-L | | |
		AUC	R@1	R@1%	AUC	R@1	R@1%	AUC	R@1	R@1%	AUC	R@1	R@1%	AUC	R@1	R@1%
`Roundabout01-03`	SC	0.942	0.880	0.947	0.786	0.648	0.681	0.119	0.077	0.130	0.007	0.008	0.018	0.017	0.018	0.030
	BTC	0.669	0.604	0.946	0.702	0.575	0.672	0.302	0.323	0.789	0.113	0.114	0.359	0.088	0.100	0.413
	RING++	0.950	0.926	0.993	0.809	0.640	0.704	0.067	0.069	0.251	0.003	0.004	0.052	0.003	0.004	0.047
	LoGG3D	0.766	0.552	0.639	0.746	0.538	0.629	0.531	0.442	0.612	0.641	0.474	0.605	0.392	0.333	0.522
`Town01-03`	SC	0.957	0.826	0.918	0.811	0.601	0.719	0.312	0.174	0.219	0.024	0.020	0.057	0.095	0.068	0.133
	BTC	0.625	0.463	0.650	0.685	0.482	0.569	0.343	0.291	0.467	0.149	0.136	0.377	0.153	0.149	0.403
	RING++	0.965	0.935	0.991	0.919	0.718	0.778	0.098	0.083	0.252	0.062	0.009	0.071	0.004	0.005	0.054
	LoGG3D	0.829	0.607	0.705	0.779	0.553	0.687	0.448	0.354	0.575	0.611	0.452	0.634	0.267	0.258	0.424
`Bridge02-03`	SC	0.713	0.786	0.948	0.666	0.712	0.925	0.108	0.091	0.183	0.019	0.018	0.035	0.012	0.013	0.041
	BTC	0.447	0.422	0.676	0.396	0.423	0.707	0.187	0.226	0.463	0.060	0.079	0.305	0.071	0.092	0.370
	RING++	0.868	0.855	0.995	0.802	0.800	0.993	0.037	0.049	0.238	0.005	0.018	0.035	0.005	0.013	0.041
	LoGG3D	0.692	0.670	0.903	0.612	0.599	0.890	0.347	0.389	0.778	0.486	0.518	0.851	0.263	0.303	0.597

- Symbol denotes LiDARs. (O: Ouster, A: Aeva, L: Livox, V: Velodyne)

(a) PR-Curve for Identical LiDAR (`Bridge`)

(b) PR-Curve for Heterogeneous LiDAR (`Bridge`)

Figure 9. PR-Curve of inter-session place recognition between `Bridge02` and `Bridge03`. (a) depicts the results when identical LiDAR systems are compared, illustrating the performance when different types of LiDARs each match with an identical LiDAR. (b) showcases the outcomes when heterogeneous LiDARs are used.

distribution. However, when evaluating BTC, similar to the original research, we accumulate 20 scans for every type of LiDAR. This approach differentiates it from other methods that typically use a single scan. LoGG3D-Net is selected for the deep learning-based place recognition method since LoGG3D-Net provides benchmark results of a mean maximum F1 score on the KITTI and MulRan datasets. The parameters and strategies for training can be found in our project page.

We evaluate each method using three inter-session pairs: `Roundabout01-03`, `Town01-03`, and `Bridge02-03`. Using Ouster as the reference database, we employ various LiDARs as queries to identify the corresponding Ouster candidates. To assess the influence of FOV, we also perform evaluations using the same LiDAR type. As detailed in Table. 5, we evaluate the four methods across three environments using five LiDAR pairings. This experimental configuration underscores the expansive scenarios of place recognition that the HeLiPR dataset can facilitate.

From Figure. 9 and Table. 5, we can discern several insights about the performance of place recognition. In terms of the dataset perspective, most methods tend to exhibit superior performance in `Roundabout` and `Town` compared to `Bridge`, likely due to the distribution of structures and inherent challenges highlighted in Figure. 6. Furthermore, for the model-based method, spinning LiDAR generally achieves better results with a higher AUC score and R@N than solid-state LiDAR. This is anticipated, given that spinning LiDARs have a large FOV. This expansive FOV equips them to adeptly handle place recognition from varied directions, including in scenarios like reverse visiting or navigating intersections.

In inter-LiDAR place recognition, we observed specific challenges associated with different LiDAR pairings: Ouster and Velodyne show resolution differences, Ouster and Aeva differ in horizontal FOV, and Ouster and Livox vary in both horizontal FOV and scanning patterns. These differences significantly affect the performance of model-based methods like RING++ and Scan Context, which are sensitive to LiDAR FOV variations as these descriptors are constructed with different FOVs in inter-LiDAR place recognition, a notable performance dip is observed in inter-LiDAR place recognition.

Contrary to other methods, BTC stands out in inter-LiDAR place recognition, particularly excelling in R@1% performance compared to other model-based approaches; however, its AUC and R@1 scores fall short of expectations. This discrepancy arises from the nearest target not having the largest similarity, meaning that the similarities between the query and target are not always aligned with the distance. Furthermore, there are instances where the overlap is significant, but the distance exceeds 7.5m, resulting in a false negative. Consequently, while the AUC and R@1 might seem underwhelming since they only utilize the best candidates, the R@1% shows a improvement

due to the enhanced possibility of identifying compatible matches among the total candidates. Unlike traditional place recognition for spinning LiDAR, place recognition of heterogeneous LiDAR may have a smaller overlap depending on the FOV or maximum range, even if the distance between query and target is close. Therefore, it is also proposed to distinguish based on overlap rather than distance as a criterion for determining true positives.

For the learning-based method, LoGG3D-Net, a slight underperformance is noted in identical LiDAR place recognition compared to model-based methods. However, it performs better inter-LiDAR place recognition, benefitting from training with super sequences that enhance its ability to distinguish between heterogeneous LiDARs. This is evident in its AUC score and R@N, with robust results in Ouster-Aeva comparisons. The similarity in the number of vertical channels between Ouster and Aeva likely contributes to better local feature aggregation. In contrast, Ouster and Livox exhibit lower scores, primarily because of the significant differences between their LiDAR characteristics, especially in comparison to Ouster and Velodyne. This emphasizes the sensitivity of LiDAR performance to resolution and FOV, with dual degradation occurring when FOV and scanning patterns differ. While the learning-based method shows promise for inter-LiDAR place recognition, effectively handling heterogeneous LiDARs remains challenging.

In summary, model-based methods excel when using identical LiDARs but fall short in inter-LiDAR scenarios. On the other hand, learning-based methods maintain consistent performance across various LiDAR combinations but only achieve partially satisfactory results. Recent learning-based methods try to perform place recognition with FOV variations (Kong et al. 2020; Vidanapathirana et al. 2021), it do not present reliable performances since FOV variances are relatively smaller than the difference between solid state and spinning LiDAR. All approaches currently need to be revised to achieve the desired performance levels. This analysis highlights the need for focused research in heterogeneous inter-LiDAR place recognition, with the HeLiPR dataset serving as a valuable resource for such investigations.

6 Conclusion

The HeLiPR dataset stands as a comprehensive resource that has been meticulously curated to showcase the remaining challenges of place recognition. It encompasses a broad spectrum of data from varied environments, including `Roundabout`, `Town`, and `Bridge`. One of the unique attributes of this dataset is its collection method; by introducing intentional time intervals and capturing data along diverse paths, we are ensuring the data reflects real-world spatiotemporal challenges. This not only mimics the dynamic nature of real-world scenarios but also enhances the application in localization and place recognition tasks. Additionally, the HeLiPR dataset overlaps with the MulRan dataset for long-term place recognition, presenting novel challenges in place recognition. With these features, the HeLiPR dataset is poised to become a valuable resource for improving place recognition and robotics applications, promoting advancements in the field.

References

Agarwal S, Vora A, Pandey G, Williams W, Kourous H and McBride J (2020) Ford multi-av seasonal dataset. *International Journal of Robotics Research* 39(12): 1367–1376.

Arandjelovic R, Gronat P, Torii A, Pajdla T and Sivic J (2016) Netvlad: Cnn architecture for weakly supervised place recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 5297–5307.

Barnes D, Gadd M, Murcutt P, Newman P and Posner I (2020) The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. In: *Proceedings of the IEEE International Conference on Robotics and Automation*.

Burnett K, Yoon DJ, Wu Y, Li AZ, Zhang H, Lu S, Qian J, Tseng WK, Lambert A, Leung KY, Schoellig AP and Barfoot TD (2023) Boreas: A multi-season autonomous driving dataset. *International Journal of Robotics Research* 42(1-2): 33–42.

Chen K, Lopez BT, Agha-mohammadi Aa and Mehta A (2022) Direct lidar odometry: Fast localization with dense point clouds. *IEEE Robotics and Automation Letters* 7(2): 2000–2007.

Chen X, Läbe T, Milioto A, Röhling T, Vysotska O, Haag A, Behley J and Stachniss C (2020) OverlapNet: Loop Closing for LiDAR-based SLAM. In: *Proceedings of the Robotics: Science & Systems Conference*.

Chung D, Kim J, Lee C and Kim J (2023) Pohang canal dataset: A multimodal maritime dataset for autonomous navigation in restricted waters. *International Journal of Robotics Research* 0(0): 02783649231191145.

Geiger A, Lenz P and Urtasun R (2012) Are we ready for autonomous driving? the kitti vision benchmark suite. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Helmberger M, Morin K, Berner B, Kumar N, Wang D, Yue Y, Cioffi G and Scaramuzza D (2021) The hilti slam challenge dataset.

Hsu L, Kubo N, Wen W, Chen W, Liu Z, Suzuki T and Meguro J (2021) Urbannav: An open-sourced multisensory dataset for benchmarking positioning algorithms designed for urban areas. In: *ION GNSS+*. pp. 226–256.

Jeong J, Cho Y, Shin YS, Roh H and Kim A (2019) Complex urban dataset with multi-level sensors from highly diverse urban environments. *International Journal of Robotics Research* 38(6): 642–657.

Jung M, Jung S and Kim A (2023) Asynchronous multiple lidar-inertial odometry using point-wise inter-lidar uncertainty propagation. *IEEE Robotics and Automation Letters* .

Kim G, Choi S and Kim A (2021) Scan context++: Structural place recognition robust to rotation and lateral variations in urban environments. *IEEE Transactions on Robotics* 38(3): 1856–1874.

Kim G, Park YS, Cho Y, Jeong J and Kim A (2020) Mulran: Multimodal range dataset for urban place recognition. In:

Proceedings of the IEEE International Conference on Robotics and Automation. pp. 6246–6253.

Knights J, Vidanapathirana K, Ramezani M, Sridharan S, Fookes C and Moghadam P (2023) Wild-places: A large-scale dataset for lidar place recognition in unstructured natural environments. In: *Proceedings of the IEEE International Conference on Robotics and Automation*.

Kong X, Yang X, Zhai G, Zhao X, Zeng X, Wang M, Liu Y, Li W and Wen F (2020) Semantic graph based place recognition for 3d point clouds. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 8216–8223.

Lee AJ and Kim A (2021) Eventvlad: Visual place recognition with reconstructed edges from event cameras. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. pp. 2247–2252.

Liu X, Yuan C and Zhang F (2022) Targetless extrinsic calibration of multiple small fov lidars and cameras using adaptive voxelization. *IEEE Transactions on Instrumentation and Measurement* 71: 1–12.

Luo L, Cao SY, Han B, Shen HL and Li J (2021) Bvmatch: Lidar-based place recognition using bird's-eye view images. *IEEE Robotics and Automation Letters* 6(3): 6076–6083.

Mueggler E, Gallego G, Rebecq H and Scaramuzza D (2018) Continuous-time visual-inertial odometry for event cameras. *IEEE Transactions on Robotics* 34(6): 1425–1440.

Nguyen TM, Yuan S, Cao M, Lyu Y, Nguyen TH and Xie L (2022) Ntu viral: A visual-inertial-ranging-lidar dataset, from an aerial vehicle viewpoint. *International Journal of Robotics Research* 41(3): 270–280.

Qingqing L, Xianjia Y, Queralta JP and Westerlund T (2022) Multi-modal lidar dataset for benchmarking general-purpose localization and mapping algorithms. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. pp. 3837–3844.

Shan T, Englot B, Duarte F, Ratti C and Rus D (2021) Robust place recognition using an imaging lidar. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. pp. 5469–5475.

Vidanapathirana K, Moghadam P, Harwood B, Zhao M, Sridharan S and Fookes C (2021) Locus: Lidar-based place recognition using spatiotemporal higher-order pooling. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. IEEE, pp. 5075–5081.

Vidanapathirana K, Ramezani M, Moghadam P, Sridharan S and Fookes C (2022) Logg3d-net: Locally guided global descriptor learning for 3d place recognition. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. pp. 2215–2221.

Wang H, Wang C and Xie L (2020) Intensity scan context: Coding intensity and geometry relations for loop closure detection. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. pp. 2095–2101.

Xu W, Cai Y, He D, Lin J and Zhang F (2022a) Fast-lio2: Fast direct lidar-inertial odometry. *IEEE Transactions on Robotics* 38(4): 2053–2073.

Xu X, Lu S, Wu J, Lu H, Zhu Q, Liao Y, Xiong R and Wang Y (2022b) Ring++: Roto-translation invariant gram for global localization on a sparse scan map. *arXiv preprint arXiv:2210.05984* .

Yuan C, Lin J, Liu Z, Wei H, Hong X and Zhang F (2024) Btc: A binary and triangle combined descriptor for 3-d place recognition. *IEEE Transactions on Robotics* 40: 1580–1599.

Yuan C, Lin J, Zou Z, Hong X and Zhang F (2023) Std: Stable triangle descriptor for 3d place recognition. In: *Proceedings of the IEEE International Conference on Robotics and Automation*.

Zhang Y, Jin R and Zhou ZH (2010) Understanding bag-of-words model: a statistical framework. *International journal of machine learning and cybernetics* 1: 43–52.