Multi-Session Mapping and Long-Term Localization for Autonomous Vehicles Using Radar

Daniel Casado Herraez, *Member, IEEE*, Matthias Zeller, *Graduate Student Member, IEEE*, Dong Wang, *Graduate Student Member, IEEE*, Jens Behley, *Member, IEEE*, Michael Heidingsfeld, and Cyrill Stachniss, *Member, IEEE*

***Abstract*—Localization of autonomous vehicles in existing maps is crucial for reliable navigation. Using previously constructed maps allows vehicles to estimate their pose without the inherent odometry drift. Building such maps involves aligning data recorded at different times and maintaining the map over time. While LiDAR sensors are commonly used for mapping due to their high accuracy, they are sensitive to adverse weather and involve high production costs. In this letter, we address the problem of long-term mapping and localization leveraging automotive radars, which are robust to weather conditions and offer a cost-effective alternative to LiDARs. In our approach, we construct maps of coinciding areas and align them by performing place recognition between them. Additionally, our multi-sequence loop detection and verification strategy for radar sensors is able to filter incorrect loop matches, enhancing trajectory alignment. Then, our novel map maintenance module handles radar noise and preserves persistent map points that remain reliable for localization. Subsequently, we estimate the robot poses in the resulting map by combining local odometry with scan-to-map matching, overcoming the complexities of sparse automotive radar data. We evaluate our method on public automotive radar datasets. The results show that our approach achieves state-of-the-art trajectory alignment, preserves persistent map points for localization, and reliably localizes within the constructed maps.**

***Index Terms*—Mapping, localization, SLAM.**

I. Introduction

AUTONOMOUS vehicles hold transformative potential for safe and efficient transportation, attracting growing research interest within the robotics community. Central to their operation is their ability to localize accurately in the environment. This task is commonly achieved by using GNSS systems or onboard simultaneous localization and mapping. However, a good GNSS signal may not always be available in tunnels, parking lots, urban canyons, or under dense foliage, and onboard pose estimation may lead to errors accumulating over time. When a map of the scene is available, however, incoming measurements can be compared to the map, leading to a reduction of the accumulated pose error. Therefore, it is crucial for localization that the maps accurately represent the scene, requiring map updates to account for potential changes in the environment over time. Map construction and localization are commonly achieved using LiDARs [23] [24] [28]. However, these sensors are not robust to adverse weather and come at high production costs. In this work, we tackle the challenge of accurate long-term mapping and localization leveraging automotive radars, which are already integrated into consumer vehicles today. Their main advantages are resilience to environmental conditions [3] [5] and low prices compared to LiDARs. Moreover, radars provide additional information beneficial for localization, including Doppler velocity and radar cross-section (RCS) of each measurement. The Doppler velocity values measure the relative radial velocity of the target, and the RCS measures the detectability of an object based on its material and angle of incidence. Nevertheless, radar scans are inherently sparse and affected by noise and multi-path propagation. As a result, directly applying LiDAR methods to radar data often fails. The sparsity of radar scans limits loop detection performance [6], and the high amounts of noise lead to incorrect point-to-point associations when detecting map differences [8]. Moreover, scan-to-map matching becomes unreliable in scenes with substantial changes, since the low point density combined with a high rate of mismatches degrades pose estimation accuracy. Existing radar solutions achieve accurate pose estimation and map construction in unknown environments [9] [35] [36]. However, these methods solely rely on information available at runtime, and do not exploit prior knowledge of locations that have been mapped in the past. This can lead to pose drift, particularly if no loops have been identified. Furthermore, their online mapping strategy does not account for map changes over time, which is critical for accurate map-based localization [24].

The main contribution of this letter is a system that enables multi-session mapping and localization leveraging automotive radars. To the best of our knowledge, this work presents one of the first long-term mapping and localization approaches leveraging this sensor modality. We exploit established components for radar SLAM, radar place recognition, and anchor map alignment, to address the limitations of LiDAR-based systems when applied to automotive radar data. Specifically, we overcome these limitations with a loop detection and filtering strategy

Received 21 August 2025; accepted 9 December 2025. Date of publication 12 January 2026; date of current version 20 January 2026. This article was recommended for publication by Associate Editor F. Caballero and Editor J. Civera upon evaluation of the reviewers' comments. *(Corresponding author: Daniel Casado Herraez.)*

Daniel Casado Herraez is with CARIAD SE, 38440 Wolfsburg, Germany, and also with the University of Bonn, 53115 Bonn, Germany (e-mail: dcasadoherraez@gmail.com).

Matthias Zeller and Michael Heidingsfeld are with CARIAD SE, 38440 Wolfsburg, Germany.

Dong Wang is with the Department of Informatics XVII Robotics, Julius Maximilians-University, 97070 Wuerzburg, Germany.

Jens Behley is with the Center for Robotics, University of Bonn, 53115 Bonn, Germany.

Cyrill Stachniss is with the Center for Robotics, University of Bonn, 53115 Bonn, Germany, and also with the Lamarr Institute for Machine Learning and Artificial Intelligence, 44227 Dortmund, Germany.

The project page of the paper is available at https://www.ipb.uni-bonn.de/html/projects/long-term-localization-and-mapping.

Digital Object Identifier 10.1109/LRA.2026.3653332

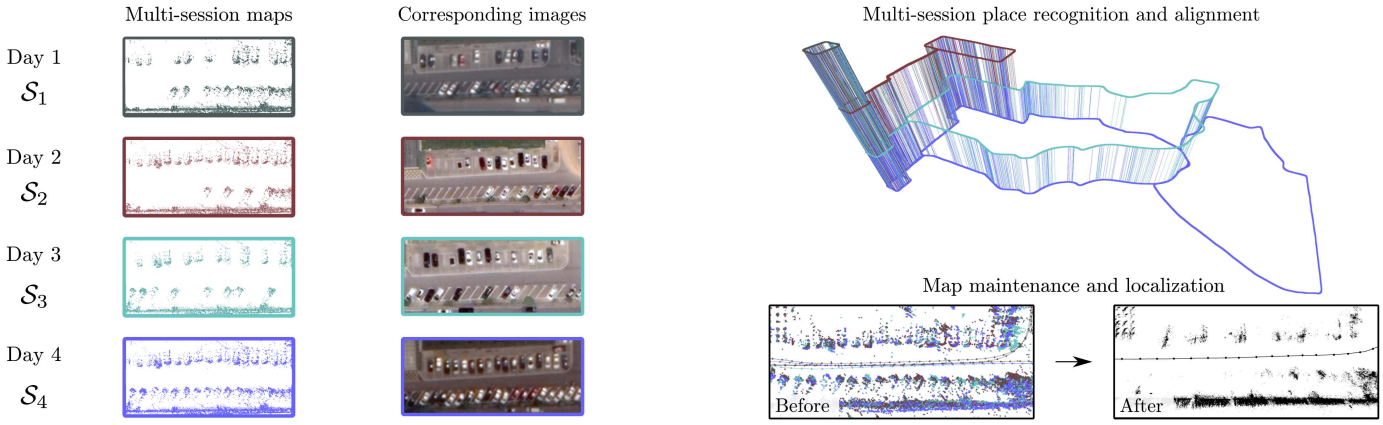

Fig. 1. Visualization of our multi-session mapping and long-term localization system on the SNAIL-Radar dataset [19]. Scenes evolve across sessions $(\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4)$ visible in the radar maps, with their corresponding Google Earth satellite image (left). We perform radar place recognition and alignment across sessions (top right) and employ map maintenance and localization within the map (bottom right).

tailored for radar point clouds, which identifies loops across all recorded sequences, and a temporal 3D grid map maintenance procedure to mitigate radar noise. We also introduce a simple, yet effective, localization method robust to environmental changes and the high sparsity of radar scans. We design each component explicitly for automotive radars, enabling multi-session mapping and localization where prior LiDAR solutions fail. In sum, we make three key claims. Our work (i) achieves state-of-the-art alignment accuracy of multiple SLAM sessions recorded at different points in time; (ii) can merge radar maps from multiple sessions, preserving persistent and reliable points within a compact map; (iii) achieves accurate radar pose estimation in the resulting map by combining odometry and localization for scan matching.

II. Related Work

We present an overview of point cloud-based methods in multi-session mapping, map maintenance, and long-term localization. Due to the limited work on radar in this context, we also include LiDAR methods and their limitations.

Multi-session mapping involves building maps from data acquired at different times and aligning them in a common reference frame. Various approaches exist to maintain global consistency of maps in large-scale environments [4] [26] or perform alignment of submaps [15] [27]. In particular, Kim et al. [24] and Yang et al. [39] address multi-session mapping on LiDAR sensors. They align multiple sessions by performing loop detections on a single reference session, which creates a high dependence on the pose accuracy and the coverage of the reference map. Moreover, they do not perform loop verification after optimization, which can lead to incorrect matches, especially in sparse radar point clouds [6]. To date, there are no multi-session mapping approaches designed for radar data, requiring specialized methods that handle the noise and sparsity of radar scans.

Map maintenance consists of identifying variations in the environment over time. Objects can be categorized into dynamic and movable objects. Dynamic objects, such as pedestrians and moving cars, are commonly identified online using learning-based techniques [7] [43] or dynamic outlier rejection [21]. Movable objects may, or may not, move during a mapping session, hence not always being identifiable at runtime. Reasons for this may be due to false negatives from failing object detectors or certain objects not being present within the training dataset. Offline LiDAR approaches [13] [23] [24] [28] [29] detect moving objects and erase them from existing maps by estimating the consistency between measurements, where multiple LiDAR maps are combined and their difference is extracted. A focus is also placed on how eliminating temporal changes affects the final map size, map compression, or the localization performance [18] [24] [30] [42]. These methods, however, are tailored for erasing outliers from dense and accurate LiDAR scans. Radars, on the contrary, output sparse scans with high amounts of random noise, clutter, and multi-path propagation, leading to challenges when computing scan-to-map correspondences and identifying map differences.

Long-term localization refers to estimating the pose of a vehicle in a previously created map within a potentially changing environment. Approaches can be categorized into global and relative localization. Global localization, also known as place recognition, estimates the global pose of the robot on the entire map and is generally used for computing the first pose for relative localization. LiDAR approaches work well with dense and accurate point clouds [10] [22] [38]. Meanwhile, radar methods exploit the nature of radar scans achieving higher performance in the radar domain [6] or localizing in different map modalities [31] [32] [40]. Relative localization methods assume a known initial pose. Although there is little published work specific to relative localization, some authors match LiDAR scans to high-definition maps [37] [41] and others directly perform scan matching to the prior map [1] [2] [4]. This kind of localization is generally achieved using scan-to-map point cloud registration [34]. However, during localization, incorrect point correspondences may lead to failure of scan matching. These failures often come from movable scene elements that have changed since the time of the creation of the map. Some long-term localization approaches [12] [33] alternate between odometry and localization based on the number of point correspondences, and others predict stable points of the environment during operation [16]. These methods, however, rely on scan-to-map matching of dense LiDAR scans and fail when applied to sparse radar scans.

We present a long-term mapping approach with map maintenance and localization specifically designed for automotive

radars. It overcomes the limitations of LiDAR-based methods [24] [34], which degrade when applied to radar data. An inter-session loop matching and verification strategy reduces incorrect matches from radar noise and sparsity, and our map maintenance procedure identifies temporal map changes from noisy radar point clouds. Finally, our localization method is robust to changes in the scene and the low point density of radar scans, enabling reliable radar long-term pose estimation.

III. Our Approach to Multi-Session Mapping and Long-Term Radar Localization

Our approach achieves accurate mapping and localization by leveraging automotive radar sensors. We first perform radar SLAM [9] over multiple sequences with overlapping locations. We find the overlap by performing place recognition [6] between all sequences and align the maps using a global offline optimization approach. We, then, perform map maintenance by comparing the sequences recorded at different times and removing non-persistent points from the map that belong to movable objects or noise. Finally, we achieve low-drift localization in the resulting map.

A. Multi-Session Mapping

Rather than relying on a single recording to construct a map of the environment, the goal of multi-session mapping is to obtain a globally aligned and accurate map by combining recordings taken at different points in time. To achieve this, we build upon the LiDAR multi-session alignment approach from LT-Mapper [24] and adapt it to operate with automotive radars. We perform SLAM to build the maps of multiple overlapping sequences, and identify coinciding locations between sessions. We build a global pose graph containing all odometry and loop constraints and perform a first optimization step. We then remove infeasible constraints and perform a second optimization step to refine the graph.

We leverage RaI-SLAM [9] to construct maps of multiple sequences with overlapping locations. Each individual sequence $\mathcal{S}_i = \{\mathcal{K}_{i,1}, \mathcal{K}_{i,2}, \ldots, \mathcal{K}_{i,M}\}$ is defined as a set of keyframes, where each keyframe $\mathcal{K}_{i,j} = (\mathsf{T}_{i,j}, \mathcal{P}_{i,j}, \mathbf{d}_{i,j})$ contains the pose $\mathsf{T}_{i,j} \in SE(3)$, the individual scans $\mathcal{P}_{i,j}$, and the place recognition descriptor $\mathbf{d}_{i,j} \in \mathbb{R}^{256}$.

Note that each sequence is recorded independently with the estimated poses given relative to the start of the sequence. We, therefore, need to identify overlapping keyframes between the sequences and align them to a common reference frame. We leverage the SPR place recognition approach [6] which exploits the radar RCS information, to identify one-to-one location correspondences. We exploit three distinct criteria to remove incorrect detections between sequences: descriptor similarity, point correspondences, and geometric distances between matched locations.

First, we measure the L2 distance similarity between two scan descriptors, $\mathbf{d}_{\text{query}}, \mathbf{d}_{\text{match}} \in \mathbb{R}^{256}$ following

$$\|\mathbf{d}_{\text{query}} - \mathbf{d}_{\text{match}}\| < \delta_{\text{PR}}, \tag{1}$$

where δ_{PR} is a predefined threshold. We, then, align both scans via iterative closest point (ICP) and perform a point correspondence check to mitigate false positives due to perceptual ambiguity. Unlike LT-Mapper [24], we leverage the correspondence matching quality, which has shown to be more effective in the radar domain than the fitness score [9]. Due to the noise of radar scans, the average fitness score may be high even with good point correspondences. We classify a location as similar if the number of correspondences within a radius r is higher than a threshold δ_{d}, following

$$\left[\frac{1}{|\mathcal{C}|}\sum_{(\mathbf{q},\mathbf{p})\in\mathcal{C}} \mathbb{I}\{\|\mathbf{q}-\mathbf{p}\| < r\}\right] > \delta_{\text{d}}, \tag{2}$$

where $\mathcal{C} = \{(\mathbf{q}, \mathbf{p})\}$ is the set of closest point correspondences between the query $\mathcal{Q} = \{\mathbf{q} \in \mathbb{R}^3\}$ and the matching point cloud $\mathcal{P} = \{\mathbf{p} \in \mathbb{R}^3\}$, and $\mathbb{I}\{c\}$ is the indicator function, returning 1 if condition c is true, and 0 otherwise.

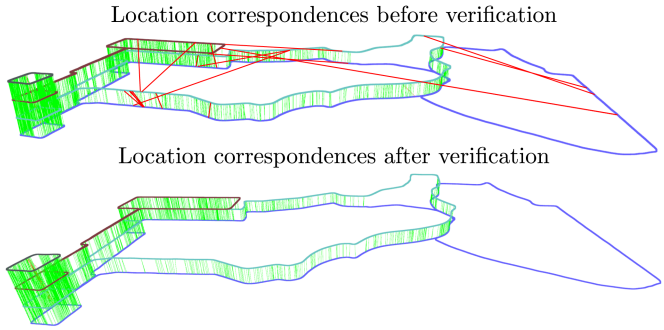

Fig. 2. Our filtering of incorrect place recognitions (red).

After identifying inter-session constraints, we construct a global pose graph following the anchor node method by Kim et al. [24]. The intra-session loop and odometry pose constraints are added to each sequence, and the inter-session correspondences are added with respect to the anchors from a central sequence. We perform an initial optimization step [20] to obtain the aligned sequences in a common frame. However, incorrect location correspondences may lead to inaccurate trajectories. We propose an additional verification step to remove ambiguous location correspondences by measuring the Euclidean distance between them, see Fig. 2. We delete all correspondences that have a distance larger than a threshold δ_{E} and perform a second global optimization step. This refines the initial results, leading to more accurate trajectories. We further extend LT-Mapper [24] by pairing all sequences. This strategy enables direct merging of multiple recordings into a single map, reducing potential deviations caused by an inaccurate central sequence. Additionally, instead of re-optimizing all trajectories when new sequences are recorded, the map can be updated incrementally by pairing each new session with the existing map as a reference.

B. Map Maintenance

Environments are dynamic, with vehicles arriving and departing, changes caused by constructions, and other elements evolving over time. To achieve reliable localization and avoid wrong point matches during scan registration, it is crucial that only persistent map points are kept. Existing LiDAR approaches compute map changes by leveraging dense and accurate point clouds, typically querying neighboring points and measuring the distances between them [16] or projecting LiDAR scans to images [24]. However, radar maps often contain noise outliers and sparse point clouds, resulting in unreliable correspondences.

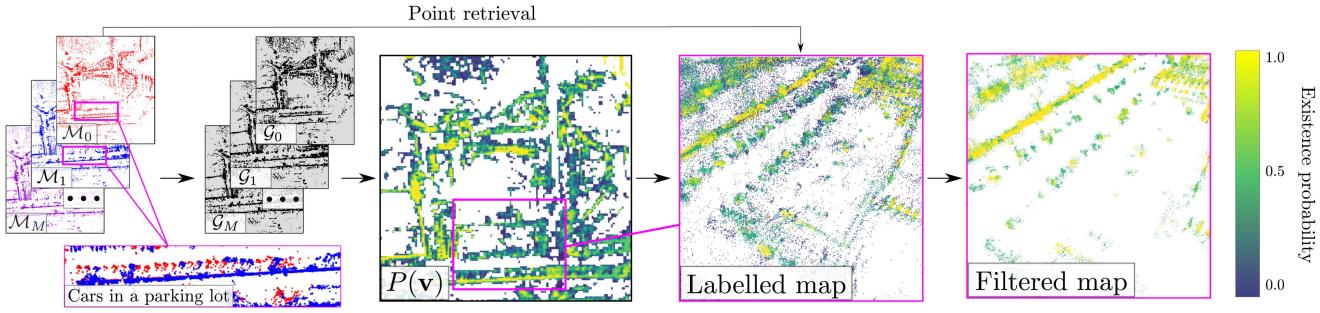

Fig. 3. Our map maintenance procedure. We convert the point cloud maps into occupancy grids, estimate the occupancy probability of each voxel, and retrieve the labeled points with their existence probability. We, then, filter out points with labels where $P(\mathbf{v}) < \beta$.

Inspired by Hroob et al. [17] and Gil et al. [13], we differentiate between three types of points in a map: dynamic, volatile, and persistent. Dynamic points belong to moving objects during SLAM. Unlike LiDARs [23] [28] [29], automotive radars provide per-point Doppler velocities. To filter out dynamic objects, we compute the instantaneous ego-velocity and discard points with a motion-compensated relative velocity higher than a threshold δ_{dyn}, following the approach by Kellner et al. [21]. Volatile points belong to objects that change their location across sessions, such as radar noise and stopped buses. Additionally, persistent points belong to objects that remain in the scene over time, which can be associated with structures like buildings or parking spots that remain occupied. However, volatile and persistent points may not be directly identifiable at runtime, requiring specialized map maintenance procedures. We propose a strategy, illustrated in Fig. 3, that leverages 3D grid maps to compute the map occupancy probability over multiple sessions. We then filter volatile points with low probability of existence, resulting in maps containing only persistent points that can be leveraged for localization.

To represent the environment, we convert the set of radar maps $\mathcal{M} = \{\mathcal{M}_0, \mathcal{M}_1, \ldots, \mathcal{M}_m, \ldots, \mathcal{M}_M\}$ into a set of 3D binary occupancy voxel grids $\mathcal{G} = \{\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_m, \ldots, \mathcal{G}_M\}$. The goal is to compute the occupancy probability of the voxels corresponding to each map. However, not all maps have the same coverage. While maps will partially overlap, some maps may contain poses covering a larger area than others. Therefore, it is essential that only maps that cover a specific voxel are taken into account when estimating its probability of being occupied. Given the set of all voxels that are present at least in one occupancy map, $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_V\}$, with each voxel defined by its center position $\mathbf{v} \in \mathbb{R}^3$, we denote the map coverage $C_m(\mathbf{v})$ as a binary indicator of whether any k^{th} position of map $\mathcal{M}_m$, defined as $\mathbf{t}_{m_k} \in \mathbb{R}^3$, is within the radar range ρ of voxel $\mathbf{v}$ such that

$$C_m(\mathbf{v}) = \mathbb{I}\{\|\mathbf{t}_{m_k} - \mathbf{v}\| < \rho\}. \tag{3}$$

We compute the number of maps N_{cov} covering a specific voxel $\mathbf{v}$ following

$$N_{\text{cov}}(\mathbf{v}) = \sum_{m \in \mathcal{M}} C_m(\mathbf{v}), \tag{4}$$

and the number of maps that cover an occupied voxel $N_{\text{cov+occ}}$ according to

$$N_{\text{cov+occ}}(\mathbf{v}) = \sum_{m \in \mathcal{M}} C_m(\mathbf{v})\, \mathcal{G}_m(\mathbf{v}), \tag{5}$$

where $\mathcal{G}_m(\mathbf{v})$ is 1 if an occupied voxel $\mathbf{v}$ exists in the grid map $\mathcal{G}_m$, and 0 otherwise. We estimate the final occupancy probability $P(\mathbf{v})$ with

$$P(\mathbf{v}) = \frac{N_{\text{cov+occ}}(\mathbf{v})}{N_{\text{cov}}(\mathbf{v})}. \tag{6}$$

Note that our approach also accounts for new structural elements. The occupancy probability increases as more sessions contain that occupied voxel, eventually considering the point as persistent. Moreover, by storing two hash maps for $N_{\text{cov}}(\mathbf{v})$ and $N_{\text{cov+occ}}(\mathbf{v})$, the probability $P(\mathbf{v})$ of affected voxels can be incrementally updated with new incoming sessions. To obtain the final map, we retrieve the points from their corresponding voxels and assign to each point the voxel's occupancy probability as its existence probability. During localization, we discard points with an existence probability below a threshold β.

C. Radar Localization

The goal of the previous steps was to create a map suitable for long-term localization. Contrary to odometry, which solely relies on an online map, having an accurate pre-built map reduces uncertainty during pose estimation. Some approaches in the LiDAR domain [1] [2] directly match the current measured scan with the pre-built map. However, radar scans are sparse, and direct scan-to-map matching leads to estimation errors in changing environments, see Fig. 4. Our simple, yet effective, approach combines scan-to-map localization with odometry and leverages radar Doppler velocities, leading to an accurate trajectory. We register the current scan $\mathcal{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N\}$ jointly against two spatial representations: the locally accumulated odometry map $\mathcal{M}_L = \{\mathbf{m}_{L_1}, \mathbf{m}_{L_2}, \ldots, \mathbf{m}_{L_M}\}$, and the globally consistent map $\mathcal{M}_G = \{\mathbf{m}_{G_1}, \mathbf{m}_{G_2}, \ldots, \mathbf{m}_{G_W}\}$ obtained following our map maintenance procedure in Section III-B. However, a naive combination of the global and the local maps into the same voxel hash map [34] can lead to a bias towards odometry, as its points tend to have more matches due to higher scan similarity, neglecting the global prior map. We approach this with a weighted scan-to-map registration strategy that explicitly biases the optimization toward the global map, while still leveraging the local map. During ICP alignment, we assign a higher

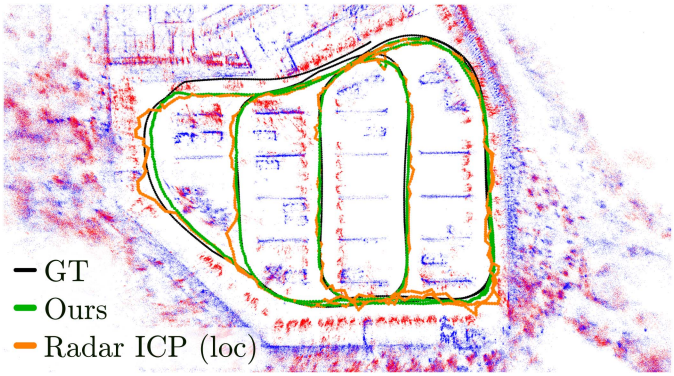

Fig. 4. Matching between current radar scans (blue) and map (red) is error-prone due to scene changes and sparsity.

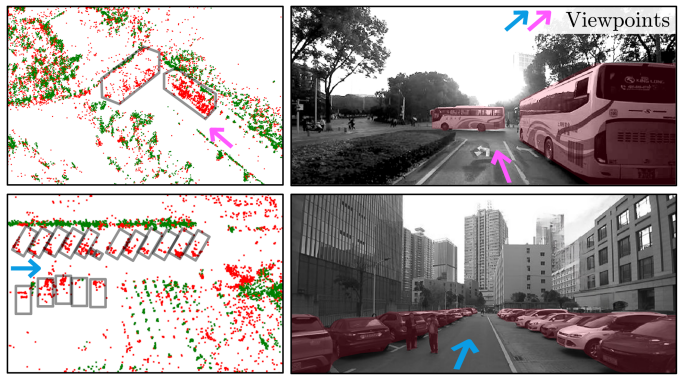

Fig. 5. (left) Radar map with volatile points in (red) and persistent points in (green). (right) Camera image with movable objects in red.

weight to correspondences from the global map $\mathcal{M}_G$, guiding the solution toward globally consistent poses. We estimate the relative transformation $\mathsf{T}^* \in SE(3)$ that combines the global and local map following

$$\mathsf{T}^* = \underset{\mathsf{T}\in \mathrm{SE}(3)}{\operatorname{argmin}} \left[w_L \cdot \sum_{(\mathbf{p},\mathbf{m}_L)\in\mathcal{C}_L} \rho\left(|\mathbf{p} - \mathsf{T}\mathbf{m}_L|\right) + w_G \cdot \sum_{(\mathbf{p},\mathbf{m}_G)\in\mathcal{C}_G} \rho\left(|\mathbf{p} - \mathsf{T}\mathbf{m}_G|\right) \right], \tag{7}$$

where $\mathcal{C}_L$, $\mathcal{C}_G$ are the sets of correspondences between the current scan and the local and global map, w_L, w_G are balancing weights for the local and global maps, and ρ is a robust Geman McClure kernel [34] for correspondence outliers. Additionally, we leverage the radar Doppler velocities to compute the initial ICP estimate, where the measured velocity $v_{r,i}$ for a static point $\mathbf{p}_i \in \mathbb{R}^3$ is related to the sensor's linear and angular velocity $\boldsymbol{v}_s, \boldsymbol{\omega}_s \in \mathbb{R}^3$ following

$$v_{r,i} = \frac{\mathbf{p}_i^\top}{\|\mathbf{p}_i\|}(-\boldsymbol{v}_s - \boldsymbol{\omega}_s \times \mathbf{p}_i). \tag{8}$$

By collecting multiple points and their velocities, we build a least-squares problem that computes the estimated sensor motion. Given the time difference Δt between scans, the instantaneous ego-motion is integrated to produce a rigid-body transform $\mathsf{T}_{\text{init}} \in SE(3)$ that serves as the initial guess for ICP scan matching. For a detailed description of this procedure, we refer the reader to RaI-SLAM [9].

IV. Experimental Evaluation

The focus of our work is a long-term mapping and localization approach that combines multiple maps, performs map maintenance, and localizes within the resulting map. The experiments show the capabilities of our method, explicitly supporting our claims that our approach: (i) achieves state-of-the-art alignment accuracy of multiple SLAM sessions recorded at different points in time; (ii) can merge radar maps from multiple sessions, preserving persistent and reliable points within a compact map; (iii) achieves accurate radar pose estimation in the resulting map by combining odometry and localization for scan matching.

A. Implementation Details

We implement our system using ROS2 for communication between modules and GTSAM [11] for factor graph optimization. For place recognition, we exploit SPR [6]. The evaluation parameters include the place recognition threshold $\delta_{\text{PR}} = 0.6$, the point matching radius $r = 0.5$ m and ratio $\delta_{\text{d}} = 0.3$, the correspondence threshold $\delta_{\text{E}} = 5$ m, the dynamic velocity threshold $\delta_{\text{dyn}} = 0.1$ m/s, a voxel grid size of 1 m, the radar range $\rho = 50$ m, a point probability threshold $\beta = 0.6$ and the global and local map weights $w_G = 10$ $w_L = 1$ for the SNAIL-Radar [19] dataset and $w_G = 2$, $w_L = 1$ for the HeRCULES [25] dataset. During map generation, we assume a planar vehicle trajectory. The evaluation measures the absolute position (APE) and rotation (ARE) error with respect to the GPS ground truth [14].

B. Experimental Setup

We evaluate our system on the SNAIL-Radar [19] and HeRCULES [25] datasets using the ARS548 radar. For multi-session alignment, we run SLAM on sequences "20240113/1","20240123/3", "20240113/3", "20240115/2" and "20240116/2" and perform our two-step global optimization between them. We also use the aforementioned sequences during our ablations to analyze the contribution of each component. Additionally, we evaluate the parking sequences from the HeRCULES dataset: "Parking Day1 1", "Parking Day2 1", "Parking Day2 2" and "Parking Night". We then perform multi-session alignment across all sequences for each dataset, treating every sequence as reference. In each sequence, we report the mean error across all alignments. Furthermore, to observe the impact of the SLAM system on our multi-session alignment, we evaluate our approach leveraging the RaI-SLAM [9] and RIV-SLAM [36] backends. We modify them to output the same pose, point cloud, and descriptor formats, and select the most reliable backend for the rest of the experiments. To evaluate our localization approach in the SNAIL-Radar dataset, we localize sequences, "20240116_eve/5" and "20240113/1", within a map built with "20231208/4", "20231213/1", "20231213/4", "20231213/5", "20240115/3", "20240116/5" and "20240123/3". We also localize sequences "20231208/5", and "20231213/2" within a map constructed from sequences "20231201/3", "20240113/1" and "20240113/3". We then perform map maintenance following Section III-B. In the HeRCULES dataset, we

TABLE I
MULTI-SESSION ALIGNMENT PERFORMANCE ON THE SNAIL-RADAR DATASET [19]. "N/A" INDICATES A FAILED RUN

	20240113/1		20240123/3		20240113/3		20240115/2		20240116/2		Mean	
	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]
RaI-SLAM [9]	0.193	0.907	3.013	1.221	12.49	1.051	6.758	0.749	45.11	2.179	13.51	1.220
RIV-SLAM [36]	0.418	1.076	4.275	1.016	3.120	0.887	5.991	0.649	N/A	N/A	-	-
LT-Mapper (RIV) [24]	0.454	1.253	1.849	1.132	3.390	1.203	3.310	0.753	N/A	N/A	-	-
Ours (RIV)	**0.206**	**1.075**	**1.568**	**1.049**	**2.512**	**1.009**	**2.544**	**0.582**	N/A	N/A	-	-
LT-Mapper (RaI) [24]	**0.197**	**0.924**	1.821	1.096	2.648	0.666	2.575	0.698	**8.998**	**1.160**	3.248	0.900
Ours (RaI)	0.242	0.956	**1.607**	**0.998**	**2.153**	**0.610**	**2.180**	**0.634**	9.472	1.229	**3.131**	**0.885**

TABLE II
MULTI-SESSION ALIGNMENT PERFORMANCE ON THE HERCULES DATASET [25]

	Parking Day1 1		Parking Day2 1		Parking Day2 2		Parking Night		Mean	
	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]
RaI-SLAM [9]	0.622	2.575	0.466	2.412	0.576	1.015	0.774	3.256	0.609	2.314
RIV-SLAM [36]	0.772	2.146	0.465	2.687	6.602	2.722	0.783	3.520	2.155	2.769
LT-Mapper (RIV) [24]	3.999	9.516	1.057	4.825	**1.283**	6.546	**1.015**	**3.384**	1.838	6.068
Ours (RIV)	**0.740**	**2.833**	**0.347**	**2.745**	3.692	**5.327**	1.064	4.658	**1.461**	**3.891**
LT-Mapper (RaI) [24]	2.001	5.353	0.990	3.323	0.660	1.874	**0.842**	**3.232**	1.123	3.445
Ours (RaI)	**0.610**	**2.516**	**0.571**	**2.530**	**0.620**	**1.383**	1.275	4.196	**0.769**	**2.656**

TABLE III
LOCALIZATION PERFORMANCE ON THE SNAIL-RADAR DATASET [19]. "N/A" INDICATES A FAILED RUN

	20240116_eve/5		20240113/1		20231208/5		20231213/2		Mean	
	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]
RaI-SLAM [9]	3.832	1.365	0.427	1.365	23.25	2.024	26.40	2.453	13.49	1.802
RIV-SLAM [36]	7.109	0.880	0.266	3.454	6.291	1.678	5.736	0.839	4.851	1.713
Radar ICP (odom) [8]	4.437	1.533	1.723	2.040	18.183	1.692	24.04	2.317	12.09	1.896
Radar ICP (loc) [8]	**1.262**	1.190	0.393	**1.298**	1.929	0.800	N/A	N/A	-	-
Ours	1.270	**1.172**	**0.328**	1.342	**1.926**	**0.776**	**1.854**	**1.387**	**1.345**	**1.170**

build, maintain, and localize sequences "Parking Day1", "Parking Day2 1", "Parking Day2 2" and "Parking Night". Similarly, we compare localization accuracy in maintained and not maintained maps, localizing "20240116_eve/5", "20240113/1", "20231208/5" and "20231213/2" in localization maps, and performing uniform voxel downsampling of the maps at 0.5, 1.0, and 1.5 m to evaluate accuracy at different resolutions. In our evaluation, we omit the results of methods with failed runs.

C. Multi-Session Trajectory Alignment

The first experiment supports our claim that our multi-session mapping approach achieves state-of-the-art alignment accuracy of multiple SLAM sessions recorded at different times. We compare our approach against the single-session SLAM results and a state-of-the-art LiDAR approach LT-Mapper [24] for multi-session alignment. Since LT-Mapper is designed for LiDAR sensors, we modify its input to take the poses coming from the radar SLAM backends. Moreover, we swap the ScanContext [22] LiDAR place recognition approach with the radar place recognition method SPR [6] for fair comparison. The results are displayed in Tables I and II, where bold indicates the best alignment for each backend. Our ablations from Fig. 6 report the mean and standard deviation under different configurations.

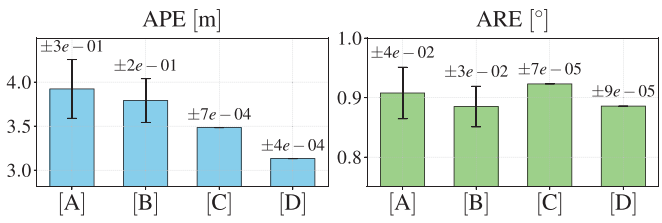

Fig. 6. Ablation study of our multi-session alignment for single sequence matching [A], single sequence with verification [B], multi-sequence matching [C], and multi-sequence with verification [D].

The results show how our method improves the average alignment accuracy in both datasets, with the RaI-SLAM backend showing superior performance. Moreover, RIV-SLAM fails in sequence "20240116/2", meaning that this sequence cannot be considered during alignment. Furthermore, multi-session alignment presents a clear advantage against single-session SLAM since it is able to correct for loop closures that may not have taken place. In addition, our ablations demonstrate how the loop verification contributes to the overall accuracy. While the improvement in positional accuracy is clear, the orientation error reflects a tradeoff from multi-sequence configurations

TABLE IV
LOCALIZATION PERFORMANCE ON THE HERCULES DATASET [25]

	Parking Day1 1		Parking Day2 1		Parking Day2 2		Parking Night		Mean	
	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]	APE [m]	ARE [°]
RaI-SLAM [9]	1.770	2.379	1.778	1.181	1.161	1.772	2.385	1.374	1.774	1.442
RIV-SLAM [36]	0.787	2.230	0.487	2.799	6.644	2.920	0.847	3.054	2.191	2.751
Radar ICP (odom) [8]	3.164	4.528	1.788	1.297	2.202	1.957	3.619	1.557	2.536	1.604
Radar ICP (loc) [8]	1.273	2.776	1.586	1.986	1.344	2.711	2.345	6.698	1.401	2.491
Ours	**1.116**	**2.064**	**1.466**	**1.574**	**1.332**	**1.770**	**2.038**	**6.314**	**1.305**	**1.803**

TABLE V
COMPARISON OF LOCALIZATION PERFORMANCE WITH RELIABLE POINT PRESERVATION IN MAINTAINED MAPS AND RAW POINT CLOUD AGGREGATION IN NON-MAINTAINED MAPS AT DIFFERENT VOXEL RESOLUTIONS IN THE SNAIL-RADAR DATASET [19]. "N/A" INDICATES A FAILED RUN

		20240116_eve/5			20240113/1			20231208/5			20231213/2		
Voxel size [m]	Map type	Map size [MB]	APE [m]	ARE [°]	Map size [MB]	APE [m]	ARE [°]	Map size [MB]	APE [m]	ARE [°]	Map size [MB]	APE [m]	ARE [°]
None	Not maintained	149.3	1.293	**1.144**	165.7	0.476	**1.119**	159.4	**1.929**	0.809	167.9	N/A	N/A
	Maintained	**106.8**	**1.254**	1.161	**119.6**	**0.420**	1.190	**124.6**	1.934	**0.797**	**131.1**	**1.816**	**1.313**
0.5	Not maintained	58.89	1.288	1.188	62.03	0.522	1.245	81.39	**1.924**	0.805	83.83	N/A	N/A
	Maintained	**34.34**	**1.259**	**1.171**	**35.92**	**0.413**	**1.233**	**57.10**	1.925	**0.800**	**58.64**	**1.817**	**1.318**
1.0	Not maintained	24.42	1.276	1.170	25.21	0.432	1.085	38.57	**1.912**	**0.799**	39.27	**1.811**	1.353
	Maintained	**12.21**	**1.274**	**1.155**	**12.46**	**0.360**	**0.977**	**24.23**	1.917	0.816	**24.58**	1.812	**1.340**
1.5	Not maintained	12.99	**1.286**	**1.122**	13.31	0.476	1.089	22.03	1.926	0.892	22.35	1.825	1.424
	Maintained	**6.478**	1.339	1.177	**6.572**	**0.464**	**1.040**	**13.49**	**1.914**	**0.835**	**13.64**	**1.806**	**1.371**

constraining the poses with respect to all trajectories. In such cases, the optimization may compromise a well-aligned trajectory to improve consistency across sessions. Therefore, single-sequence alignment may present superior performance when relying on a single session with high trajectory accuracy, but will substantially degrade if the reference session is inaccurate. Our approach, on the contrary, is able to mitigate reference trajectory inaccuracies, visible in the reduced standard deviation.

D. Localization in Radar Maps

The second experiment supports our claim that our radar localization method achieves accurate radar trajectory estimates in the resulting map. Due to the lack of map-based localization approaches, we adapt Radar-ICP [8] to perform scan matching against a pre-built map and use it as our baseline. We display the best map-based localization approach in bold, and also provide online pose estimation methods that do not leverage a prior map including the original Radar-ICP [8], RaI-SLAM [9] and RIV-SLAM [36]. The results are displayed in Tables III, IV, and Fig. 4.

The evaluation shows that odometry, which operates without any prior map, leads to substantial drift. Online SLAM approaches, on the contrary, achieve a high accuracy in shorter parking lot sequences, but can result in drift over time. Localization approaches, however, are able to achieve reduced drift by leveraging a pre-computed map. Moreover, as displayed in Fig. 4, our method is able to overcome the sparsity of automotive radar scans and substantial environmental changes, where simple scan-to-map matching fails.

E. Maintained Maps for Localization and Storage

The third experiment supports our claim that our map maintenance approach can merge radar maps from multiple sessions, preserving persistent and reliable points within a compact map. To validate our method, we perform standard scan-to-map localization without relying on the odometry. We verify the performance on "Maintained" maps, which include only persistent points, and "not Maintained" maps, which are simply constructed by merging all session maps without filtering. We display the results in Table V, where bold highlights the best performing method.

The evaluation demonstrates that our map maintenance approach is able to reduce the map size while still maintaining superior overall localization accuracy. Interestingly, in sequence "20231213/2", localization in the non-maintained map fails at low compression rates due to the high amount of incorrect point correspondences, but manages to work after heavy downsampling, when many of these wrong correspondences are removed. Additionally, as shown in Fig. 5, our map maintenance strategy effectively removes volatile points from the environment belonging to movable objects and preserves structural elements like buildings and trees.

V. CONCLUSION

In this work, we presented a novel approach for multi-session map construction and localization of autonomous vehicles leveraging automotive radars. We address the limitations of LiDAR-based techniques when applied to radar data. Specifically, our radar strategy performs loop detection and verification, reducing the dependence on a single reference sequence, leverages a temporal occupancy grid to filter out noise and movable objects from radar maps, and provides reliable vehicle pose estimation by combining localization, odometry, and radar Doppler velocities. We implemented and evaluated our approach on real-world scenarios supporting all claims made in this letter. We demonstrated how our multi-session

alignment technique achieves state-of-the-art performance on automotive radar maps and how our map maintenance procedure is able to identify temporal changes in the aligned maps. Finally, our localization approach achieves accurate pose estimation in sparse and maintained radar maps, collectively enhancing the capabilities of long-term vehicle navigation with automotive radar data.

REFERENCES

[1] I. Alamanos, G. P. Moustris, and C. S. Tzafestas, "Localization and offline mapping of high-voltage substations in rough terrain using a ground vehicle," in *Proc. Mediterranean Conf. Control Automat.*, 2024, pp. 107–112.

[2] J. Baek, J. Park, S. Cho, and C. Lee, "3D global localization in the underground mine environment using mobile LiDAR mapping and point cloud registration," *Sensors*, vol. 22, no. 8, 2022, Art. no. 2873.

[3] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, "The Oxford radar robotcar dataset: A radar extension to the Oxford robotcar dataset," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2020, pp. 6433–6438.

[4] J. L. Blanco-Claraco, "A modular optimization framework for localization and mapping," in *Proc. Robot.: Sci. Syst.*, 2019.

[5] K. Burnett, Y. Wu, D. J. Yoon, A. P. Schoellig, and T. D. Barfoot, "Are we ready for radar to replace LiDAR in all-weather mapping and localization?," *IEEE Robot. Automat. Lett.*, vol. 7, no. 4, pp. 10328–10335, Oct. 2022.

[6] D. Casado Herraez et al., "SPR: Single-scan radar place recognition," *IEEE Robot. Automat. Lett.*, vol. 9, no. 10, pp. 9079–9086, Oct. 2024.

[7] D. Casado Herraez, Y. Kawasaki, and M. Takahashi, "Onboard semantic mapping for action graph estimation," in *Proc. 13th Int. Workshop Robot Motion Control*, 2024, pp. 76–81.

[8] D. Casado Herraez, M. Zeller, L. Chang, I. Vizzo, M. Heidingsfeld, and C. Stachniss, "Radar-only odometry and mapping for autonomous vehicles," in *Proc. 2024 IEEE Int. Conf. Robot. Automat.*, 2024, pp. 10275–10282.

[9] D. Casado Herraez, M. Zeller, D. Wang, J. Behley, M. Heidingsfeld, and C. Stachniss, "Rai-SLAM: Radar-inertial SLAM for autonomous vehicles," *IEEE Robot. Automat. Lett.*, vol. 10, no. 6, pp. 5257–5264, Jun. 2025.

[10] X. Chen, I. Vizzo, T. Läbe, J. Behley, and C. Stachniss, "Range image-based LiDAR localization for autonomous vehicles," in *Proc. 2021 IEEE Int. Conf. Robot. Automat.*, 2021, pp. 5802–5808.

[11] F. Dellaert and M. Kaess, "Factor graphs for robot perception," *Found. Trends Robot.*, vol. 6, no. 1, pp. 1–139, 2017.

[12] Y. Fang, Y. Li, K. Qian, F. Tombari, Y. Wang, and G. H. Lee, "LiLoc: Lifelong localization using adaptive submap joining and egocentric factor graph," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2024, pp. 8041–8047.

[13] H. Gil, D. Lee, G. Kim, and A. Kim, "Ephemerality meets LiDAR-based lifelong mapping," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2025, pp. 3312–3319.

[14] M. Grupp, "EVO: Python package for the evaluation of odometry and slam," 2017. [Online]. Available: https://github.com/MichaelGrupp/evo

[15] S. Gupta, T. Guadagnino, B. Mersch, I. Vizzo, and C. Stachniss, "Effectively detecting loop closures using point cloud density maps," in *Proc. 2024 IEEE Int. Conf. Robot. Automat.*, 2024, pp. 10260–10266.

[16] I. Hroob, S. Molina, R. Polvara, G. Cielniak, and M. Hanheide, "Learned long-term stability scan filtering for robust robot localisation in continuously changing environments," in *Proc. 2023 Eur. Conf. Mobile Robot.*, 2023, pp. 1–8.

[17] I. Hroob, S. Molina, R. Polvara, G. Cielniak, and M. Hanheide, "Adaptive robot localization in dynamic environments through self-learnt long-term 3D stable points segmentation," *J. Robot. Auton. Syst.*, vol. 181, no. C, 2024, Art. no. 104786.

[18] X. Hu et al., "MS-Mapping: An uncertainty-aware large-scale multi-session LiDAR mapping system," 2024, arXiv:2408.03723.

[19] J. Huai et al., "Snail-radar: A large-scale diverse dataset for the evaluation of 4D-radar-based slam systems," *Int. J. Robot. Res.*, vol. 44, no. 12, pp. 1941–1958, 2024.

[20] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, "ISAM2: Incremental smoothing and mapping with fluid relinearization and incremental variable reordering," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 3281–3288.

[21] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous ego-motion estimation using Doppler radar," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, 2013, pp. 869–874.

[22] G. Kim and A. Kim, "Scan context: Egocentric spatial descriptor for place recognition within 3D point cloud map," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2018, 4802–4809.

[23] G. Kim and A. Kim, "Remove, then revert: Static point cloud map construction using multiresolution range images," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2020, pp. 10758–10765.

[24] G. Kim and A. Kim, "LT-Mapper: A modular framework for LiDAR-based lifelong mapping," in *Proc. 2022 IEEE Int. Conf. Robot. Automat.*, 2022, pp. 7995–8002.

[25] H. Kim et al., "HeRCULES: Heterogeneous radar dataset in complex urban environment for multi-session radar SLAM," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2025, pp. 4649–4656.

[26] M. T. Lázaro, R. Capobianco, and G. Grisetti, "Efficient long-term mapping in dynamic environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2018, pp. 153–160.

[27] M. T. Lazaro, L. M. Paz, P. Pinies, J. A. Castellanos, and G. Grisetti, "Multi-robot SLAM using condensed measurements," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2013, pp. 1069–1076.

[28] H. Lim, S. Hwang, and H. Myung, "ERASOR: Egocentric ratio of pseudo occupancy-based dynamic object removal for static 3D point cloud map building," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 2272–2279, Apr. 2021.

[29] H. Lim et al., "ERASOR2: Instance-aware robust 3D mapping of the static world in dynamic scenes," in *Proc. Robot.: Sci. Syst.*, 2023.

[30] W. Lyu, W. Ke, H. Sheng, X. Ma, and H. Zhang, "Dynamic downsampling algorithm for 3D point cloud map based on voxel filtering," *Appl. Sci.*, vol. 14, no. 8, 2024, Art. no. 3160.

[31] Y. Ma, X. Zhao, H. Li, Y. Gu, X. Lang, and Y. Liu, "RoLM: Radar on LiDAR map localization," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2023, pp. 3976–3982.

[32] A. Nayak, D. Cattaneo, and A. Valada, "RALF: Flow-based global and metric radar localization in LiDAR maps," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2024, pp. 5097–5103.

[33] B. Peng, H. Xie, and W. Chen, "Roll: Long-term robust LiDAR-based localization with temporary mapping in changing environments," in *Proc. 2022 IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2022, pp. 2841–2847.

[34] I. Vizzo, T. Guadagnino, B. Mersch, L. Wiesmann, J. Behley, and C. Stachniss, "KISS-ICP: In defense of point-to-point icp–simple, accurate, and robust registration if done the right way," *IEEE Robot. Automat. Lett.*, vol. 8, no. 2, pp. 1029–1036, Feb. 2023.

[35] D. Wang et al., "Doppler-SLAM: Doppler-aided radar-inertial and LiDAR-inertial simultaneous localization and mapping," *IEEE Robot. Automat. Lett.*, vol. 10, no. 9, pp. 9438–9445, Sep. 2025.

[36] D. Wang, S. May, and A. Nüchter, "RIV-SLAM: Radar-inertial-velocity optimization based graph SLAM," in *Proc. Int. Conf. Automat. Sci. Eng.*, 2024, pp. 774–781.

[37] J. Wen, J. Tang, H. Liu, C. Qian, and X. Fan, "Real-time scan-to-map matching localization system based on lightweight pre-built occupancy high-definition map," *Remote Sens.*, vol. 15, no. 3, 2023, Art. no. 595.

[38] L. Wiesmann, L. Nunes, J. Behley, and C. Stachniss, "KPPR: Exploiting momentum contrast for point cloud-based place recognition," *IEEE Robot. Automat. Lett.*, vol. 8, no. 2, pp. 592–599, Feb. 2023.

[39] L. Yang, S. M. Prakhya, S. Zhu, and Z. Liu, "Lifelong 3D mapping framework for hand-held & robot-mounted LiDAR mapping," *IEEE Robot. Automat. Lett.*, vol. 9, no. 11, pp. 9446–9453, Nov. 2024.

[40] H. Yin, X. Xu, Y. Wang, and R. Xiong, "Radar-to-LiDAR: Heterogeneous place recognition via joint learning," *Front. Robot. AI*, vol. 8, 2021, Art. no. 661199.

[41] K. Yoneda, H. Tehrani, T. Ogawa, N. Hukuyama, and S. Mita, "LiDAR scan feature for localization with highly precise 3-D map," in *Proc. 2014 IEEE Intell. Veh. Symp.*, 2014, pp. 1345–1350.

[42] Z. Yu, Z. Qiao, W. Liu, H. Yin, and S. Shen, "SLIM: Scalable and lightweight LiDAR mapping in urban environments," *IEEE Trans. Robot.*, vol. 41, pp. 2569–2588, 2025.

[43] M. Zeller, J. Behley, M. Heidingsfeld, and C. Stachniss, "Gaussian radar transformer for semantic segmentation in noisy radar data," *IEEE Robot. Automat. Lett.*, vol. 8, no. 1, pp. 344–351, Jan. 2023.