2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

LiDAR-based HD Map Localization using Semantic Generalized ICP with Road Marking Detection

Yansong Gong, Xinglian Zhang, Jingyi Feng, Xiao He and Dan Zhang*

***Abstract*—In GPS-denied scenarios, a robust environmental perception and localization system becomes crucial for autonomous driving. In this paper, a LiDAR-based online localization system is developed, incorporating road marking detection and registration on a high-definition (HD) map. Within our system, a road marking detection approach is proposed with real-time performance, in which an adaptive segmentation technique is first introduced to isolate high-reflectance points correlated with road markings, enhancing real-time efficiency. Then, a spatio-temporal probabilistic local map is formed by aggregating historical LiDAR scans, providing a dense point cloud. Finally, a LiDAR bird's-eye view (LiBEV) image is generated, and an instance segmentation network is applied to accurately label the road markings. For road marking registration, a semantic generalized iterative closest point (SG-ICP) algorithm is designed. Linear road markings are modeled as 1-manifolds embedded in 2D space, mitigating the influence of constraints along the linear direction, addressing the under-constrained problem and achieving a lower localization errors on HD maps than ICP. Extensive experiments are conducted in real-world scenarios, demonstrating the effectiveness and robustness of our system.**

***Index Terms*—Localization, autonomous vehicles, road markings, LiDAR, HD map.**

I. Introduction

Accurate localization is a prerequisite for autonomous driving. In unsheltered open-air environments, the global positioning system (GPS) is the predominant technology for accurate localization. However, GPS-provided poses become unstable when satellite signals are obstructed by ceilings or viaducts. Therefore, localization through environmental perception using observation sensors, such as cameras and light detection and ranging (LiDAR) sensors, becomes necessary for autonomous vehicles.

In open environments, such as roads near airport aprons, the absence of visual texture and structural information causes vehicles to rely heavily on road markings for self-localization. Therefore, road marking detection is crucial for autonomous vehicle navigation. [1], [2]. Cameras have been widely used for road marking detection [3], [4], because camera images contain rich texture information of environments. However, cameras are limited by the susceptibility of illumination variations and distortions in bird's-eye view (BEV) lane representation, rendering them less robust for practical applications [5], [6].

In contrast, LiDAR sensors exhibit reduced sensitivity to varying illumination conditions and provide a precise 3D representation of the environment. Meanwhile, road markings can be extracted from road surfaces using LiDAR point clouds, leveraging the characteristic of their high reflectance from the retro-reflective materials [7], [8]. However, LiDAR-based methods face challenges to balance the need for denser point cloud with the essential requirement for real-time performance.

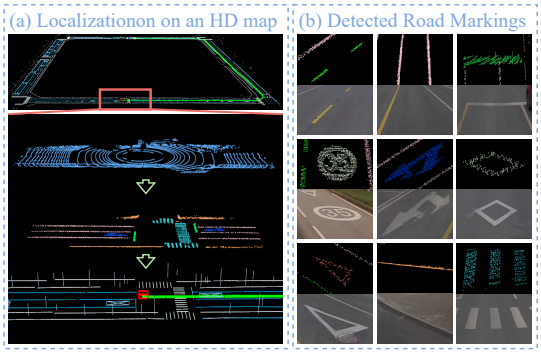


Fig. 1: (a) The HD map localization of our approach is visualized, where the trajectory of vehicle localization is marked in green, and the current pose of the vehicle is represented by a red cube. The blue point cloud represents ground points from a single-frame LiDAR data. These ground points are adaptively segmented to identify highly reflective points. Subsequently, they are aggregated by successive frames of data to form a denser point cloud. Finally, semantic segmentation is applied to obtain a semantic point cloud, which is then registered with the HD map to estimate the vehicle's pose. (b) Road markings extracted using our approach are visualized, encompassing dashed lanes, solid lanes, stop lines, texts, arrows, diamond signs, triangle signs, curbs, and crosswalks.

To address the challenges, a real-time LiDAR-based approach is proposed for road marking detection and registration with high-definition (HD) maps, as visualized in Fig. 1 (a). For road marking detection, an adaptive segmentation technique is first employed to efficiently isolate points correlated with road markings. Then, a spatio-temporal probabilistic local map is established by aggregating segmented points from historical scans, resulting in a dense point cloud representation of road markings. Finally, a LiDAR bird's-eye view (LiBEV) image is generated by partitioning the local map into grid cells, and a proficiently trained instance segmentation network (CenterMask [9] is selected in our implementation) is applied to accurately detect 9 different types of road markings, as shown in Fig. 1 (b).

Regarding the registration of road markings, the existing scan matching methods, such as ICP or NDT and their variants, do not fully utilize semantic information and the distribution of ground markings, which results in significant registration errors and can even lead to degeneration issues.

To address the problem, a semantic generalized iterative closest point (SG-ICP) algorithm is specifically designed to robustly align the detected road markings with the HD map by leveraging both their semantic and geometric attributes. In the proposed SG-ICP registration, the linear types of road markings are modeled as 1-manifolds embedded in the 2D space, making the constraints along the linear direction have minimal influence on the ultimate solution.

The contributions of this paper are summarized as follows.

1) A LiDAR-based road marking detection approach is proposed for online environmental perception, in which point density and real-time performance are balanced by adaptively segmenting high-reflectance points and updating spatio-temporal probabilistic local map. Finally, a LiBEV image is generated, and 9 different types of road markings can be detected accurately using an instance segmentation network on the LiBEV image.
2) A novel road marking registration algorithm is proposed for localization of autonomous vehicles on HD maps, in which linear road markings are represented as 1-manifolds embedded in 2D space. This representation can provide a robust and accurate solution for the registration problem with minimal influence on the under-constrained dimensions. Compared with the widely-used ICP, SG-ICP achieves higher accuracy of localization.
3) Comprehensive experiments are conducted in real-world scenarios, demonstrating effectiveness and real-time performance of our system. Furthermore, experimental results indicate the approach's adaptability to various types of LiDAR sensors, as well as its robustness under different vehicle speeds and weather conditions.

II. Related work

In urban autonomous driving scenarios, the detection of road markings stands out as a crucial method for environmental perception. The road markings, typically painted on asphalt roads using retro-reflective materials, play a vital role in guiding autonomous vehicles. Leveraging the near-infrared wavelength of laser pulses, road markings exhibit higher reflectance compared to unpainted road surfaces [7]. As a result, the LiDAR sensor's ability to capture intensity measurements becomes instrumental in detecting these road markings [8].

LiDAR-based road marking detection is extensively applied in HD map generation [10], [11]. Since the data for HD map generation is processed offline, consecutive scans are aggregated into a point cloud with a high density of points, capturing detailed information about the surroundings [10]. However, processing such high-density points is time-consuming, rendering existing methods applied in HD map generation impractical for the online environmental perception.

In existing studies on real-time perception, the detection of road markings is achieved by thresholding the measured intensities within a single LiDAR scan. A lane markings detection approach was developed by Team AnnieWAY for the DARPA Urban Challenge 2007, which detected the painted lane markings from the single scans by thresholding the points with high-reflectance gradients [12]. Similarly, the approach proposed in [13] detected highly reflective lane markings by employing a polar lane detector grid. In [14], a modified Otsu thresholding technique was employed to segment high-reflectance points obtained from a multilayer LiDAR into distinct categories. Due to the sparsity of LiDAR measurements, the single-scan-based approaches face challenges detecting complete road markings, making the detection results susceptible to noise and lack robustness.

The approach proposed in [15] accumulated two consecutive frames of segmented road points, and then applied a fixed intensity threshold to isolate lane marking points. In the subsequent works [16], [17], the approach was extended to detect various types of high-reflectance landmarks, such as road signs and guard rail reflectors, to reduce localization errors. However, these methods utilize a fixed intensity threshold to segment road marking points, which is sensitive to changes in environmental conditions and sensor types.

Recently, the deep learning approaches have been widely-used in the road marking detection, such as [5], [18] and [19]. The global feature correlation (LLDN-GFC) was introduced in [20] which leveraged the spatial characteristics of lane lines including sparsity, thinness, and elongation across the entirety of the ground points. This method was further improved in [21], resulting in a substantial reduction in computational cost. Nevertheless, LLDN-GFC focuses solely on extracting lane lines, overlooking other types of road markings. This limitation implies that the extracted lane lines can only provide lateral constraints on the vehicle's poses, potentially contributing to a degeneracy problem during the localization.

III. Methodology

In response to the limitations in previous researches, we propose a LiDAR-based road marking detection system for real-time environmental perception. Additionally, a novel road marking registration algorithm is introduced to reduce localization errors of autonomous vehicles with HD maps. The flowchart of the proposed system is illustrated in Fig. 2.

A. LiDAR-based Real-Time Road Marking Detection

Limited by the sparse distribution of LiDAR points, the robust detection of road markings proves challenging when relying on individual frame of data. To overcome this limitation, successive LiDAR scans are aggregated into a local map, generating a denser point cloud for effective road marking detection. In consideration of online requirements and high-reflectance road markings, the aggregation process can selectively extract points with higher intensities from the ground plane. This approach ensures the construction of a local map optimized for road marking detection, striking a balance between computational efficiency and information richness.

1) High-Reflectance Point Segmentation: This procedure aims to adaptively identify points with high reflectance, which are often correlated with road markings painted using retro-reflective materials. To ensure adaptability across diverse sensors and scenarios, we introduce an adaptive segmentation

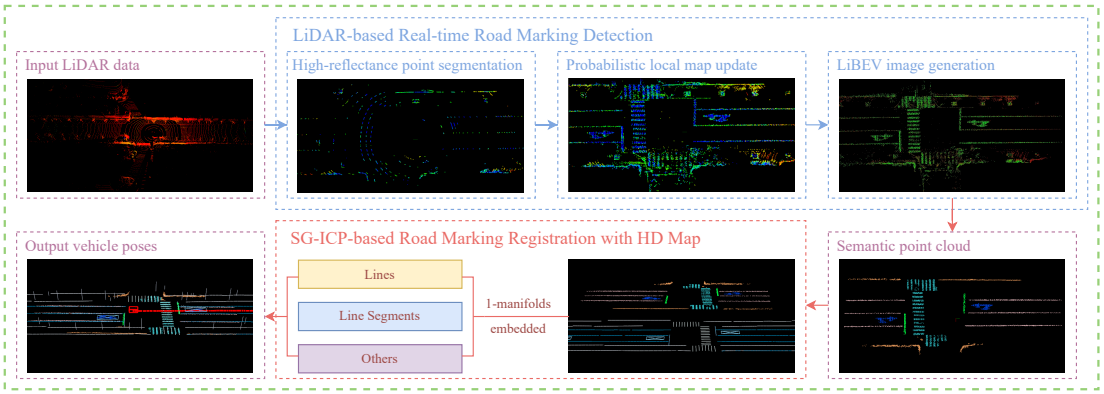


Fig. 2: The flowchart of the proposed approach.

approach designed to isolate high-reflectance points. This enhancement contributes to a more robust system overall.

For the efficiency of the system, only ground points are considered, which are extracted from the LiDAR scan utilizing the methodology in [22]. This approach segments ground points based on height information and subsequently extracts them by partially fitting the ground plane. Then, a coefficient ρ_k is introduced to distinguish high-reflectance points from the ground points in the k-th scan. Specifically, points with intensities below ρ_k are excluded from the scan. Notably, the segmentation coefficient ρ_k is not predetermined manually. Instead, it is dynamically estimated and continuously updated using a Kalman filter. The state of the Kalman filter is evolved according to the state-transition model and measurement model denoted by (1) and (2), respectively.

$$\rho_k = \rho_{k-1} + w_k \tag{1}$$

$$z_k = \rho_k + v_k \tag{2}$$

where the $w_k \sim \mathcal{N}(0, Q_k)$ and $v_k \sim \mathcal{N}(0, R_k)$ are the process noise and measurement noise, respectively. In each LiDAR scan, the mean μ_k and variance σ_k of the intensities of the ground points are calculated. The measurement for the innovation computation is then determined as $z_k = \mu_k + 2\sigma_k$.

This adaptive approach relies on two assumptions. Firstly, it presumes that nearby consecutive roads should possess similar segmentation coefficients owing to the consistency in ground materials. Secondly, it assumes that the majority of LiDAR points lie on the common asphalt surface, while road marking points exhibit statistically higher intensities. These two assumptions are satisfied in most urban and airport road environments, ensuring the effectiveness of the approach. Furthermore, segmenting these high-reflectance points is pivotal for optimizing the efficiency, strategically mitigating the computational load by excluding a significant volume of data points unrelated to road markings.

2) Probabilistic Local Map Update: A local map is constructed through the aggregation of spatio-temporally successive LiDAR scans using an odometry, incorporating high-reflectance points to generate a dense point cloud for road marking detection. However, with the accumulation of scan data, the volume of information grows substantially, leading to an increasing computational burden.

To achieve real-time performance, a novel approach for probabilistically updating the local map is introduced. This approach employs a probabilistic discarding strategy, wherein each point in the map is selectively removed based on a calculated probability value. The probability assigned to the i-th point in the local map, denoted as p_i, is computed by

$$p_i = \frac{1}{1 + (|k - k_i| / \eta)^2}, \tag{3}$$

where k denotes the index of the current frame, and k_i represents the frame from which the i-th point originates. η is a manually-set parameter to determine the probability of discarding old points. As η increases, old points are more likely to be retained, resulting in a higher density of points in the probabilistic local map.

As indicated by (3), higher retaining probability values are assigned to newly observed points by the LiDAR sensor. This strategy effectively ensures the spatio-temporal consistency of the local map, alleviates the impact of accumulated errors over time. Moreover, when contrasted with the aggregation method employing scans within a fixed window, the proposed approach ensures a more seamless transition in the local map data, thereby yielding higher-quality LiBEV images. Finally, points that are too distant from the vehicle are excluded from the local map. This is achieved by maintaining a rectangular region centered on the vehicle, extending w meters laterally and l meters longitudinally. This approach further reduces the number of points in the local map, thereby improving the efficiency of subsequent processing.

3) LiBEV Image Generation: The generation of the LiBEV image involves dividing the local map into grid cells on the ground plane, where each cell corresponds to a pixel in the LiBEV image. Within each cell, the RGB value of the corresponding pixel is determined by mapping the maximum intensity value among the enclosed points using a color map.

In our implementation, a instance segmentation network, CenterMask [9], is used to accurately identify semantic road markings in the LiBEV images which is selected for its balance of accuracy, efficiency, and ease of deployment on em-

bedded automotive platforms. Notably, the proposed approach is not dependent on the specific implementation of CenterMask, any instance segmentation network can be effectively integrated into our pipeline.

Subsequently, points located within the grid cells corresponding to the segmented pixels are extracted from the local map. The extraction yields a semantic point cloud wherein each point is labeled with a specific road marking category. Notably, our approach is designed to accommodate the segmentation of up to 9 types of road markings, including dashed lanes, solid lanes, stop lines, texts, arrows, diamond signs, triangle sighs, curbs and crosswalks, as shown in Fig. 1 (b). The incorporation of diverse semantic road markings, in contrast to approaches solely focused on lane lines, significantly enhances the robustness of pose estimation. In addition, because annotating image semantic segmentation is more convenient than annotating point clouds, the proposed LiBEV image generation approach is better suited for practical applications than directly processing point cloud data.

B. SG-ICP-based Road Marking Registration with HD Map

After road marking detection, the detected road markings can be associated with their corresponding elements in the HD map. Finally, road marking registration is employed to estimate the pose of the vehicle in 2D space. In this subsection, the SG-ICP algorithm is introduced for robust registration of detected road markings with semantic elements in the HD map. In our proposed SG-ICP, detected road markings are divided three categories, including lines, line segments and others. Solid lanes and curbs exhibit a linear distribution in their point clouds and lack distinct endpoints, and thus they are classified as lines. Dashed lanes, sidewalks, and stop lines also have a linear distribution but possess endpoints, leading to their classification as line segments. Texts, arrows, diamond signs and triangle signs do not have linear point cloud distribution, and thus they are classified as others.

For lines, the lack of endpoints leads to the complete loss of constraints along the linear direction. For line segments, endpoints can provide constraints along the linear direction. However, due to inaccurate endpoint estimation, registration between endpoints leads to significant localization errors along the direction of the line segment. Consequently, for linear markings, constraints along their linear direction need to have minimal influence on the pose estimation, mitigating the effect of under-constraint issues. In our algorithm, the registration of the three different categories of markings is organized into a unified representation using the objective function of GICP.

The GICP algorithm incorporates a probabilistic model into the optimization procedure, as defined by

$$\begin{aligned}\boldsymbol{T}^* = \arg\min_{\boldsymbol{T}} \Bigg(& \sum_{i=1}^{n} (\boldsymbol{q}_{mi} - \boldsymbol{T} \cdot \boldsymbol{q}_{Li})^T \\ & (\boldsymbol{C}_{mi} + \boldsymbol{R}\boldsymbol{C}_{Li}\boldsymbol{R})^{-1} (\boldsymbol{q}_{mi} - \boldsymbol{T} \cdot \boldsymbol{q}_{Li}) \Bigg),\end{aligned} \tag{4}$$

where $\boldsymbol{q}_{mi}$ and $\boldsymbol{q}_{Li}$ represent a pair of corresponding points, belonging respectively to the HD map element and the labeled point cloud. Their correspondences are established through the nearest neighbor search strategy. $\boldsymbol{C}_{mi}$ and $\boldsymbol{C}_{Li}$ represent the covariance matrices of points from map and labeled point cloud, respectively, which are appropriately constructed in SG-ICP to mitigate the influence of under-constrained direction.

In our SG-ICP, the probabilistic model is designed by exploiting the semantic and geometric attributes inherent in semantic road markings. For the points lying on the i-th detected road marking instance, the covariance matrix is estimated by

$$\tilde{\boldsymbol{C}}_{Li} = \frac{1}{n_i - 1} \sum_{j}^{n_i} (\boldsymbol{p}_{L(i,j)} - \tilde{\boldsymbol{p}}_{Li}) \cdot (\boldsymbol{p}_{L(i,j)} - \tilde{\boldsymbol{p}}_{Li})^T, \tag{5}$$

where $\boldsymbol{p}_{L(i,j)}$ represents the j-th point of the i-th road marking instance, $\tilde{\boldsymbol{p}}_{Li}$ represents the centroid of the points. Then, the singular value decomposition (SVD) is performed on $\boldsymbol{C}_{Li}$.

$$\tilde{\boldsymbol{C}}_{Li} = \boldsymbol{U}_i \tilde{\boldsymbol{\Sigma}}_i \boldsymbol{V}_i, \quad \tilde{\boldsymbol{\Sigma}}_i = \mathrm{diag}(\sigma_1^2, \sigma_2^2), \tag{6}$$

σ_1 and σ_2 satisfy $\sigma_1 \geq \sigma_2$. Then, a matrix $\boldsymbol{\Sigma}_i = \mathrm{diag}(1, \epsilon)$ is constructed, with ϵ satisfying

$$\epsilon = \begin{cases} 1e-6, & \text{if the marking is classified lines;} \\ 1e-1, & \text{if the marking is classified line segments;} \\ 1, & \text{if the marking is classified others.} \end{cases} \tag{7}$$

The three categories of road markings have distinct values of ϵ, representing the different constraints along the line direction. A value of ϵ closer to 1.0 indicates a stronger constraint along the line direction. The final covariance matrix corresponding to the i-th road marking can be calculated by

$$\boldsymbol{C}_{Li} = \boldsymbol{U}_i \boldsymbol{\Sigma}_i \boldsymbol{V}_i. \tag{8}$$

The i-th semantic element in the HD map is represented as $\{\boldsymbol{v}_{mi}, l_{mi}, P_{mi}\}$, where $\boldsymbol{v}_{mi}$, l_{mi} and $P_{mi} = \{\boldsymbol{p}_{m(i,j)}, j = 1, 2, \cdots, n_{mi}\}$ denote the main direction, the semantic label and the point set of the map element, respectively. The rotation that rotates the basis vector $\boldsymbol{e}_1 = [1, 0]^T$ to the direction $\boldsymbol{v}_{mi}$ can be calculated by

$$\boldsymbol{R}_{vi} = \cos\theta \cdot \boldsymbol{I} + (1 - \cos\theta)\boldsymbol{r}\boldsymbol{r}^T + \sin\theta \cdot [\boldsymbol{r}]_\times, \tag{9}$$

$$\boldsymbol{r} = [\boldsymbol{e}_1]_\times \cdot \boldsymbol{v}_{mi}, \quad \theta = \arccos(\boldsymbol{e}_1^T \boldsymbol{v}_{mi}). \tag{10}$$

The symbol $[\boldsymbol{r}]_\times$ denote the skew-symmetric matrix associated with the vector $\boldsymbol{r}$. The covariance matrix corresponding to the i-th semantic element is calculated by

$$\boldsymbol{C}_{mi} = \boldsymbol{R}_{vi} \boldsymbol{\Sigma}_i \boldsymbol{R}_{vi}. \tag{11}$$

Finally, associations can be established between the semantic point cloud and the closest points of the map elements shared the same semantic label. Meanwhile, their corresponding covariance matrices calculated in (8) and (11) are then substituted into the objective function (4) to initiate the optimization and iteration process. The probabilistic model from SG-ICP characterizes both the semantic and geometric

attributes for road marking registration, which improves the accuracy of pose estimation.

IV. Experimental Evaluation

In this section, extensive experiments are conducted using data collected from diverse scenarios and vehicular platforms, demonstrating the accuracy and robustness of the proposed approach across different scenes and types of LiDAR sensors.

A. Experimental Setup

All experiments are conducted on the NVIDIA Jetson AGX Xavier. The acquisition frequency of LiDAR data is set to $10Hz$. The global localization results of vehicles are recorded using Real-Time Kinematic (RTK) and temporally synchronized with the LiDAR data. These RTK results are used as ground truths. The experimental scenarios and the corresponding HD maps are shown in Fig. 3. *Fangshan1* and *Fangshan2* represent two open urban scenarios in Beijing Fangshan, which covers a $0.30km \times 0.25km$ area and spans a length of $2.0km$, respectively. *Jiashan* depicts an internal road measuring $0.20km$ located in a test field in Zhejiang Jiashan. *Airport* represents an internal road spanning a length of $4.0km$ located within an airport. For the parameters of our approach, the initial variances of the state-transition model and the measurement model were experimentally set to 0.1 and 2.0 in the Kalman filter, respectively. η to determine the discarding probability of local map points was set empirically to 50.0. The length l and width w of the rectangular region for the local map were set to 60 m and 30 m, respectively. In the subsequent experiments, using the same parameters across all 8 sequences could yield accurate results, demonstrating the robustness of the algorithm to parameter selection.

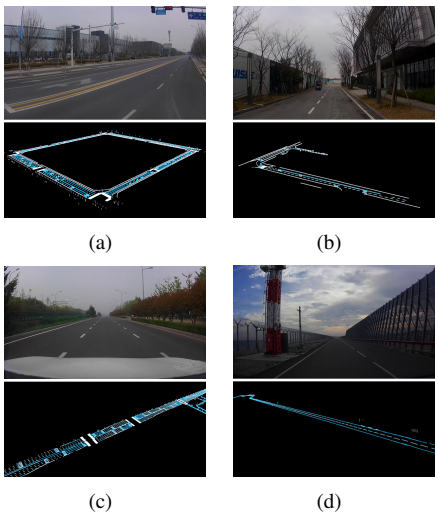


Fig. 3: The experimental scenarios (top) and their corresponding HD maps (bottom). (a) *Fangshan1* (b) *Jiashan* (c) *Fangshan2* (d) *Airport*.

B. Evaluation on Road Marking Detection

In this subsection, an experiment is conducted to assess the performance of our road marking detection approach using precision-recall metrics. To ensure a comprehensive evaluation, a total of 9949 labeled LiBEV images of urban scenes are available, with 80% used for training and the remaining 20% used for testing. The manual annotations serve as the ground truths against which we evaluate the precision and recall of our approach in detecting road markings. A true positive sample is identified when the Intersection over Union (IoU) between the detected instance and its corresponding annotated instance exceeds 0.5, and both instances shared the same semantic label. Conversely, a false positive sample represent a detection result for which no corresponding instance with the same semantic label and an IoU greater than 0.5 could be found in the ground truths. Meanwhile, a false negative sample indicates that an instance present in the ground truth is not successfully detected by our approach.

The precision, recall and F1-score for all types of road markings supported by the approach are presented in TABLE I. The proposed approach successfully detects 9 distinct types of road markings, assigning semantic labels to each point in the LiDAR data, as visually depicted in Fig. 1 (b). The experimental results demonstrate the effectiveness of our approach in successfully detecting common road elements, achieving high precision and recall rates. Notably, certain elements such as curbs and crosswalks exhibit a slight decrease in precision, attributed to their visual similarity to lane markings in LiBEV images. However, the subsequent HD map registration steps effectively mitigate the impact of these false positives on localization. Moreover, the proposed detection approach is highly efficient, meeting real-time perception requirements for vehicles, which is detailed in Section IV-D.

C. Evaluation on Localization

The registration performance of the proposed SG-ICP is evaluated using absolute translation error (ATE) and absolute rotation error (ARE), calculated using the mean absolute errors (MAE) approach. The ground truth for comparison is obtained from RTK data. The translation error is decomposed into longitudinal and lateral components, abbreviated as Lo. ATE and La. ATE, respectively, representing the translation error along the vehicle's orientation and the direction perpendicular to the vehicle's orientation. This is because the shape of road markings are usually strongly correlated with the road direction, resulting in significant differences in translation errors in the two directions. As presented in TABLE II, the evaluation encompasses eight experimental sequences, spanning four scenarios and employing seven different LiDAR configurations, demonstrating the flexibility of the proposed approach. The widely-used ICP algorithm is chosen as the baseline for evaluation. Notably, SG-ICP is replaced with ICP only in the HD map registration step, while all other steps and parameters remained unchanged to ensure a fair comparison. Additionally, the proportions of ATE $<$ 0.3 m and ARE $<$ 1.0° are also reported in Table II, which typically represent the ratio of achieving good localization results.

As indicated in the table, SG-ICP outperforms ICP in most sequences. SG-ICP has a clear superiority in terms of La.

TABLE I: The precision, recall and F1-score for all types of road marking supported by our approach.

	Dashed lane	Solid lane	Stop line	Text	Arrow	Diamond sign	Triangle sign	Curb	Crosswalk	Average
Precision	94.05%	86.11%	81.16%	89.19%	91.70%	94.66%	88.23%	64.52%	72.77%	86.31%
Recall	96.50%	85.48%	75.47%	100.00%	96.57%	97.37%	100.00%	76.11%	71.96%	88.56%
F1-score	95.27%	85.79%	78.21%	94.29%	94.07%	95.99%	93.75%	69.83%	72.36%	85.20%

TABLE II: Average localization error compared to the ground-truth obtained through RTK.

Seq.	Scene	LiDAR type	Lo. ATE (m)		La. ATE (m)		ARE (deg)		ATE <0.3m		ARE <1.0°	
			ICP	SG-ICP	ICP	SG-ICP	ICP	SG-ICP	ICP	SG-ICP	ICP	SG-ICP
S1	*Fangshan1*	2#Hesai-Pandar64	0.158	**0.137**	0.050	**0.043**	0.233	**0.208**	86.7%	**89.8%**	98.4%	**98.9%**
S2	*Fangshan1*	1#Hesai-Pandar64	0.165	**0.160**	0.051	**0.041**	0.386	**0.346**	84.6%	**86.9%**	**93.9%**	93.6%
S3	*Fangshan1*	2#VLP-32C	0.167	**0.130**	0.127	**0.043**	0.302	**0.188**	80.1%	**89.7%**	96.7%	**98.9%**
S4	*Fangshan1*	2#HAP	0.164	**0.139**	0.080	**0.058**	0.299	**0.218**	79.8%	**88.1%**	93.3%	**97.1%**
S5	*Jiashan*	2#VLP-32C&1#VLP-16	0.082	**0.077**	0.055	**0.050**	0.547	**0.401**	**98.5%**	97.8%	87.8%	**96.5%**
S6	*Jiashan*	3#HAP	**0.099**	0.106	0.082	**0.062**	0.330	**0.309**	94.8%	**95.9%**	**96.6%**	95.8%
S7	*Fangshan2*	1#Hesai-Pandar64	**0.124**	0.125	0.091	**0.040**	0.277	**0.184**	88.3%	**91.0%**	97.8%	**99.6%**
S8	*Airport*	2#Hesai-XT16	0.129	**0.128**	0.116	**0.050**	0.447	**0.230**	84.2%	**91.5%**	95.7%	**99.1%**

ATE and ARE, due to the emphasis placed on the sufficiently constrained direction during the SG-ICP calculation process. Fig. 4 depicts the visualized trajectories estimated by SG-ICP-based and ICP-based approaches, respectively, in comparison with the ground-truths acquired through RTK. It is worth noting that the substantial localization error of SG-ICP and ICP are marked with purple and red lines, respectively, where estimated translation errors exceed 2.0 m or rotation error surpass 5.0°. It is evident from Fig. 4 that SG-ICP demonstrates significantly fewer occurrences of substantial localization errors compared to ICP across all sequences.

In conclusion, the proposed approach achieves centimeter-level La. ATEs in a variety of environmental scenarios and with different types of LiDAR sensors. The tested sensors encompass not only traditional mechanical LiDARs like VLP-32C, Hesai-Pandar64, and Hesai-XT16 but also solid-state LiDARs such as HAP. The comprehensive experiments illustrate the robustness and adaptability of the approach across diverse scenarios and sensor types. In addition, it is worth noting, as indicated in TABLE II, that the Lo. ATE is slightly larger than the La. ATE. In the urban road scenario where autonomous driving occurs, the majority of road markings exhibit a linear shape along the longitudinal direction. Consequently, the stronger influence of lateral constraints, compared to longitudinal constraints, contributes to a more precise lateral localization outcome. Nevertheless, our approach ensures that the Lo. ATE remains below 0.20 m, thereby ensuring its effectiveness in autonomous driving applications.

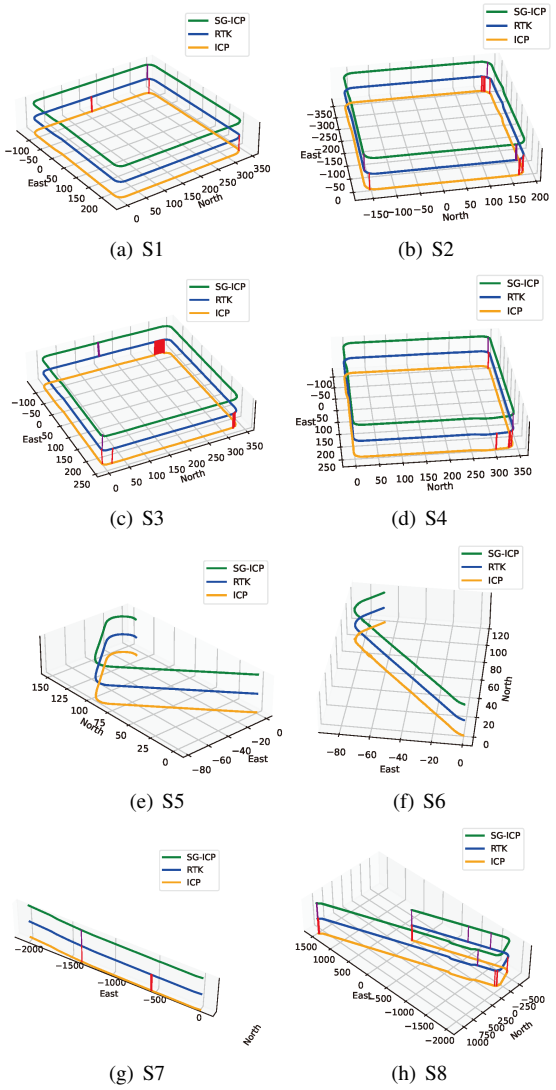
D. Evaluation on Runtime

During the experiments conducted on the eight sequences, the runtime for each sub-step of our approach is detailed in TABLE III. The corresponding box-plot depicting the statistical results can be observed in Fig. 5. It is worth noting that the runtime of the detection sub-step is divided into CPU time and GPU time. CPU time refers to the time consumed by the steps processed by the CPU, including high-reflectance point segmentation, probabilistic local map update, and LiBEV image generation. GPU time refers to the inference time of the instance segmentation of the LiBEV image. The registration sub-step is processed only by the CPU. It can be seen that, when utilizing the onboard processor XAVIER, the average and maximum runtime of the overall approach consistently remains below 50 ms and 200 ms across various scenes and types of LiDAR sensors. Consequently, the efficiency of the proposed system proves sufficient for real-time perception and localization in autonomous vehicle applications.

TABLE III: The runtime (ms) for each sub-step of our approach.

Seq.	Detection (CPU)	Detection (GPU)	Registration	Totals
S1	28.20	16.47	4.63	49.30
S2	20.34	16.36	4.25	40.95
S3	22.13	16.42	4.43	42.98
S4	19.66	18.21	3.48	41.35
S5	22.20	16.55	3.96	42.71
S6	17.46	19.13	4.73	41.32
S7	18.69	14.74	3.46	36.89
S8	19.98	16.75	2.63	39.36

Moreover, it is worth highlighting that the runtime on the *S1* sequence is only 8.35 ms longer than that on the *S2* sequence, despite the fact that the data quantity of *S1* is twice that of *S2* (as indicated in the *LiDAR type* column in TABLE II). This observation demonstrates that the runtime does not exhibit a linear increase with the quantity of the point cloud data, because the substantial reduction in the quantity of the aggregated local map points is achieved through a probabilistic discarding strategy.



(a) S1 (b) S2

(c) S3 (d) S4

(e) S5 (f) S6

(g) S7 (h) S8

Fig. 4: Comparison between the trajectories estimated by SG-ICP and ICP is conducted using ground-truth trajectories provided by RTK. The substantial localization error of SG-ICP and ICP are marked with purple and red lines, respectively, where estimated translation errors exceed 2.0 m or rotation error surpass 5.0°.

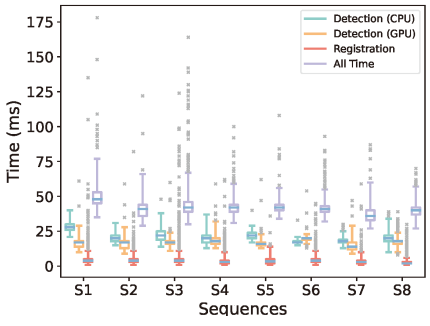


Fig. 5: A box plot illustrating the time consumption for each sub-step of the proposed approach.

E. Evaluation on Robustness

To demonstrate the robustness of our approach, we evaluate the localization errors at different vehicle speeds, as in TABLE IV. The vehicle was driven at speeds of 20 km/h, 40 km/h, and 60 km/h in the *Fangshan1* scenario using 1 Hesai-Pandar64 LiDAR. The results were compared against the ground-truth provided by RTK. As evident from TABLE IV, there is a slight increase in the localization error with higher driving speeds. This can be attributed to the fact that, as the driving speed increases, the point cloud data captured by the LiDAR sensors is more prone to motion distortions. Despite the slight increase in localization error with higher driving speeds, the proposed approach consistently maintains a relatively low level of localization error. This demonstrates the robustness of the approach across varying vehicle speeds. Regarding real-time performance, as indicated in TABLE IV, the system runtime is minimally affected by increases in driving speed. This further highlights the robustness of the system in handling variations in vehicle speed.

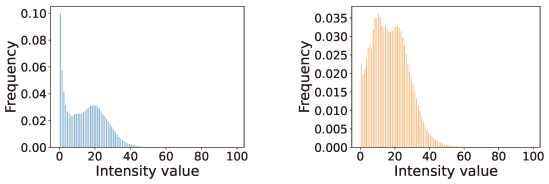
TABLE IV: The localization errors at various driving speeds.

Speeds	Lo. ATE (m)	La. ATE (m)	ARE (deg)	Time (ms)
20 km/h	0.166	0.048	0.366	42.56
40 km/h	0.124	0.067	0.620	44.40
60 km/h	0.153	0.091	0.775	44.92

To illustrate the robustness of our approach under varying weather conditions, experiments were conducted in different settings. As depicted in Fig. 6, the intensity distribution of LiDAR point clouds on dry and wet road surfaces (on sunny and rainy days) typically exhibits significant differences. As a result, rainy weather poses considerable challenges to intensity-based road marking extraction, particularly for methods relying on fixed intensity thresholds. The LiBEV images generated under both sunny and rainy weather conditions are depicted in Fig. 7. It is evident that the proposed adaptive threshold-based approach consistently provides stable and accurate segmentation results, even in the presence of significantly different intensity distributions caused by varying weather conditions. TABLE V presents a comparison of localization errors under both dry and wet ground conditions in the *Fangshan1* scenario, employing 1 Hesai-Pandar64 LiDAR. Although more noise in LiBEV images causes an increase in localization error when driving on wet ground, it can still ensure average La. ATE within 0.10 m and Lo. ATE within 0.20 m. These results demonstrate the robustness of our approach in addressing challenging weather conditions.

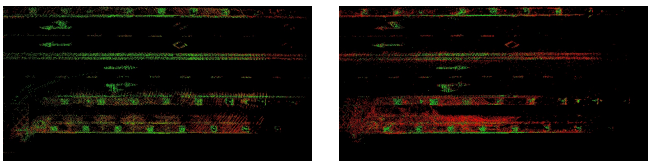
TABLE V: The localization errors at various road conditions.

Speeds	Lo. ATE (m)	La. ATE (m)	ARE (deg)
Dry road	0.160	0.041	0.346
Wet road	0.199	0.056	0.390



(a) (b)

Fig. 6: Histograms illustrating the LiDAR intensity distribution on (a) sunny and (b) rainy days, respectively, in the *Fangshan1* scenario.



(a) (b)

Fig. 7: The LiBEV images on (a) sunny and (b) rainy days in the *Fangshan1* scenario. Pixel coloring is based on intensity values, with green representing high intensity and red representing low intensity.

V. Conclusion

In this paper, we introduce a LiDAR-based online environmental perception and localization system with high efficiency and robustness. The proposed road marking detection approach employs a novel adaptive segmentation technique to enhance efficiency, and utilize a spatio-temporal probabilistic local map to ensure the density of points. For road marking registration, an SG-ICP algorithm is designed, modeling linear road markings as 1-manifolds embedded in 2D space. Our approach minimizes the influence of constraints along the linear direction of markings, to address the under-constrained problem, and thus reduce localization errors. Extensive experiments conducted in real-world urban environments demonstrate the effectiveness and robustness of the proposed system, showcasing its potential for reliable online environmental perception and localization. However, our approach cannot be applied to roads without road markings on the ground surface, due to the lack of high-reflectance points. In future work, we will explore the effective utilization of above-ground information (such as buildings, poles, trees, etc.) retrived from point-cloud maps, to improve the robustness of localization.

References

[1] A. Gupta and A. Choudhary, "A framework for camera-based real-time lane and road surface marking detection and recognition," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 4, pp. 476–485, 2018.

[2] Z. Feng, M. Li, M. Stolz, M. Kunert, and W. Wiesbeck, "Lane detection with a high-resolution automotive radar by introducing a new type of road marking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 7, pp. 2430–2447, 2019.

[3] C. Yan, C. Zheng, C. Gao, W. Yu, Y. Cai, and C. Ma, "Lane information perception network for HD maps," in *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2020, pp. 1–6.

[4] J. Zhou, Y. Guo, Y. Bian, Y. Huang, and B. Li, "Lane information extraction for high definition maps using crowdsourced data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 7, pp. 7780–7790, 2023.

[5] M. Bai, G. Mattyus, N. Homayounfar, S. Wang, S. K. Lakshmikanth, and R. Urtasun, "Deep multi-sensor lane detection," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018, pp. 3102–3109.

[6] C. Wang, H. Zhang, M. Yang, X. Wang, L. Ye, and C. Guo, "Automatic parking based on a bird's eye view vision system," *Advances in Mechanical Engineering*, vol. 6, 2014.

[7] B. Riveiro, H. González-Jorge, J. Martínez-Sánchez, L. Díaz-Vilariño, and P. Arias, "Automatic detection of zebra crossings from mobile LiDAR data," *Optics & Laser Technology*, vol. 70, pp. 63–70, 2015.

[8] M. Soilán, B. Riveiro, J. Martínez-Sánchez, and P. Arias, "Segmentation and classification of road markings using MLS data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 123, pp. 94–103, 2017.

[9] Y. Lee and J. Park, "CenterMask: Real-time anchor-free instance segmentation," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 13 903–13 912.

[10] J. B. Martirena, M. N. Doncel, A. C. Vidal, O. O. Madurga, J. F. Esnal, and M. G. Romay, "Automated annotation of lane markings using LIDAR and odometry," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 3115–3125, 2022.

[11] H. Xiong, T. Zhu, Y. Liu, Y. Pan, S. Wu, and L. Chen, "Road-model-based road boundary extraction for high definition map via LIDAR," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 18 456–18 465, 2022.

[12] S. Kammel and B. Pitzer, "LiDAR-based lane marker detection and mapping," in *IEEE Intelligent Vehicles Symposium*, 2008, pp. 1137–1142.

[13] P. Lindner, E. Richter, G. Wanielik, K. Takagi, and A. Isogai, "Multi-channel lidar processing for lane detection and estimation," in *2009 12th International IEEE Conference on Intelligent Transportation Systems*, 2009, pp. 1–6.

[14] A. Hata and D. Wolf, "Road marking detection using lidar reflective intensity data and its application to vehicle localization," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2014, pp. 584–589.

[15] F. Ghallabi, F. Nashashibi, G. El-Haj-Shhade, and M.-A. Mittet, "Lidar-based lane marking detection for vehicle positioning in an hd map," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 2209–2214.

[16] F. Ghallabi, G. El-Haj-Shhade, M.-A. Mittet, and F. Nashashibi, "Lidar-based road signs detection for vehicle localization in an hd map," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, 2019, pp. 1484–1490.

[17] F. Ghallabi, M.-A. MITTET, G. EL-HAJ-SHHADE, and F. Nashashibi, "Lidar-based high reflective landmarks (hrl)s for vehicle localization in an hd map," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 4412–4418.

[18] T. Qin, Y. Zheng, T. Chen, Y. Chen, and Q. Su, "A light-weight semantic map for visual localization towards autonomous driving," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 11 248–11 254.

[19] T. Suleymanov, M. Gadd, L. Kunze, and P. Newman, "Lidar lateral localisation despite challenging occlusion from traffic," in *IEEE/ION Position, Location and Navigation Symposium*, 2020, pp. 334–341.

[20] D.-H. Paek, S.-H. Kong, and K. T. Wijaya, "K-lane: Lidar lane dataset and benchmark for urban roads and highways," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2022, pp. 4449–4458.

[21] D.-H. Paek, K. T. Wijaya, and S.-H. Kong, "Row-wise lidar lane detection network with lane correlation refinement," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022, pp. 4328–4334.

[22] D. Zermas, I. Izzat, and N. Papanikolopoulos, "Fast segmentation of 3d point clouds: A paradigm on lidar data for autonomous vehicle applications," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 5067–5073.