RALoc: Enhancing Outdoor LiDAR Localization via Rotation Awareness —Supplementary Material—

Yuyang Yang[1,2*] Wen Li[1,2*] Sheng Ao[1,2] Qingshan Xu[3] Shangshu Yu[4] Yu Guo[1,2]
Yin Zhou[5] Siqi Shen[1,2] Cheng Wang[1,2†]
[1]Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University
[2]Key Laboratory of Multimedia Trusted Perception and Efficient Computing,
Ministry of Education of China, Xiamen University
[3]Nanyang Technological University [4]Northeastern University [5]GAC R&D Center

In this supplementary material, we begin by providing a proof of the equivariance of the Gram-Schmidt process (Sec. 1). Next, we offer a detailed description of our proposed dataset (Sec. 2). We then outline the pipeline for point cloud canonicalization (Sec. 3). Following this, we present a comparison with the scanning range of the MulRan dataset and additional visualization results of the Oxford Radar RobotCar dataset (Sec. 4). Finally, we provide more results on the Oxford and BiLiLo datasets (Sec. 5).

1. Proof: The Gram-Schmidt is equivariant

Proof: Let $\{v_1, v_2, \ldots, v_k\}$ be a linearly independent set of vectors in $\mathbb{R}^n$.

Let define the rotated vectors $v_i' = Rv_i$, where R is an orthogonal matrix.

The Gram-Schmidt process on $\{v_1, v_2, \ldots, v_k\}$ gives:

$$\begin{aligned} u_1 &= v_1, \\ u_2 &= v_2 - \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1, \\ &\vdots \\ u_k &= v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, u_i \rangle}{\langle u_i, u_i \rangle} u_i, \end{aligned} \tag{1}$$

The Gram-Schmidt process to the rotated vectors $v_i' = Rv_i$:

$$u_1' = v_1' = Rv_1 = Ru_1, \tag{2}$$

For u_2', we compute:

$$\begin{aligned} u_2' &= v_2' - \frac{\langle v_2', u_1' \rangle}{\langle u_1', u_1' \rangle} u_1' \\ &= Rv_2 - \frac{\langle Rv_2, Rv_1 \rangle}{\langle Rv_1, Rv_1 \rangle} Rv_1 \\ &= R\left(v_2 - \frac{\langle v_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 \right) = Ru_2, \end{aligned} \tag{3}$$

Similarly, for $u_3', \ldots, u_k'$, we have:

$$u_i' = Ru_i, \tag{4}$$

Thus, the Gram-Schmidt process is rotation equivariant.

2. Dataset Details

We evaluate our proposed RALoc for LiDAR localization on two large-scale outdoor benchmark datasets: Oxford RobotCar dataset [1] and our collected BiLiLo dataset.

2.1. Oxford Radar RobotCar

The Oxford dataset is a geospatial dataset collected using sensors mounted on an autonomous-capable Nissan LEAF vehicle [4]. The dataset primarily features point cloud data captured using dual Velodyne HDL-32E LiDAR sensors. Notably, many studies [2, 3, 7–11], focus on utilizing the point cloud data from only the left-side LiDAR sensor. This dataset is widely used for research in localization [5], mapping, and autonomous navigation tasks.

We list the corresponding data split as shown in Tab. 1.

2.2. BiLiLo

The BiLiLo dataset is specifically designed to address the current lack of datasets featuring significant rotations in LiDAR localization benchmarks. We report the training and testing trajectories on BiLiLo, along with their lengths and

*Equal contribution
†Corresponding author

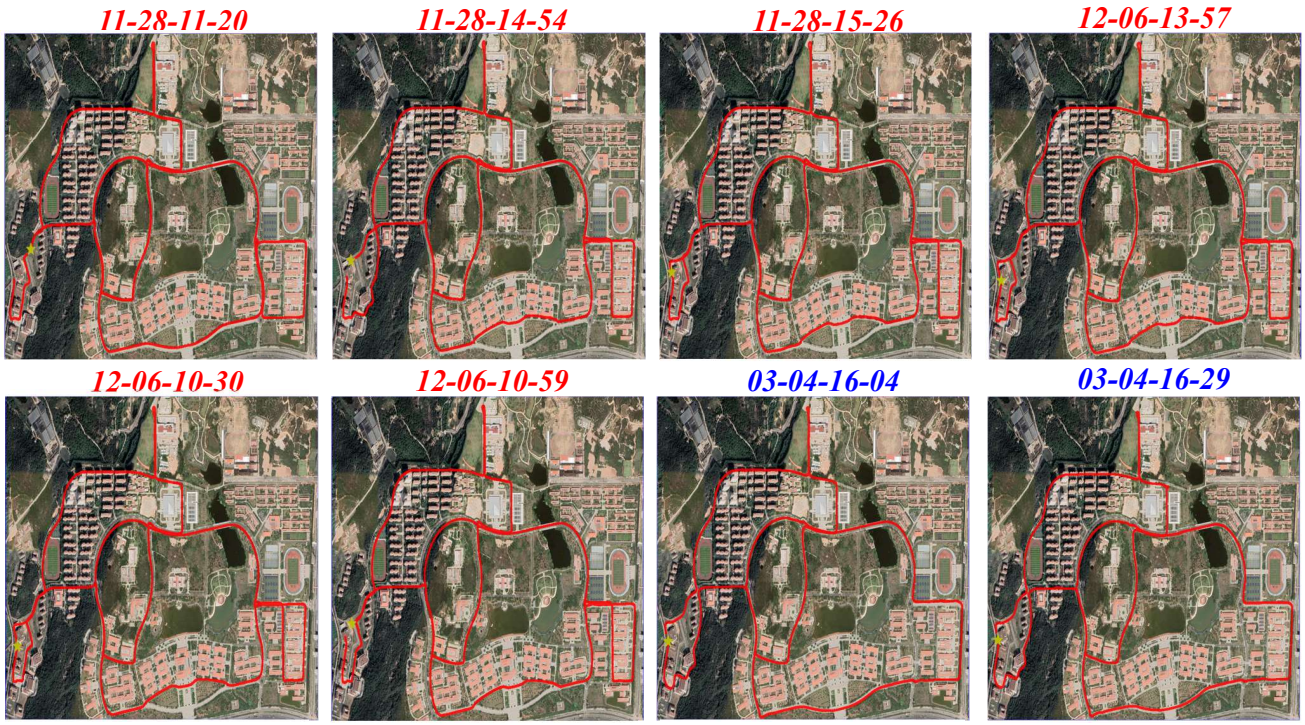

Figure 1. **Visualization dataset traversals.** We are visualizing the trajectory shapes of each sequence on the remote sensing image, with forward direction highlighted in red and the reverse direction shown in blue. The star shape is the starting point.

Sequence	Length	Tag	Training	Test
11-14-02-26	9.37km	sunny	✓	
14-12-05-52	9.22km	overcast	✓	
14-14-48-55	9.04km	overcast	✓	
18-15-20-12	9.04km	overcast	✓	
15-13-06-37	8.85km	overcast		✓
17-13-26-39	9.02km	sunny		✓
17-14-03-00	9.02km	sunny		✓
18-14-14-42	9.04km	overcast		✓

Table 1. Dataset details on the Oxford dataset.

Sequence	Length	Direction	Training	Test
11-28-11-20	9.54km	Forward	✓	
11-28-14-54	9.69km	Forward	✓	
11-28-15-26	9.76km	Forward	✓	
12-06-13-57	9.64km	Forward	✓	
12-06-10-30	9.76km	Forward		✓
12-06-10-59	9.76km	Forward		✓
03-04-16-04	9.04km	**Reverse**		✓
03-04-16-29	9.08km	**Reverse**		✓

Table 2. Dataset details on the BiLiLo dataset. The forward direction highlighted in red and the reverse direction shown in blue.

direction in Tab. 2. The relevant collection parameters, along with an illustration of the sensor layout, are presented in Fig. 2.

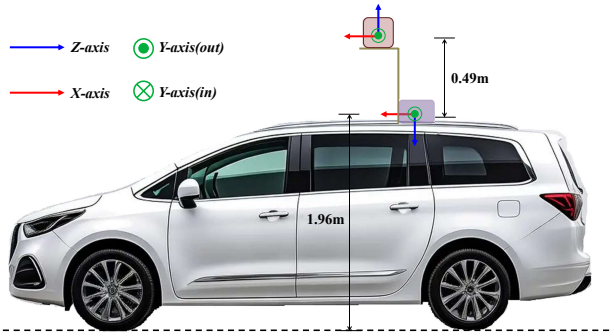

Figure 2. **The BiLiLo platform.** Coordinate frames show the origin and direction of each sensor mounted on the vehicle with the convention: x (red), y (green), z (blue). The dark red area is the LiDAR, and the purple is the M39.

We collect eight trajectories for training and testing. The visualization of each trajectory on the remote sensing map is shown in the Fig. 1.

Data Organization. The BiLiLo dataset is divide into sequences, which include LiDAR point cloud and ground truth poses from a single drive. Sequences are identified by the time at which they are collected see Tab. 2 for details.

Timestamps. The name of each file corresponds to its

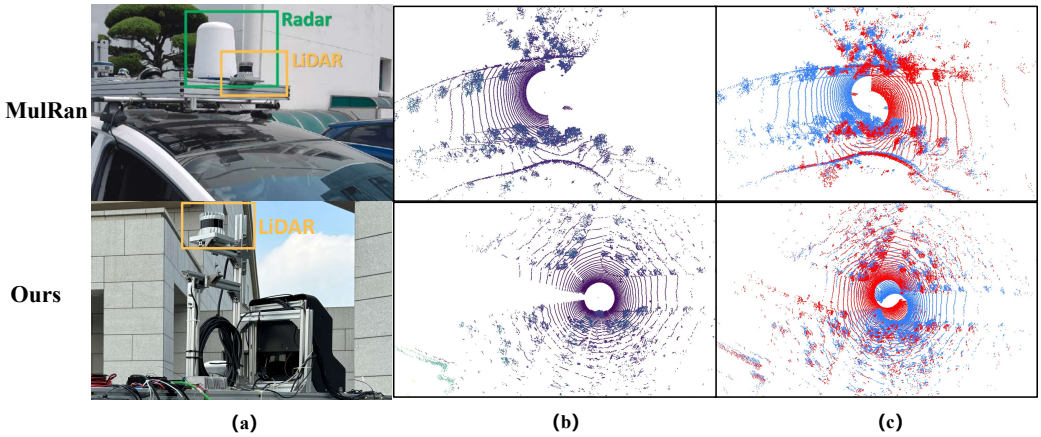

Figure 3. **Visualization comparison of datasets. a.** Data acquisition equipment installation. **b.** Visualization of acquisition LiDAR point cloud. **c.** Scanning comparison of same locations in forward and reverse directions.

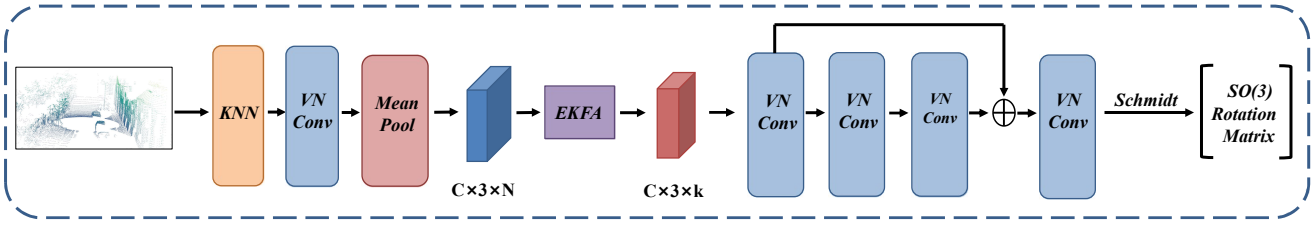

Figure 4. **The Pipline of point cloud canonicalization.** The VN-Conv layer consists of a VN-Linear layer, a VN-BatchNorm layer, and a VN-LeakyReLU activation function. The MeanPool operation applies average pooling along the last dimension.

timestamp, which is given as UNIX epoch times in microseconds. All sensor timestamps are synchronized to the coordinated universal time (UTC) reported by the M39. The Ouster LiDAR is synchronized using a standard hardwired connection to the M39, which carries a pulse-per-second (PPS) signal and NMEA messages. The data-recording computer is synchronized to UTC time using an RS-232 serial cable that carries a PPS signal and NMEA messages.

File Formats. LiDAR point cloud are stored in a binary format to minimize storage requirements. Each point has four fields: $[x,y,z,i]$ where (x,y,z) is the position of the point with respect to LiDAR, i is the intensity of the reflected infrared signal. Each binary file is named according to its corresponding timestamp.

Ground Truth. Ground truth poses are obtained by post-processing GNSS, IMU and an RTK. Positions and velocities are given with respect to a fixed East-North-Down frame $\mathrm{END}_{\mathrm{ref}}$. Each sensor frame's ground truth is stored with the following format: $[t, x, y, z, v_x, v_y, v_z, r, p, y, \omega_z, \omega_y, \omega_x]$ where t is the epoch timestamp in microseconds that matches the filename, $[x, y, z]^T$ is the position of the sensor with respect to $\mathrm{END}_{\mathrm{ref}}$, $[v_x, v_y, v_z]^T$ is the velocity of the sensor with respect to $\mathrm{END}_{\mathrm{ref}}$, (r, p, y) are the roll, pitch, and yaw angles, which can be converted into a rotation matrix between the sensor frame and $\mathrm{END}_{\mathrm{ref}}$. $[\omega_x, \omega_y, \omega_z]^T$ are the angular velocities of the sensor with respect to $\mathrm{END}_{\mathrm{ref}}$ as measured in the sensor frame.

3. Network Architecture

The detailed architecture of the proposed point cloud canonicalization is illustrated in Fig. 4.

4. Visualization

First, we compared our dataset with the MulRan dataset, as shown in Fig. 3. (a) displays our data acquisition equipment; in the MulRan dataset, there is a radar behind the LiDAR that obstructs most of the rear field of view. (b) We visualized the scanned point cloud from the MulRan dataset, revealing a blank area at the rear of the scene point cloud. (c) We overlaid the point clouds from the same location in forward and reverse directions, distinguished by red and blue colors respectively. It can be observed that the regions visible in both directions at the same location in our dataset (areas covered by both red and blue) are significantly greater than those in the MulRan dataset.

We show more visualization results of RALoc on the Oxford dataset in Fig. 5. The first row exhibits the test trajectory without rotation, whereas the second row illustrates the

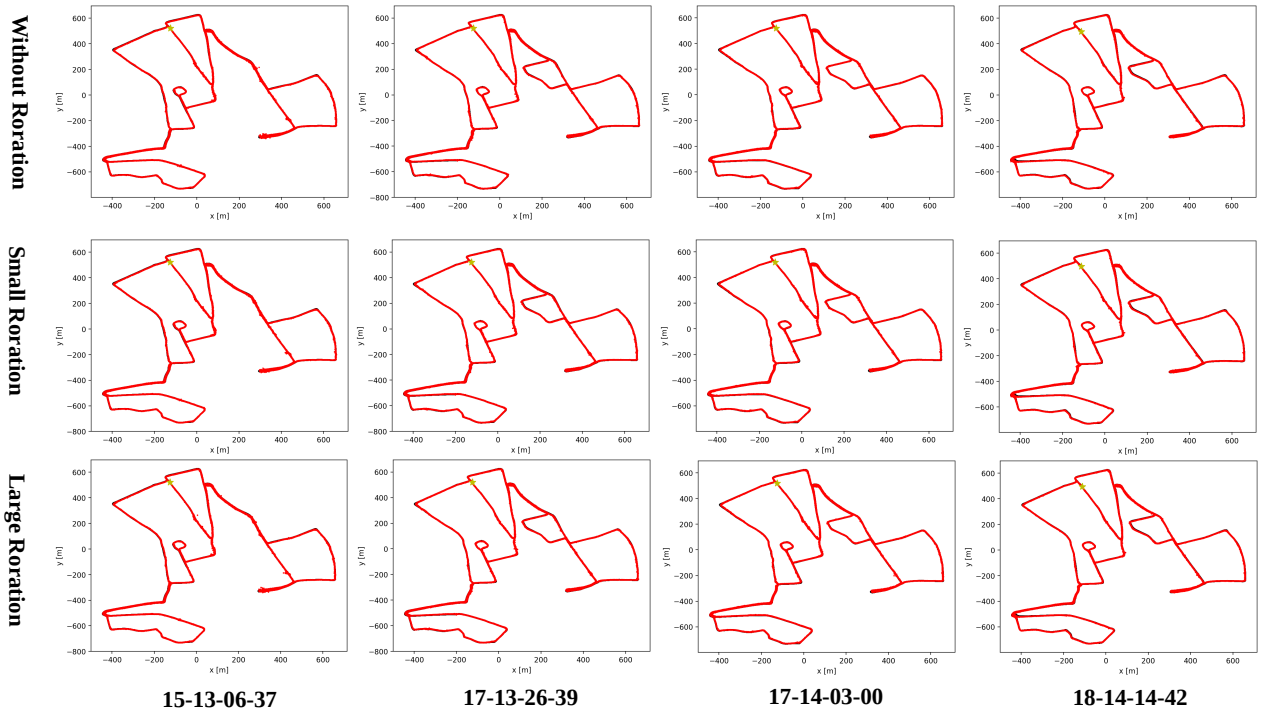

Figure 5. **The evaluation of RALoc on the Oxford dataset.** We visualize the RALoc's results on the Oxford dataset, in line with the main paper, by conducting and visualizing the results under three different rotation scenarios: without rotation, small rotation, and large rotation.

Methods	SO(3) Rotation Augmentation			Without Augmentation			
	STCLoc	NIDALoc	HypLiLoc	STCLoc	NIDALoc	HypLiLoc	**RALoc**
15-13-06-37	36.56m, 23.38°	15.64m, 17.47°	23.07m, 19.43°	43.54m, 13.31°	42.24m, 9.85°	18.69m, 7.81°	**3.17m, 4.08°**
17-13-26-39	38.80m, 21.88°	15.10m, 17.43°	24.34m, 19.87°	40.66m, 12.52°	41.89m, 9.57°	17.81m, 8.34°	**3.87m, 3.96°**
17-14-03-00	38.09m, 21.62°	15.90m, 17.06°	23.96m, 19.49°	39.55m, 12.21°	37.49m, 9.20°	16.16m, 8.13°	**3.32m, 3.83°**
18-14-14-42	32.53m, 20.74°	14.08m, 16.41°	19.50m, 19.58°	35.54m, 12.08°	37.15m, 9.54°	13.46m, 7.67°	**2.58m, 3.70°**
Average	36.50m, 21.91°	15.18m, 17.09°	22.72m, 19.59°	39.82m, 12.53°	39.69m, 9.54°	16.53m, 7.99°	**3.24m, 3.89°**

Table 3. **Results on the Oxford dataset with small rotation.** Mean position error (m) and mean orientation error (°) for various methods are reported, with best results in **bold** and second best results underlined.

Methods	SO(3) Rotation Augmentation			Without Augmentation			
	STCLoc	NIDALoc	HypLiLoc	STCLoc	NIDALoc	HypLiLoc	**RALoc**
15-13-06-37	29.97m, 24.31°	15.34m, 14.39°	28.22m, 29.28°	339.29m, 83.40°	321.16m, 78.81°	44.55m, 90.65°	**3.19m, 4.07°**
17-13-26-39	32.22m, 25.10°	15.08m, 14.90°	28.12m, 29.50°	338.52m, 84.31°	322.29m, 76.83°	43.35m, 89.81°	**3.87m, 3.96°**
17-14-03-00	32.33m, 25.56°	15.19m, 15.30°	27.41m, 29.75°	324.63m, 85.33°	314.47m, 75.21°	39.80m, 90.42°	**3.33m, 3.88°**
18-14-14-42	26.39m, 26.48°	14.35m, 14.10°	22.94m, 29.72°	331.12m, 84.21°	316.01m, 78.25°	38.78m, 89.82°	**2.58m, 3.73°**
Average	30.23m, 25.36°	14.99m, 14.67°	26.67m, 29.56°	333.39m, 84.31°	318.48m, 77.28°	41.62m, 90.18°	**3.24m, 3.91°**

Table 4. **Results on the Oxford dataset with large rotation.** Mean position error (m) and mean orientation error (°) for various methods are reported, with best results in **bold** and second best results underlined.

trajectory under small rotational perturbations (less than 10 degrees around each of the Roll, Pitch, and Yaw, in accordance with the settings described in the main paper). The final row evaluates the performance during large rotations (random rotations around the Yaw axis, as outlined in the main paper). It is evident that the introduction of rotation perception into RALoc results in negligible impact on positioning accuracy. Our model consistently maintains high precision in localization despite varying degrees of rotation.

Methods	SO(3) Rotation Augmentation			Without Augmentation			
	STCLoc	NIDALoc	HypLiLoc	STCLoc	NIDALoc	HypLiLoc	**RALoc**
15-13-06-37	38.02m, 22.78°	15.77m, 17.49°	32.28m, 32.14°	6.93m, 1.48°	5.45m, 1.40°	6.88m, **1.09**°	**3.19m**, 4.10°
17-13-26-39	40.36m, 20.98°	15.20m, 17.72°	31.93m, 32.98°	7.55m, **1.23**°	7.63m, 1.56°	6.79m, 1.29°	**3.87m**, 3.96°
17-14-03-00	39.92m, 20.16°	16.10m, 16.32°	30.57m, 31.01°	7.44m, 1.24°	6.68m, 1.26°	5.82m, **0.97**°	**3.32m**, 3.87°
18-14-14-42	33.88m, 19.50°	14.95m, 16.17°	25.70m, 32.91°	6.13m, 1.15°	4.80m, 1.18°	3.45m, **0.84**°	**2.59m**, 3.71°
Average	38.05m, 20.86°	15.51m, 16.93°	30.12m, 32.26°	7.01m, 1.28°	6.14m, 1.35°	5.74m, **1.05**°	**3.24m**, 3.91°

Table 5. **Results on the Oxford dataset without rotation.** Mean position error (m) and mean orientation error (°) for various methods are reported, with best results in **bold** and second best results underlined.

Methods	SO(3) Rotation Augmentation			Without Augmentation			
	STCLoc	NIDALoc	HypLiLoc	STCLoc	NIDALoc	HypLiLoc	**RALoc**
12-06-10-30	41.13m, 87.44°	18.56m, 94.95°	49.21m, 85.52°	9.27m, 2.47°	8.71m, 2.77°	4.01m, **1.51**°	**1.91m**, 2.90°
12-06-10-59	41.09m, 87.32°	18.70m, 94.81°	48.22m, 81.56°	8.88m, 2.51°	8.43m, 2.74°	3.45m, **1.36**°	**1.89m**, 2.75°
03-04-16-04	116.89m, 52.23°	75.47m, 111.09°	146.86m, 103.87°	426.08m, 87.14°	389.59m, 96.62°	439.12m, 96.87°	**5.24m**, **9.59**°
03-04-16-29	146.05m, 112.30°	103.44m, 113.78°	168.33m, 106.53°	440.37m, 88.79°	394.74m, 98.40°	355.45m, 104.06°	**6.97m**, **11.02**°
Average	86.29m, 84.82°	54.04m, 103.66°	103.16m, 94.37°	221.15m, 45.23°	200.37m, 50.13°	200.51m, 50.95°	**4.00m**, **6.57**°

Table 6. **Results on the BiLiLo dataset.** Mean position error (m) and mean orientation error (°) for various methods are reported, with best results in **bold** and second best results underlined.

5. Additional Validation

To conduct a more comprehensive evaluation, we also tested several additional baselines [6, 9, 11] using the same experimental setup as described in the main text. The test results for the Oxford dataset under rotation conditions are listed in Tab. 3 and Tab. 4. The results for the Oxford dataset are presented in Tab. 5, while the results for the BiLiLo dataset are shown in Tab. 6.

References

[1] Dan Barnes, Matthew Gadd, Paul Murcutt, Paul Newman, and Ingmar Posner. The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. In *ICRA*, pages 6433–6438. IEEE, 2020. 1

[2] Wen Li, Shangshu Yu, Cheng Wang, Guosheng Hu, Siqi Shen, and Chenglu Wen. Sgloc: Scene geometry encoding for outdoor lidar localization. In *CVPR*, pages 9286–9295, 2023. 1

[3] Wen Li, Yuyang Yang, Shangshu Yu, Guosheng Hu, Chenglu Wen, Ming Cheng, and Cheng Wang. Diffloc: Diffusion model for outdoor lidar localization. In *CVPR*, pages 15045–15054, 2024. 1

[4] Will Maddern, Geoffrey Pascoe, Chris Linegar, and Paul Newman. 1 year, 1000 km: The oxford robotcar dataset. *The International Journal of Robotics Research*, 36(1):3–15, 2017. 1

[5] Mikaela Angelina Uy and Gim Hee Lee. Pointnetvlad: Deep point cloud based retrieval for large-scale place recognition. In *CVPR*, pages 4470–4479, 2018. 1

[6] Sijie Wang, Qiyu Kang, Rui She, Wei Wang, Kai Zhao, Yang Song, and Wee Peng Tay. Hypliloc: Towards effective lidar pose regression with hyperbolic fusion. In *CVPR*, pages 5176–5185, 2023. 5

[7] Wei Wang, Bing Wang, Peijun Zhao, Changhao Chen, Ronald Clark, Bo Yang, Andrew Markham, and Niki Trigoni. Pointloc: Deep pose regressor for lidar point cloud localization. *IEEE Sensors Journal*, 22(1):959–968, 2021. 1

[8] Bochun Yang, Zijun Li, Wen Li, Zhipeng Cai, Chenglu Wen, Yu Zang, Matthias Muller, and Cheng Wang. Lisa: Lidar localization with semantic awareness. In *CVPR*, pages 15271–15280, 2024.

[9] Shangshu Yu, Cheng Wang, Yitai Lin, Chenglu Wen, Ming Cheng, and Guosheng Hu. Stcloc: Deep lidar localization with spatio-temporal constraints. *IEEE TITS*, 24(1):489–500, 2022. 5

[10] Shangshu Yu, Cheng Wang, Chenglu Wen, Ming Cheng, Minghao Liu, Zhihong Zhang, and Xin Li. Lidar-based localization using universal encoding and memory-aware regression. *PR*, 128:108685, 2022.

[11] Shangshu Yu, Xiaotian Sun, Wen Li, Chenglu Wen, Yunuo Yang, Bailu Si, Guosheng Hu, and Cheng Wang. Nidaloc: Neurobiologically inspired deep lidar localization. *IEEE TITS*, 2023. 1, 5