ALITA: A Large-scale Place Recognition Dataset for Long-term Autonomy

Peng Yin[1], Shiqi Zhao[2,+], Ruohai Ge[1,+], Ivan Cisneros[1], Ji Zhang[1], Howie Choset[1] and Sebastian Scherer[1]

Abstract

For long-term autonomy, most place recognition methods are mainly evaluated on simplified scenarios or simulated datasets, which cannot provide solid evidence to evaluate the readiness for current Simultaneous Localization and Mapping (SLAM). This paper presents a long-term place recognition dataset for use in mobile localization under large-scale dynamic environments. This dataset includes a campus-scale track and a city-scale track. The campus track focuses on the long-term property and is recorded with a LiDAR device and an omnidirectional camera on 10 trajectories. Each trajectory is repeatedly recorded 8 times under variant illumination conditions. The city track focuses on the large-scale property and is recorded only with the LiDAR device on 120km trajectory, which contains open streets, residential areas, natural terrains, etc. They include 200 hours of raw data of all kinds of scenarios within urban environments. The ground truth position for both tracks is provided on each trajectory, obtained from the Global Position System with an additional General ICP-based point cloud refinement. To simplify the evaluation procedure, we also provide the Python-API with a set of place recognition metrics proposed to quickly load our dataset and evaluate the recognition performance against different methods. This dataset targets finding methods with high place recognition accuracy and robustness and providing real robotic systems with long-term autonomy. We provide both the dataset and tools at https://github.com/MetaSLAM/ALITA.

Introduction

Place recognition or loop closure detection (LCD) is one of the most fundamental tasks in simultaneous localization and mapping (SLAM) and is also a key factor for long-term autonomy. In real-world environments, nevertheless, place recognition has been studied for decades Lowry et al. (2016), reliable long-term and large-scale localization is still an unsolved problem. Recent years, noticeable developments in autonomous driving and last-mile delivery along with the increased demand for long-term, large-scale, and repeated (2LR) localization have been witnessed. Unlike the short-term SLAM tasks, the 2LR localization includes both spatial and temporary differences, and the existing datasets are either too complicated or too simple for evaluation methods. To assess the performance in localization tasks, most new recognition methods must be evaluated on previous exiting datasets, even with unique place feature encoding ability.

With recent developments in computer vision, new learning- and non-learning-based, vision- and LiDAR-based place recognition methods have been proposed to improve the recognition performance under viewpoint and appearance differences. All methods have their pros and cons according to the different environmental conditions. Transitional non-learning based place recognition methods, such FABMAP Nowakowski et al. (2017), CoHoG Zaffar et al. (2020) and CALC Merrill and Huang (2018) for visual inputs, or ScanContext Kim and Kim (2018), M2DP He et al. (2016) and 3DSIFT Mondal et al. (2014) for LiDAR inputs, have been well studied in the recent years but require careful parameter tuning. In contrast, learning-based place recognition methods, such as NetVLAD Arandjelovic et al. (2016), NetVLAD Nowakowski et al. (2017), PointNetVLAD Uy and Lee (2018), OverlapNet Chen et al. (2021) have shown improved place localization performance under complicate 3D/2D environment, such as the well-known *KITTI* and *NCLT* datasets.

But currently, limited datasets can hardly evaluate the accuracy, robustness, and generalization ability of current methods under the (1) large-scale, (2) long-term, and (3) changing perspective environments. Collecting a dataset that could cover the above three properties for real mobile platforms is a complicated task; the collection platform is hard and expensive to design, and the procedure for long-term and large-scale recording requires long-time preparation. Many existing methods are mainly evaluated on specific scenarios, and new researchers can hardly judge

[1] Robotics Institute, Carnegie Mellon University, USA.

[2] University of California San Diego, USA.

[+] Authors Shiqi Zhao and Ruohai Ge contributed equally to this paper.

Corresponding author:
Peng Yin, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA.
Email: pyin2@andrew.cmu.edu

Figure 1. ALITA Dataset(Urban). The *Urban* dataset includes four zones (colored in yellow, green, cyan, and blue) covering the downtown, residential, suburban, and commercial areas of Pittsburgh, Pennsylvania.

Figure 2. ALITA Dataset(Campus). The *Campus* dataset covers the main campus area of Carnegie Mellon University and contains diverse 3D scenes including the buildings, parking lot, corridors, courts, and parks.

Table 1. Comparison of different map merging approaches. Total length is the multiplication of geographical coverage with the number of traversing. Temporal Diversity includes seasonal changes and day-night changes. "-" means not applicable.

Method	Scenarios	Total Length	Temporal Diversity	Viewpoint Diversity	Structural Diversity
Freiburg Steder et al. (2010)	Campus	∼ 0.7	One time	−	⋆
Ford Campus Pandey et al. (2011)	Campus	−	One time	−	⋆
KITTI Geiger et al. (2013)	Urban	∼ 39.2	One time	⋆	⋆⋆
NCLT Carlevaris-Bianco et al. (2016)	Campus	∼ 148.5	Season	⋆	⋆
Oxford RobotCar Maddern et al. (2017)	Urban	∼ 1000	Day-Night	⋆	⋆
MulRan Kim et al. (2020)	Campus & Urban	∼ 123.9	Multi times	⋆ ⋆ ⋆	⋆ ⋆ ⋆
KITTI360 Liao et al. (2021)	Urban	∼ 1000	Day-Night	⋆⋆	⋆⋆
ALITA(Urban)(ours)	Urban	∼ 120	One time	⋆⋆	⋆ ⋆ ⋆
ALITA(Campus)(ours)	Campus	∼ 50	Day-Night	⋆ ⋆ ⋆	⋆

the performance in real applications. Even the most recent place recognition approaches, LPDNet Liu et al. (2019) and OverlapNet Chen et al. (2021) are mainly claimed state-of-the-art performance over limited test datasets, and few methods have a reasonable number of real-world evaluations.

This paper presents ALITA, a dataset set for long-term place recognition in large-scale environments. Our datasets contain two tracks: (1) *Urban* dataset, which records LiDAR data inputs in a city-scale urban-like area for 50 segments and $120km$ trajectory in total. And (2) *Campus* dataset, recorded under a campus-scale environment, where we gathered the omnidirectional visual inputs and LiDAR inputs on 10 different trajectories for 8 repeated times, under different illuminations and viewpoints; this dataset targets long-term localization challenge. Figure. 1 and Figure. 2 gives a better visualization of its scale, and Table. 1 shows the comparison of different datasets. Most datasets are targeted at short-term, fixed conditions or viewpoints place recognition tasks, so it is hard to evaluate the localization performance in real-world long-term, large-scale applications. Compared to existing

datasets, our *Urban* dataset covers variant 3D scenarios for comprehensive 3D place recognition evaluation and multi-session SLAM Van Opdenbosch and Steinbach (2019); Tian et al. (2022). And our *Campus* dataset repeatedly covers diverse campus areas with dynamic objects, illumination, and viewpoint differences, which is suitable to evaluate long-term re-localization or incremental learning ability.

Both *Urban* and *Campus* datasets provide the ground truth for the exact place recognition to help evaluate different methods. The *Urban* dataset has been used in the IEEE ICRA 2022 *General Place Recognition Competition* to benchmark the current new 3D place recognition approaches. This paper describes the details of both datasets and provides the Python-API for the whole pipeline of data processing and localization evaluation in https://github.com/MetaSLAM/ALITA.

Related works

Place Recognition

Based on the natural robustness of the LiDAR against illumination variant, loads of LiDAR-based place recognition methods have been developed. The success of PointNet Qi et al. (2017) makes it possible to extract the features from the point cloud directly and, therefore Uy and Lee (2018) convert the local features utilizing PointNet into a global descriptor via a NetVLAD layer Arandjelovic et al. (2016). However, point-based methods suffer from the fixed point number of input which cannot provide many structural details, and giant model size, which leads to low computation efficiency and high computation cost. Komorowski (2021) use sparse 3D convolutions on a voxelized point cloud to extract local features. The introduction of sparse 3d convolution accelerates the process of local feature extraction and extracts the features within the points in each local neighborhood. Nevertheless, most point-based and voxel-based methods are sensitive to viewpoint differences common in real-life robot navigation. Some projection-based methods Kim and Kim (2018), Chen et al. (2021), Yin et al. (2020), Ma et al. (2022) claim viewpoint invariant. Our dataset aims to offer good test cases for models to test their robustness under different viewpoints and translation differences. In the Urban dataset, we provide python API to generate query and database frames based on the user's requirements. The Campus dataset provides forward and reversed loop closures with dynamic object disturbance.

Existing Datasets

Table. 1 summaries a set of outdoor place recognition datasets. *Freiburg* and *Ford Campus* are both collected in campus environments, but the insufficient size makes it hard to train large networks. *KITTI* Geiger et al. (2013) is recorded by a data collection device, including a 64-beam LiDAR (Velodyne HDL-64E) mounted on the car in Karlsruhe. There are slight viewpoint changes, and most of the revisit is in the same direction. *KITTI360* Liao et al. (2021) contains more reverse revisit, but temporal diversity is also not considered. *NCLT* Carlevaris-Bianco et al. (2016) contains times changes in the Campus for 15 months, which includes seasonal changes and repeated sequences that can improve

the model's robustness during training. *Oxford* Maddern et al. (2017) is collected by a car-mounted device and covers the same $10km$ route twice every week for a year. However, the scene these two datasets contain is limited to a single place, which lacks geographical diversity. Our proposed dataset contains around $170km$ long trajectories split into two sections Urban and Campus. *Urban* dataset contains $120km$ of trajectories which provide abundant frames and structural diversity(building, forest, etc.). The $50km$ *Campus* dataset provides deliberate revisit in both directions along with temporal and illumination diversity (day and night)

Place recognition methods can not only contribute to the SLAM system to alleviate localization shift but be essential to large map merging systems. Yin et al. (2022) has proven the application of place recognition in offline and real-time map merging. *MulRan* Kim et al. (2020) contains various types of revisit, such as reverse revisit and lane-level revisit. In addition, the dataset is recorded in different environments. Even though loads of datasets emphasize deliberately designed revisit, they omit the revisit between different trajectories, and the spatial scale is limited to relatively small areas. In ALITA, we provide trajectories with overlaps with adjacent ones both in *Urban* and *Campus*, which makes it possible for methods to evaluate either city-scale or campus-scale environments.

The Platform

Our data collection platform contains a Velodyne VLP-16 LiDAR scanner, Xsens MTI-300 inertial measurement units, and an Nvidia Jetson TX2 onboard computer. For the extrinsic calibration between LiDAR and inertial measurement units, we follow the method mentioned in https://github.com/ethz-asl/lidar_align. For the *Urban* dataset, we mount our platform onto the top of a mobile vehicle and parallel with the GNSS position system to record the ground truth positions in the city-scale environments. As shown in Fig. 3, for the *Campus* dataset, we mount the same platform onto the mobile rover robot, and with an additional THETA V omnidirectional camera on the top of the LiDAR device; this setup can provide time-synced LiDAR inputs and 360 visual inputs.
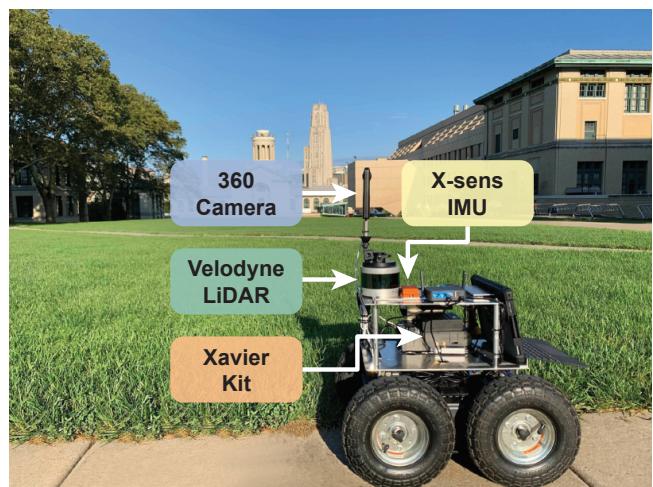

Figure 3. Data-collection platform

a) Time and Viewpoint Change in Campus Dataset

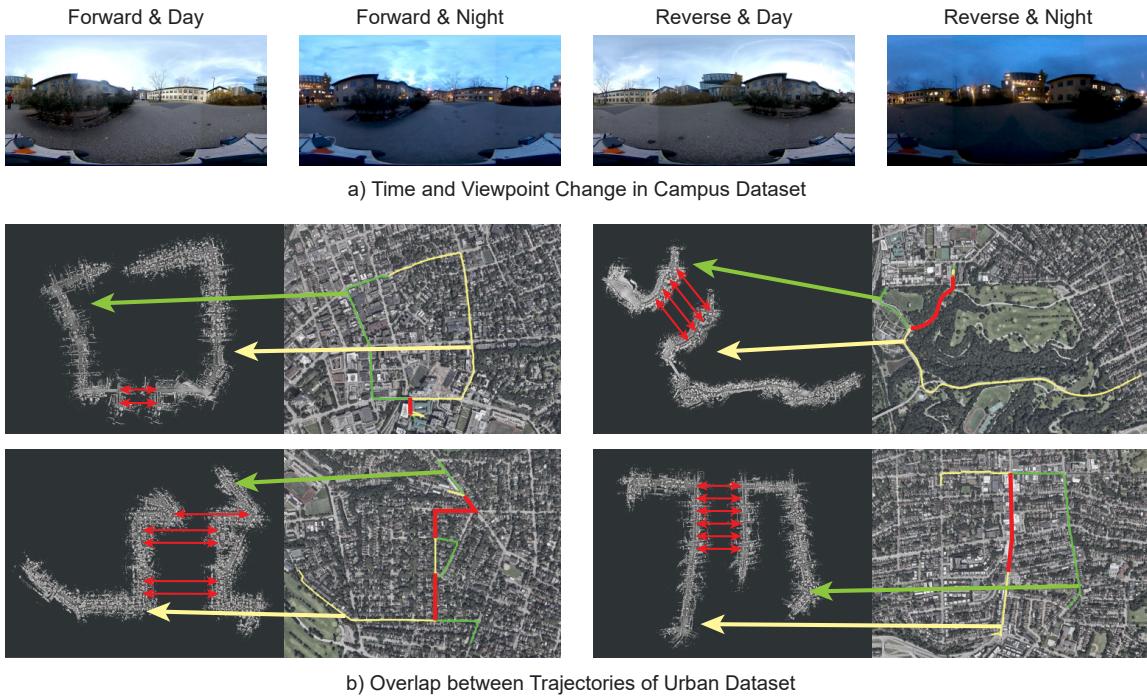b) Overlap between Trajectories of Urban Dataset

Figure 4. a) The Campus is equipped with an omnidirectional camera. As the data is recorded at different times of the day and from different viewpoints, the Campus can be utilized to test the re-localization ability under diverse illumination conditions. b) There are significant overlaps within trajectories of the Urban dataset.

Dataset

Dataset Overview

The present dataset contains two sub-datasets:

Urban dataset is composed of 50 vehicle trajectories, as shown in Fig. 1, covering $120km$ in the city of Pittsburgh, Pennsylvania; The range of *Urban* provides a sufficient quantity of data for extensive network training and high structural diversity, including commercial, residential, downtown, and suburban areas, for improving network robustness. Especially, *Urban* are also designed for the map-merging systems. As shown in Fig. 4, each trajectory is at least overlapped at one junction with the others, and there are 158 overlaps in total within the dataset.

Campus dataset consists of 10 trajectories collected within the campus area of Carnegie Mellon University(CMU), and the total length is around $36km$. Each trajectory is recorded eight times under different conditions(illumination, direction): as shown in Fig. 4, we have four types of combinations and recorded them two times each. Even within a relatively small area, *Campus* contains buildings, corridors, and crossroads, providing sufficient structural diversity for place recognition evaluation. Moreover, omnidirectional pictures with position labels are provided, *Campus* are also be utilized for visual place recognition evaluation under different viewpoint and illumination conditions. Same as *Urban*, there is a total of 9 overlaps.

Data Description and Format

Each trajectory of the Urban dataset consists of 3 types of data, described as follows:

- *Global Maps*: Global maps are processed to contain the 3D structure of each trajectory, which is provided

in Point Cloud Data (`PCD`) file format. We use self-developed LiDAR-Inertial odometry based on LOAM Zhang and Singh (2014) to generate global maps and process the maps with a VoxelGrid filter.
- *odometry*: We save the key poses generated by our SLAM algorithm as odometry information and provide them in (`TXT`) file format. The key poses are within the local coordinate of each trajectory, and the distance between adjacent poses is around $1m$.
- *Raw Data*: In order to offer convenience for map merging tasks, the raw data is provided in `rosbag` ROS package. Inside the raw data, two ROS topics */imu/data* and */velodyne_packets* reveal the inertial measurement unit and LiDAR data. The frequency of two ROS topics is $200hz$ and $10hz$.

Each trajectory of the Campus dataset consists of 8 sequences, and each sequence includes four types of data:

- *Global Maps* and *Odometry* are the same with *Urban*.
- *Unified Odometry*: We utilize interactive SLAM Koide et al. (2021) to find the geometric relations between the key poses of different sequences within the same trajectory and unify them into the same global coordinate. The data is provided in (`TXT`) file format.
- *Omnidirectional Pictures*: For each key pose, a corresponding omnidirectional picture with resolution of 1024×512 is provided in (`PNG`) file format.

Using the Dataset

The `PCD` files for global maps can be easily visualized using PCL Rusu and Cousins (2011) or Open3D Zhou et al. (2018) packages and the `bag` can be played back in the command line by `rosbag` ROS package. In addition,
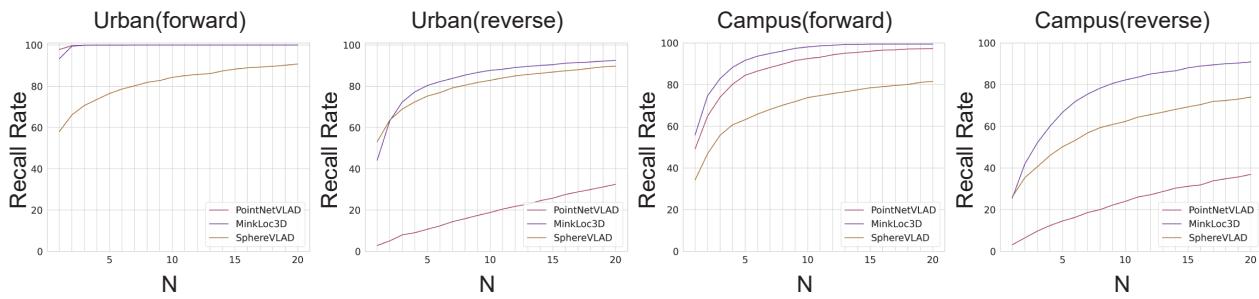
Figure 5. Methods of point-based, voxel-based and projection-based methods are selected to be tested on both datasets

we provide Python-API to access the data and generate training and validation sets based on global maps and corresponding odometry. For *Urban* dataset, we further provide highly personalized query and database frames generation to evaluate the performance of models on any translation and viewpoint differences. As the *Urban* dataset also can be used in large-scale map merging tasks, the ground truth of overlap between trajectories is provided in our API. To offer convenience for models to compare with the state-of-art methods, we provide an online evaluation in AIcrowd to record each submission.

Benchmark Experiments

Models and Evaluation Methods

We select PointNetVLAD Uy and Lee (2018), MinkLoc3D Komorowski (2021) and SphereVLAD Yin et al. (2020) as our baseline models. Implementations in their official Github repository are utilized to train the networks from scratch on Urban dataset trajectory($01 \sim 10 \& 16 \sim 20$) and the unified training data are available at *General Place Recognition Competition*. All models are tested in trajectory($21 \sim 41$) of the Urban dataset for robustness under two combinations of translation and viewpoint differences denoted as forward and reverse. Queries frames are generated by uniformly sampling in the range of $[-3m \sim 3m]$ and $[-5° \sim 5°]$ based on database frames for forward and $[-3m \sim 3m]$ and $[175° \sim 185°]$ for reverse. Furthermore, trajectory($01 \sim 06$) of *Campus* are utilized to evaluate the generalization ability. We use Average Recall curve of top 20 candidates to show the performance of each method, and the Recall@1 is selected to compare each technique's retrieval and generalization ability. A successful retrieval is defined as retrieving a point cloud within 5m for Pittsburgh and 3m for the Campus dataset.

Results

In the Urban dataset, PointNetVLAD outperforms the other methods at Recall@1 forward, and SphereVLAD exceeds the other methods at Recall@1 backward. Our Python-API can help researchers quickly analyze the recognition performance under variant viewpoints. Because the distance between adjacent database frames is only around 3m, the Campus dataset can be used to test the precision of re-localization. As shown in Fig. 5, all the models cannot achieve high recall@1 as in *Urban* and MinkLoc3D outperforms other methods both in forward and backward.

4. Summary and Future Work

In this paper, we presented ALITA dataset which aims to long-term place recognition tasks in large-scale environments. We believe that this dataset will be helpful in place recognition research in handling illumination and viewpoint changes and expect future LiDAR-Image(Omnidirectional) fusion-based robotics research. Since the presented ALITA dataset provides abundant overlaps between trajectories, we also expect future usage for map merging systems. We provided codes to help with using the dataset and evaluating new methods with it.

5. Acknowledgment

References

Stephanie Lowry, Niko SÃ¼nderhauf, Paul Newman, John J. Leonard, David Cox, Peter Corke, and Michael J. Milford. Visual place recognition: A survey. *IEEE Transactions on Robotics*, 32(1):1–19, 2016. doi: 10.1109/TRO.2015.2496823.

Mathieu Nowakowski, Cyril Joly, SÃ©bastien Dalibard, Nicolas Garcia, and Fabien Moutarde. Topological localization using wi-fi and vision merged into fabmap framework. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3339–3344, 2017. doi: 10.1109/IROS.2017.8206171.

Mubariz Zaffar, Shoaib Ehsan, Michael Milford, and Klaus McDonald-Maier. Cohog: A light-weight, compute-efficient, and training-free visual place recognition technique for changing environments. *IEEE Robotics and Automation Letters*, 5(2):1835–1842, 2020. doi: 10.1109/LRA.2020.2969917.

Nate Merrill and Guoquan Huang. Lightweight unsupervised deep loop closure. *CoRR*, abs/1805.07703, 2018. URL http://arxiv.org/abs/1805.07703.

Giseop Kim and Ayoung Kim. Scan context: Egocentric spatial descriptor for place recognition within 3d point cloud map. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4802–4809, 2018. doi: 10.1109/IROS.2018.8593953.

Li He, Xiaolong Wang, and Hong Zhang. M2dp: A novel 3d point cloud descriptor and its application in loop closure detection. In

2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 231–237, 2016. doi: 10.1109/IROS. 2016.7759060.

Prasenjit Mondal, Jayanta Mukhopadhyay, Shamik Sural, and Pinak Pani Bhattacharyya. 3d-sift feature based brain atlas generation: An application to early diagnosis of alzheimer's disease. In *2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*, pages 342–347, 2014. doi: 10.1109/MedCom.2014.7006030.

Relja Arandjelovic, Petr Gronat, Akihiko Torii, Tomas Pajdla, and Josef Sivic. Netvlad: Cnn architecture for weakly supervised place recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5297–5307, 2016. doi: 10.1109/CVPR.2016.572.

Mikaela Angelina Uy and Gim Hee Lee. Pointnetvlad: Deep point cloud based retrieval for large-scale place recognition. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4470–4479, 2018. doi: 10.1109/CVPR. 2018.00470.

Xieyuanli Chen, Thomas Läbe, Andres Milioto, Timo Röhling, Olga Vysotska, Alexandre Haag, Jens Behley, and Cyrill Stachniss. Overlapnet: Loop closing for lidar-based SLAM. *CoRR*, abs/2105.11344, 2021. URL https://arxiv. org/abs/2105.11344.

Zhe Liu, Shunbo Zhou, Chuanzhe Suo, Peng Yin, Wen Chen, Hesheng Wang, Haoang Li, and Yunhui Liu. Lpd-net: 3d point cloud learning for large-scale place recognition and environment analysis. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 2831–2840, 2019. doi: 10.1109/ICCV.2019.00292.

Bastian Steder, Giorgio Grisetti, and Wolfram Burgard. Robust place recognition for 3d range data based on point features. In *2010 IEEE International Conference on Robotics and Automation*, pages 1400–1405, 2010. doi: 10.1109/ROBOT. 2010.5509401.

Gaurav Pandey, James R McBride, and Ryan M Eustice. Ford campus vision and lidar data set. *The International Journal of Robotics Research*, 30(13):1543–1552, 2011. doi: 10. 1177/0278364911400640. URL https://doi.org/10. 1177/0278364911400640.

A Geiger, P Lenz, C Stiller, and R Urtasun. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11):1231–1237, 2013. doi: 10. 1177/0278364913491297. URL https://doi.org/10. 1177/0278364913491297.

Nicholas Carlevaris-Bianco, Arash K Ushani, and Ryan M Eustice. University of michigan north campus long-term vision and lidar dataset. *The International Journal of Robotics Research*, 35(9):1023–1035, 2016. doi: 10. 1177/0278364915614638. URL https://doi.org/10. 1177/0278364915614638.

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. doi: 10.1177/0278364916679498. URL https://doi. org/10.1177/0278364916679498.

Giseop Kim, Yeong Sang Park, Younghun Cho, Jinyong Jeong, and Ayoung Kim. Mulran: Multimodal range dataset for urban place recognition. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6246–6253, 2020. doi:

10.1109/ICRA40945.2020.9197298.

Yiyi Liao, Jun Xie, and Andreas Geiger. KITTI-360: A novel dataset and benchmarks for urban scene understanding in 2d and 3d. *CoRR*, abs/2109.13410, 2021. URL https:// arxiv.org/abs/2109.13410.

Dominik Van Opdenbosch and Eckehard Steinbach. Collaborative visual slam using compressed feature exchange. *IEEE Robotics and Automation Letters*, 4(1):57–64, 2019. doi: 10.1109/LRA. 2018.2878920.

Yulun Tian, Yun Chang, Fernando Herrera Arias, Carlos Nieto-Granda, Jonathan P. How, and Luca Carlone. Kimera-multi: Robust, distributed, dense metric-semantic slam for multi-robot systems. *IEEE Transactions on Robotics*, pages 1–17, 2022. doi: 10.1109/TRO.2021.3137751.

Charles Ruizhongtai Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 77–85. IEEE Computer Society, 2017. doi: 10.1109/CVPR.2017.16. URL https://doi.org/10.1109/CVPR.2017.16.

Jacek Komorowski. Minkloc3d: Point cloud based large-scale place recognition. In *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1789–1798, 2021. doi: 10.1109/WACV48630.2021.00183.

Peng Yin, Fuying Wang, Anton Egorov, Jiafan Hou, Ji Zhang, and Howie Choset. Seqspherevlad: Sequence matching enhanced orientation-invariant place recognition. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5024–5029, 2020. doi: 10.1109/IROS45743. 2020.9341727.

Junyi Ma, Jun Zhang, Jintao Xu, Rui Ai, Weihao Gu, and Xieyuanli Chen. Overlaptransformer: An efficient and yaw-angle-invariant transformer network for lidar-based place recognition. *IEEE Robotics and Automation Letters*, 7(3): 6958–6965, 2022. doi: 10.1109/LRA.2022.3178797.

Peng Yin, Haowen Lai, Shiqi Zhao, Ruijie Fu, Ivan Cisneros, Ruohai Ge, Ji Zhang, Howie Choset, and Sebastian Scherer. Automerge: A framework for map assembling and smoothing in city-scale environments, 2022. URL https://arxiv. org/abs/2207.06965.

Ji Zhang and Sanjiv Singh. LOAM: lidar odometry and mapping in real-time. In Dieter Fox, Lydia E. Kavraki, and Hanna Kurniawati, editors, *Robotics: Science and Systems X, University of California, Berkeley, USA, July 12-16, 2014*, 2014. doi: 10.15607/RSS.2014.X.007. URL http://www. roboticsproceedings.org/rss10/p07.html.

Kenji Koide, Jun Miura, Masashi Yokozuka, Shuji Oishi, and Atsuhiko Banno. Interactive 3d graph slam for map correction. *IEEE Robotics and Automation Letters*, 6(1):40–47, 2021. doi: 10.1109/LRA.2020.3028828.

Radu Bogdan Rusu and Steve Cousins. 3d is here: Point cloud library (pcl). In *2011 IEEE International Conference on Robotics and Automation*, pages 1–4, 2011. doi: 10.1109/ ICRA.2011.5980567.

Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Open3d: A modern library for 3d data processing, 2018. URL https://arxiv. org/abs/1801.09847.