DiTer++: Diverse Terrain and Multi-modal Dataset for Multi-Robot Navigation in Multi-session Outdoor Environments

Juwon Kim[1*], Seokhwan Jeong[1*], Hogyun Kim[1] and Younggun Cho[1†]

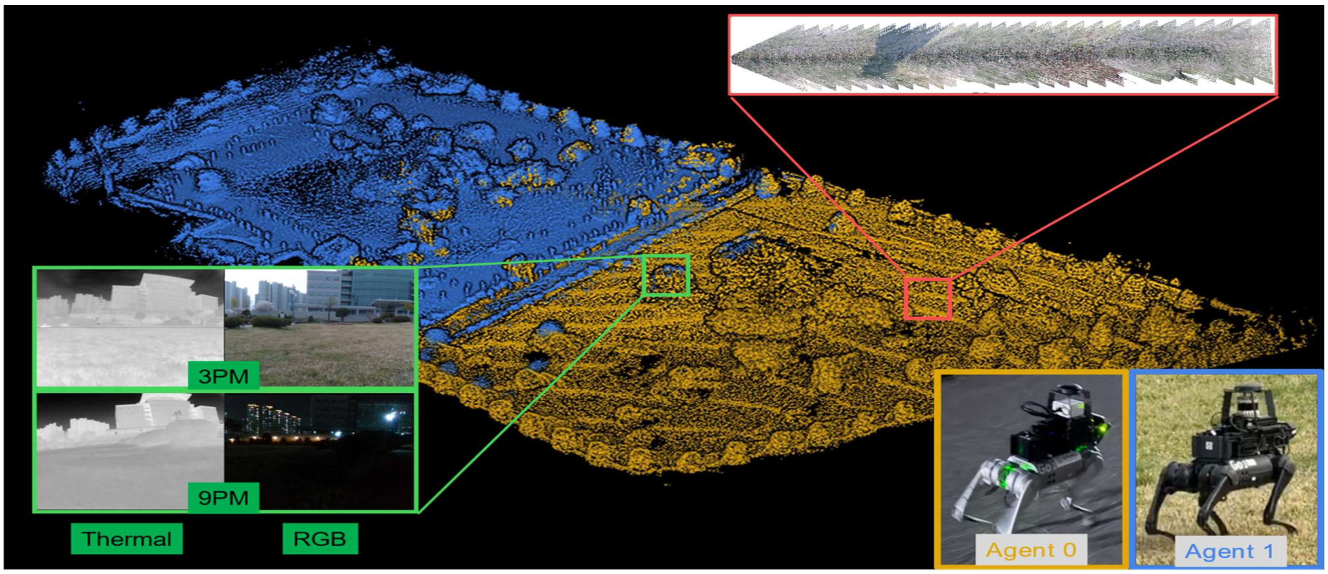

Fig. 1. A large-scale map from heterogeneous LiDARs. Yellow represents the map acquired using the SOSLAB's ML-X LiDAR (solid state LiDAR) with an `Agent 0` robot, and blue represents the map acquired using the Ouster OS1-64 (spinning LiDAR) with an `Agent 1` robot. The green square represents a thermal image that shows temperature changes and an RGB image that shows illumination changes between day and night. Additionally, the red square represents a local ground map obtained through a ground-facing RGB-D camera.

***Abstract*— Collaboration of multiple field robots is necessary for the navigation and mapping of large-scale environments. While traversing, traversability estimation considering each robot's nature is essential for keeping the robot safe and ensuring its performance. Even in a structured environment, driving without considering terrain information can lead to serious damage to the platform, such as slipping due to steep slopes or falling caused by sudden height changes. To address this challenge, we present DiTer++, multi-robot, multi-session, and multi-modal datasets, including ground-level information. The dataset is obtained with a forward-facing RGB camera and ground-facing RGB-D camera, a thermal camera, two types of LiDARs, IMU, GPS, and robot motion sensors. The dataset and supplement materials are available at `https://sites.google.com/view/diter-plusplus/`.**

I. Introduction

Multi-robot collaborative simultaneous localization and mapping (SLAM) plays a groundbreaking role in the efficient mapping of large-scale outdoor environments such as construction site where hazards and dynamic objects are spread. Therefore, sharing the understanding of perception among robots is necessary to conduct mapping tasks in such challenging environments.

Recent studies have reported about [4]-[11] multi-robot management in outdoor environments. However, since each robot has distinct hardware characteristics that determines the traversability in its operating area, ground-level information of each robot is necessary. Also, as the RGB camera is blind at night, including the thermal camera data from both day and night to complement the tracking loss is also needed. Unlike the aforementioned datasets that only employ spinning light detection and ranging (LiDAR), we utilize heterogeneous (spinning, solid) LiDAR to enrich the dataset.

Our previous version of this article [1] proposed multi-modal and multi-session dataset consisted of terrain information obtained from a ground-facing RGB-D camera, but challenges from multi-robot, night navigation, and heterogeneous LiDAR still remained. To overcome the challenges and to address more various situations, we present new contributions as follows:

- **Multi-robot:** We employ two quadruped robots for data acquisition. Each robot is assigned to traverse a certain coverage with diverse terrain, including lawn,

[1*]Juwon Kim, [1*]Seokhwan Jeong, [1]Hogyun Kim and [1†]Younggun Cho are with the Electrical and Computer Engineering, Inha University, Incheon, South Korea `lambertkim@naver.com, [eric5709, hg.kim]@inha.edu, yg.cho@inha.ac.kr`

(*) represents equal contribution.

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (RS-2023-00302589) and Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00448).

TABLE I

VARIOUS MULTI-ROBOT AND DIVERSE TERRAIN DATASETS

Datasets	Hardware	Environments	Multi-Robot	Multi-Session	Ground RGB-D	RGB	Thermal	Heterogeneous LiDARs	GPS	IMU	Robot Motion
DiTer [1]	Quadruped	Diverse Terrain	x	✓	✓	✓	✓	x	✓	✓	✓
Katwijk [2]	Rover	Soft Terrain	x	x	✓	✓	x	x	✓	✓	x
TAIL [3]	Wheeled / Quadruped	Soft Terrain	x	x	✓	✓	x	x	✓	✓	✓
WildPlaces [4]	Handheld	Jungle	x	✓	x	x	x	x	x	✓	x
CWT [5]	Excavator	Construction site	x	x	x	✓	x	x	x	x	x
PARK [6]	Wheeled	Outdoor	✓	x	x	x	x	x	✓	✓	x
LAMP 2.0 [7]	Wheeled / Quadruped	In/Outdoor	✓	x	x	x	x	x	x	x	x
Kimera-Multi [8]	Wheeled	In/Outdoor	✓	x	x	✓	x	x	x	✓	x
GRACO [9]	Wheeled / Drone	Outdoor	✓	x	x	✓	x	x	✓	✓	x
S3E [10]	Wheeled	In/Outdoor	✓	x	x	✓	x	x	✓	✓	x
RING++ [11]	Quadruped	Outdoor	✓	x	x	x	x	x	x	x	x
DiTer++	Quadruped	Diverse Terrain	✓	✓	✓	✓	✓	✓	✓	✓	✓

curb, asphalt, etc.

- **Multi-session:** Our dataset contains day and night scenarios for each sequence.
- **Multi-modal:** Our dataset is collected by multiple perceptual sensors, including RGB, RGB-D, thermal camera, and LiDARs with heterogeneous scan patterns as shown in Fig. 1. Additionally, we equip the robot with sensors that assist in navigation, including a built-in motion sensor, Global Positioning System (GPS), and inertial measurement unit (IMU). A brief setup of sensor and acquired data is represented in the Fig. 2.

II. RELATED WORKS

As represented in Table I, we review various works related to multi-robot and diverse terrain datasets with *DiTer++*, an extension of the previous work [1].

1) Multi-Robot datasets: PARK [6] consists of sequences in which three robots equipped with LiDAR navigate. *LAMP* 2.0 is a dataset used in the *DARPA SubT Challenge*. It provides sequences from a variety of situations, with a single LiDAR mounted on the robot. *Kimera-Multi* is a dataset focused on vision-based distributed semantic SLAM. However, it equips each robot with LiDAR to provide ground truth depth. *GRACO* is a multi-session and multi-modal driving dataset for drones and unmanned ground vehicle (UGV). It reports the results of distributed SLAM, thereby proving its value as a synthetic distributed dataset. *S3E* provides a variety of indoor/outdoor sequences. It demonstrates the utility of this dataset through the application of DCL SLAM [12]. *RING++* efficiently conducts multi-robot mapping through its own quadruped robot dataset.

2) Diverse Terrain datasets: DiTer is an initial version of this paper, focusing on diverse sensors acquiring data over the long term. It aims to address the problems of multi-session scenarios, but due to the absence of sequences captured at night, overcoming lighting changes between day and night through this dataset is challenging. *Katwijk*, *TAIL* [2, 3] acquire various sensor data in soft terrains, such as sand. Each dataset utilizes a rover or quadruped robot to overcome challenges in these environments. However, they also face issues with illumination changes between day and night due to the absence of thermal cameras. Additionally, since the datasets are not acquired over a long period, they are not suitable for addressing multi-session problems. *Wild-Places* is a dataset acquired over a long period but is only equipped with LiDAR sensors. Since it does not capture dense information, the map building results sparse, hindering map to contain rich features.

III. SYSTEM SETUP AND CALIBRATION

A. Sensor measurement, setup, and topics

As shown in Fig. 2, each robot shares a partially identical sensor setup, followed by different LiDAR and IMU. Sensor specifications for each agent are configured in Table II. To inherit the main purpose of our previous work [1], both agents have the identical structure of a forward-looking sensor system consisting of RGB, thermal, and tilted RGB-D cameras. The following sensors are placed heading forward to obtain sufficient visual measurement and ground information simultaneously.

TABLE II

SENSOR SPECIFICATIONS AND ROSTOPIC NAME

Agent ID	Hardware	Sensors	Specifications	Topic name
Agent 0	Intel NUC	RGB-D	Intel Realsense D435i	/agent0/ground/depth/image_raw /agent0/ground/depth/camera_info /agent0/ground/color/image_raw /agent0/ground/color/camera_info
		RGB	Intel Realsense D435i	/agent0/front/color/image_raw /agent0/front/color/camera_info
		Thermal	FLIR Boson ADK	/agent0/flir_boson/image_raw /agent0/flir_boson/camera_info
		LiDAR	SOSLAB ML-X	/agent0/ml/points
		GPS	Ublox C099-F9P	/agent0/ublox_gps/fix
		9DoF-IMU	Microstrain 3DM-CV7	/agent0/imu/data
	Unitree-GO1	6DoF-IMU	Built-in Robot	/agent0/odom /agent0/imu
Agent 1	Intel NUC	RGB-D	Intel Realsense D435i	/agent1/ground/depth/image_raw /agent1/ground/depth/camera_info /agent1/ground/color/image_raw /agent1/ground/color/camera_info
		RGB	Intel Realsense D435i	/agent1/front/color/image_raw /agent1/front/color/camera_info
		Thermal	FLIR Boson ADK	/agent1/flir_boson/image_raw /agent1/flir_boson/camera_info
		LiDAR	Ouster OS1-64	/agent1/ouster/points
		GPS	Ublox C099-F9P	/agent1/ublox_gps/fix
		9DoF-IMU	Microstrain 3DM-GX5-25	/agent1/imu/data
	Unitree-GO1	6DoF-IMU	Built-in Robot	/agent1/odom /agent1/imu

Subsequently, the LiDAR sensor is located at the rear side of the visual measurement mount, providing relatively accurate geometric measurement. Each agent utilizes LiDAR with heterogeneous scan patterns, measurement range, and FOV. `Agent 0` provides full azimuth measurement, enabling full recognition of the surrounding environment. `Agent 1` provides limited FOV, but still has accurate range measurement from forward. GPS, IMU, built-in IMU, and odometry are

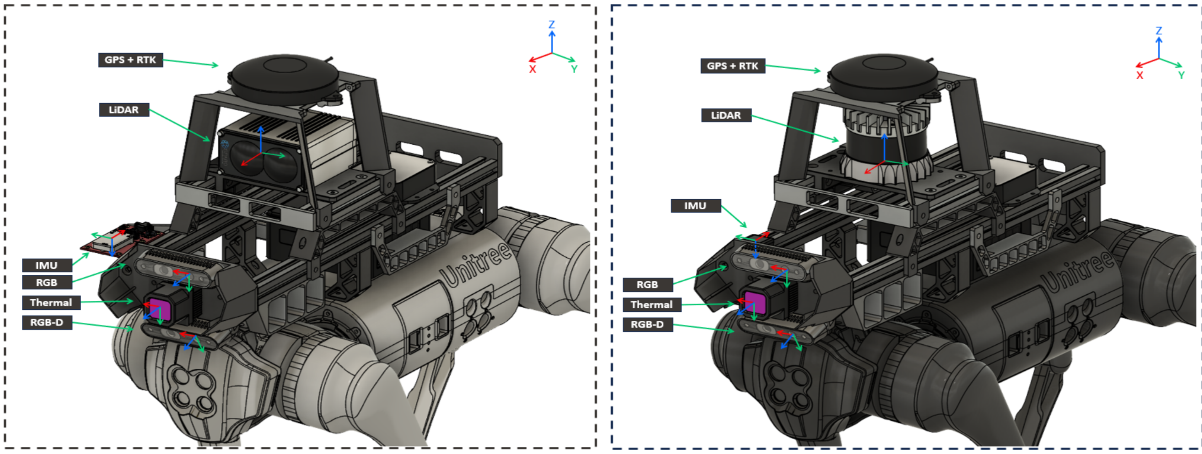

Fig. 2. Illustration of robots with holistic and perceptual sensors with different LiDAR setups: solid-state and mechanical types.

also provided considering the robot's navigation. Agent notation is done to distinguish acquired data with namespace, which can also be found in the obtained data topics.

B. Calibration

1) Intrinsic Calibration: For calibrating camera's intrinsic parameter, we utilize `ROS` camera calibration. In terms of thermal camera, pixel value inversion is applied to the thermal image since the values of the black and white pattern are opposed. For more details, we suggest to refer [1].

2) Extrinsic Calibration: For extrinsic calibration, we apply marker-based LiDAR-camera calibration[13] for forward-facing RGB and thermal camera between LiDAR. For LiDAR-IMU calibration, we utilize the robust real-time LiDAR-inertial initialization method proposed in [14].

Since calibration between ground-facing RGB and LiDAR is not straightforward due to the non-overlapping region, we adopt extrinsic parameters for the two RGB cameras from the CAD model we designed. The output of the merged point clouds is represented in Fig. 3.

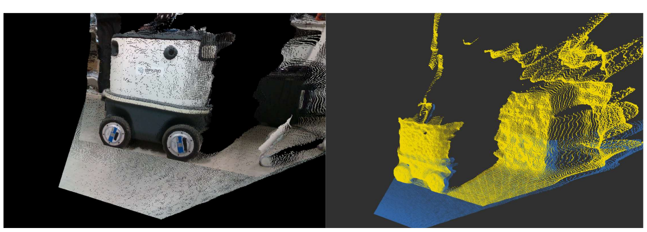

Fig. 3. (Left) Merged point cloud of two depth cameras. (Right) The blue color means a point cloud from the ground-facing RGB-D camera. The yellow color represents a point cloud from the forward-facing camera.

IV. DATASET

All sequences are obtained within the outdoor sites and are conducted at two different places, noted as LAWN and PARK, respectively. Each sequence contains multi-session and multi-robot scenarios that deploy agents with varying setups of sensors. LAWN is conducted by `Agent 0` and `Agent 1`, which are distinguishable by sensor setups. PARK is conducted only with `Agent 1`. Table III summarizes terrain, etc. Details of each sequence are as follows:

1) LAWN is obtained from an extremely unstructured environment with terrains showing high level of vegetation. In this sequence, `Agent 0` and `Agent 1` travels around the boundary of the upper and lower region of the LAWN sequence, then moves to the center of each region. Data obtained from both agents include scenes with occlusion caused by vegetation, and challenging traverse scenarios against stones and narrow bushes.
2) PARK is obtained around a campus building that contains multiple environments. Both inner and outer regions of the PARK are covered with `Agent 1` setup. The inner region of the PARK consists of a narrow underground area and unstructured terrain in the center, while the outer region consists of pedestrian walkway, urban area, and a route around the construction site. PARK sequence includes scene with multiple dynamic objects and various illumination conditions from environment changes.

Sequences contain overlapping regions and shared perspectives of view from each sensor. As shown in Fig. 4, the overlapping region is comparatively large in the LAWN sequence, while the PARK sequence only shows a limited amount of overlap near the start point. Fig. 5 shows the alignment of maps acquired from each robot, leveraged by Point-LIO[15] to check our dataset's capability of collaborative map building visually.

V. CONCLUSION

We propose DiTer++, an extension of our previous work [1]. DiTer++ exploits two quadrupedal robots, which provide both proprioceptive measurements of the robot and exteroceptive measurement from perceptive sensors. Our sequences contain multiple challenges for generic mobile platforms to be traversed, including rapid environment changes, unstructured terrains with extreme height changes, and etc. Rich scenarios and comprehensive measurement of our dataset allows research on long-term autonomy, environment change

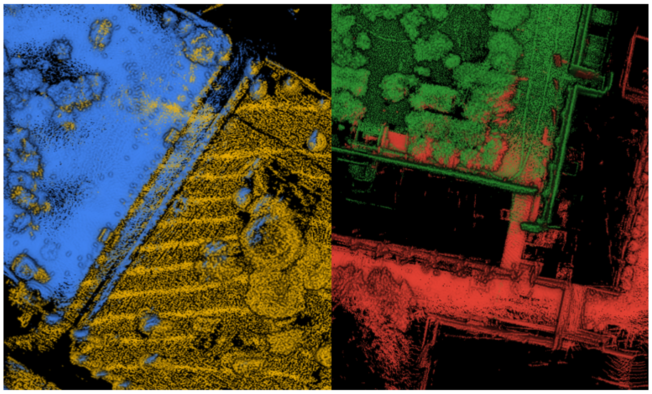

Fig. 4. Overlapping region visualization from both sequences (LAWN and PARK).

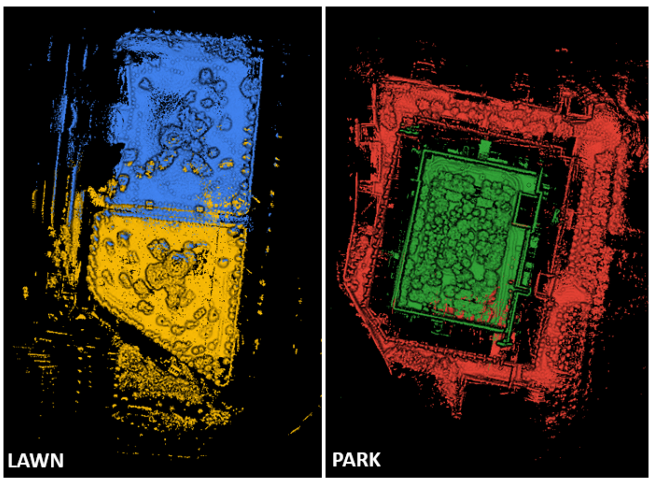

Fig. 5. Global map visualization of both sequences (PARK and LAWN).

detection, and autonomous survey task in construction environments. Our dataset especially targets development of long-term SLAM and traversability mapping algorithm for both single and multi-robot in construction.

TABLE III
SUMMARY OF OUR MULTI-SESSION DATASET

Sequence Name	Terrain	Length	Duration
LAWN 0 & 1	Lawn with hills	351m / 204m	790s / 239s
PARK 0 & 1	Vegetation / Sidewalk and Asphalt	446m / 244m	1070s / 459s

REFERENCES

[1] S. Jeong, H. Kim, and Y. Cho, "Diter: Diverse terrain and multi-modal dataset for field robot navigation in outdoor environments," *IEEE Sensors Letters*, 2024.

[2] R. A. Hewitt, E. Boukas, M. Azkarate, M. Pagnamenta, J. A. Marshall, A. Gasteratos, and G. Visentin, "The katwijk beach planetary rover dataset," *The International Journal of Robotics Research*, vol. 37, no. 1, pp. 3–12, 2018.

[3] C. Yao, Y. Ge, G. Shi, Z. Wang, N. Yang, Z. Zhu, H. Wei, Y. Zhao, J. Wu, and Z. Jia, "Tail: A terrain-aware multi-modal slam dataset for robot locomotion in deformable granular environments," *arXiv preprint arXiv:2403.16875*, 2024.

[4] J. Knights, K. Vidanapathirana, M. Ramezani, S. Sridharan, C. Fookes, and P. Moghadam, "Wild-places: A large-scale dataset for lidar place recognition in unstructured natural environments," in *2023 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2023, pp. 11 322–11 328.

[5] T. Guan, Z. He, R. Song, D. Manocha, and L. Zhang, "Tns: Terrain traversability mapping and navigation system for autonomous excavators," *arXiv preprint arXiv:2109.06250*, 2021.

[6] Y. Huang, T. Shan, F. Chen, and B. Englot, "Disco-slam: Distributed scan context-enabled multi-robot lidar slam with two-stage global-local graph optimization," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1150–1157, 2021.

[7] Y. Chang, K. Ebadi, C. E. Denniston, M. F. Ginting, A. Rosinol, A. Reinke, M. Palieri, J. Shi, A. Chatterjee, B. Morrell *et al.*, "Lamp 2.0: A robust multi-robot slam system for operation in challenging large-scale underground environments," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9175–9182, 2022.

[8] Y. Tian, Y. Chang, F. H. Arias, C. Nieto-Granda, J. P. How, and L. Carlone, "Kimera-multi: Robust, distributed, dense metric-semantic slam for multi-robot systems," *IEEE Transactions on Robotics*, vol. 38, no. 4, 2022.

[9] Y. Zhu, Y. Kong, Y. Jie, S. Xu, and H. Cheng, "Graco: A multimodal dataset for ground and aerial cooperative localization and mapping," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 966–973, 2023.

[10] D. Feng, Y. Qi, S. Zhong, Z. Chen, Y. Jiao, Q. Chen, T. Jiang, and H. Chen, "S3e: A large-scale multi-modal dataset for collaborative slam," *arXiv preprint arXiv:2210.13723*, 2022.

[11] X. Xu, S. Lu, J. Wu, H. Lu, Q. Zhu, Y. Liao, R. Xiong, and Y. Wang, "Ring++: Roto-translation-invariant gram for global localization on a sparse scan map," *IEEE Transactions on Robotics*, 2023.

[12] S. Zhong, Y. Qi, Z. Chen, J. Wu, H. Chen, and M. Liu, "Dcl-slam: A distributed collaborative lidar slam framework for a robotic swarm," *IEEE Sensors Journal*, 2023.

[13] D. Tsai, S. Worrall, M. Shan, A. Lohr, and E. Nebot, "Optimising the selection of samples for robust lidar camera calibration," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2021, pp. 2631–2638.

[14] F. Zhu, Y. Ren, and F. Zhang, "Robust real-time lidar-inertial initialization," in *Proc. IEEE/RSJ Intl. Conf. on Intell. Robots and Sys.* IEEE, 2022, pp. 3948–3955.

[15] D. He, W. Xu, N. Chen, F. Kong, C. Yuan, and F. Zhang, "Point-lio: Robust high-bandwidth light detection and ranging inertial odometry," vol. 5, no. 7, 2023, p. 2200459.