Self-Supervised Terrain Representation Learning from Unconstrained Robot Experience

Anonymous Author(s)
Affiliation
Address
email

Abstract: *Terrain awareness*, i.e., the ability to *identify* and *distinguish* different types of terrain, is a critical ability that robots must have to succeed at autonomous off-road navigation. Current approaches that provide robots with this awareness either rely on labeled data which is expensive to collect, engineered features and cost functions that may not generalize, or expert human demonstrations which may not be available. Towards endowing robots with terrain awareness without these limitations, we introduce *Self-supervised TErrain Representation LearnING* (STERLING), a novel approach for learning terrain representations that relies solely on easy-to-collect, unconstrained (e.g., non-expert), and unlabelled robot experience, with no additional constraints on data collection. STERLING employs a novel multi-modal self-supervision objective through non-contrastive representation learning to learn relevant terrain representations for terrain-aware navigation. Through physical robot experiments in off-road environments, we evaluate STERLING features on the task of preference-aligned visual navigation and find that STERLING features perform on par with fully-supervised approaches and outperform other state-of-the-art methods with respect to preference alignment. Additionally, we perform a large-scale experiment of autonomously hiking a 3-mile long trail which STERLING completes successfully with only two manual interventions, demonstrating its robustness to real-world off-road conditions.

Keywords: Vision-Based Navigation, Representation Learning.

1 Introduction

Off-road navigation is emerging as a crucial capability for autonomous mobile robots envisioned for use in a growing number of outdoor applications such as agricultural operations [1], package delivery [2], and search and rescue [3]. The ability to visually identify distinct terrain features that are relevant to the navigation task, referred to here as *terrain awareness*, is necessary for enabling off-road navigation.

Endowing robots with terrain awareness is a challenging problem in autonomous off-road navigation [4, 5]. Prior work on terrain awareness has typically relied on difficult-to-collect curated datasets [6, 7, 8, 9] or has been focused on particular tasks [10, 11, 12, 13, 14] and is not amenable to downstream task changes [15, 14, 16]. To enable off-road autonomy at scale, new approaches are needed that address the perception problem of learning relevant terrain representations in a self-supervised way from easy-to-collect unconstrained robot experiences gathered without requiring an expert agent in the field. Moreover, the learned representations need to be general enough to support a variety of downstream navigation tasks (e.g., changing preferences over terrain types). Learning such rich visual features of the terrain through self-supervision by leveraging multi-modal sensor data collected onboard a mobile robot can help tackle many off-road navigation problems such as high-speed navigation [4, 17], traversing challenging terrain [18, 19], and preference-aligned planning [20, 9].

Submitted to the 7th Conference on Robot Learning (CoRL 2023). Do not distribute.

Towards robot terrain awareness from unconstrained and unlabelled robot experience, we introduce *Self-supervised TErrain Representation LearnING* (STERLING), a novel approach to learning terrain representations for off-road navigation. STERLING utilizes easy-to-collect unconstrained robot experiences, without any additional restrictions on data collection (e.g., that the trajectories be of expert demonstrations), and employs a novel multi-modal self-supervision objective through non-contrastive, non-reconstructive representation learning to learn relevant terrain representations useful for visual terrain-aware navigation. The key contributions of this paper are— 1) *Self-supervised TErrain Representation LearnING* (STERLING), a novel approach that learns terrain representations from easy-to-collect unconstrained robot experiences, 2) Detailed evaluation of STERLING against baseline methods on the task of operator preference-aligned off-road navigation, and 3) A large-scale qualitative experiment of autonomously hiking a 3-mile long trail, demonstrating the effectiveness of STERLING-features.

We execute all of our off-road physical robot experiments on a Boston Dynamics Spot and compare features learned using STERLING to existing visual representation learning-based methods SE-R [5] and GANav [9], self-supervised cost learning approach RCA [16], and a fully-supervised oracle. We find that STERLING performs on par with or better than baseline approaches in all the outdoor environments with respect to preference alignment, without requiring expert labels or demonstrations.

2 Related Work

In this section, we review related work on terrain-aware visual off-road navigation. We specifically focus on approaches that learn to navigate off-road conditions using supervised and self-supervised learning.

2.1 Supervised Methods

Several approaches in the past have proposed using supervised learning from large-scale data to navigate off-road environments. We divide them into two categories as follows.

End-to-End Learning: The initial success of applying learning-based solutions to off-road terrain-aware navigation was by LeCun et al. [21] who used a convolutional network to learn to drive in off-road conditions. More recently, Bojarski et al. [10] trained a deep neural network end-to-end using several miles of driving data collected on a vehicle in the real world. While both approaches were promising in urban and off-road environments, end-to-end methods require large amounts of data and are well-known to suffer from domain and covariate shifts [22].

Image Segmentation: Unlike end-to-end approaches that learn behaviors, segmentation-based approaches either classify or pixel-wise segment terrains into known classes [23, 9, 24], trained using supervised learning on large labeled datasets. Guan et al. [9] propose GANav, a transformer-based architecture to pixel-wise segment terrains, trained on RELLIS [7] and RUGD [6] datasets, with manually assigned terrain costs. While effective at terrain awareness, segmentation-based methods are fixed to the specific terrain types available in the datasets and require additional labeling effort to generalize to novel terrains. In STERLING, we do not require semantically labeled datasets and learn terrain representations from unconstrained experience collected onboard a mobile robot.

2.2 Self-Supervised Learning

To alleviate the need for extensive human labeling, self-supervised learning methods have been proposed to either learn terrain representations or costs from data gathered onboard a mobile robot.

Representation Learning: Brooks et al. [25] utilize contact vibrations and visual sensors to classify terrains via self-supervision. Zürn et al. [5] introduce SE-R which utilizes acoustic and visual sensors on the robot to segment terrains using a self-supervised triplet-contrastive learning framework. Using triplet-based contrastive learning methods requires negative samples which may not be available when learning using unlabeled data. In STERLING, we use recently proposed non-contrastive

unsupervised learning approaches such as VICReg [26] that do not require any negative samples and instead rely on correlations between data modalities to learn relevant terrain representations.

Cost Learning: Several methods have applied self-supervision to assign traversability costs for the downstream off-road navigation task [16, 27, 15, 28]. Specifically, these methods rely on inertial spectral features [16], future predictive models [15], inertial-odometry errors [27], or foothold positions [28] as self-supervision signals to learn a traversability cost map, used to evaluate candidate actions. Instead of inferring costs or rewards using self-supervision for a fixed task, in this work, we focus on learning relevant visual features from unconstrained robot experiences that could be used in downstream tasks. This framework allows a designer to reuse features across tasks without retraining entirely from scratch.

Hybrid Methods: The approach closest to ours is VRL-PAP [20] which requires human expert teleoperated demonstrations of a particular trajectory pattern to both explicitly learn visual terrain representations as well as to infer terrain preference costs. However, in this work, we focus on learning terrain features from unconstrained robot experiences without requiring human experts in the field for demonstrations, which is a more general problem than the one considered by VRL-PAP.

3 Approach

In this section, we introduce the self-supervised terrain representation learning approach, STERLING, proposed in this work. We first describe the offline pre-processing performed on recorded unconstrained robot data and then summarize the self-supervision objectives. Finally, we describe the problem formulation for preference-aligned off-road navigation and present how features learned using STERLING can be utilized within a planner for terrain-aware and preference-aligned navigation.

3.1 Data-Collection and Pre-Processing

STERLING learns terrain representations from unconstrained, unlabeled robot experiences collected using any navigation policy. This policy may be, for instance, non-expert human teleoperation, curiosity-driven exploration, or point-to-point navigation using any underlying planner. Compared to requiring a human expert to provide teleoperated demonstrations and labels, collecting this type of robot experience is cheap and easy. We additionally assume that the robot is equipped with *multiple sensors*, e.g., an egocentric RGB camera, odometry sensors, onboard IMU, proprioceptive, and tactile sensors, that together provide rich multi-modal observations as the robot traverses over different terrains collecting experience. STERLING leverages this multi-modal data by using the correlation between different modalities to inform the learned terrain representations.

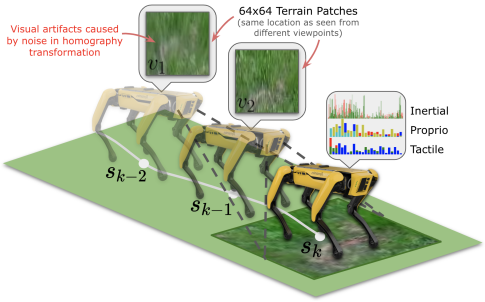

Figure 1: An illustration of the pre-processing performed on unconstrained robot experience. Image patches of traversed terrain at location s_k are extracted from bird's eye view observations at prior locations s_{k-1}, s_{k-2} along the trajectory. The corresponding IPT observations at s_k are transformed from time series to PSD signals.

In order to learn terrain representations using STERLING, we begin by pre-processing the visual and non-visual observations, which are explained in detail below.

Visual Patch Extraction: The egocentric camera observations are homography-projected into a virtual bird's eye view (BEV) frame, assuming that the ground is a flat plane, using the intrinsic and extrinsic camera matrices. As shown in Fig. 1, we project the robot's trajectory onto the bird's eye view frame and extract 64-by-64 pixels (equivalent to the robot's footprint of 0.5-by-0.5 meters)

square visual patches of the terrain along with the corresponding inertial, proprioceptive, and tactile measurements at the same location, along the trajectory. Since the terrain at a particular location s_k is unobservable when the robot itself is at s_k (i.e., it is underneath the robot), we extract terrain image patches corresponding to s_k from BEV observations made by the robot from previous locations $s_{k-1}, s_{k-2}, \ldots$ along its trajectory. Fig. 1 illustrates the offline patch extraction process from two previous viewpoints, however, we extract patches from up to 20 previous viewpoints within 2 meters. Although just one viewpoint is sufficient to learn the correlation between visual and other sensor observations, when planning to navigate, the robot will need to visually evaluate terrain at future locations, and therefore STERLING also seeks representations that are invariant to patch differences due to viewpoint.

IPT Pre-Processing: For the inertial, proprioceptive, and tactile (IPT) observations, we retain up to 2-second history and convert the time-series signals into power-spectral density (PSD) representation in the frequency domain. This ensures the IPT time-series data representations used as input to STERLING are invariant to differences in length and phase in the recorded signals.

3.2 Non-Contrastive Terrain Representation Learning

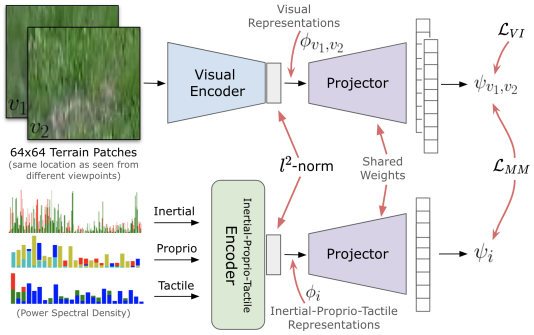

Figure 2: Overview of the training architecture in STERLING. Terrain patches v_1 and v_2 from different viewpoints of the same location are encoded as ϕ_{v_1} and ϕ_{v_2} respectively, and mapped into embeddings ψ_{v_1} and ψ_{v_2}. Similarly, inertial, proprio, tactile signals are encoded as ϕ_i, and mapped as ψ_i. Self-supervision objectives $\mathcal{L}_{VI}$ for viewpoint-invariance and $\mathcal{L}_{MM}$ for multi-modal correlation are computed on the minibatch to perform gradient descent.

It is desired for learned representations of terrains to be such that representations of similar terrain are close together in the embedding space and that representations of different terrains are sufficiently far apart. Although we do not possess information such as semantic labels of terrains for training, the visual and kinodynamic observations experienced by the robot reflect similarities and differences between terrain samples. For instance, traversing a smooth terrain that a human may refer to as `cement sidewalk` may lead to relatively smooth motion by the robot's joints, whereas a rough terrain such as what might be referred to as `marble rocks` may correspond to jerkier motion. STERLING leverages this multi-modal experience observed by the robot and computes a correlation objective between visual and inertial-proprio-tactile signals to learn desired terrain representations. Additionally, STERLING uses viewpoint invariance as an objective unique to the visual component of the experience to learn viewpoint-invariant terrain representations.

Fig. 2 provides an overview of the self-supervised representation learning framework adopted in STERLING. A parameterized visual encoder (4-layer CNN with 0.25 million parameters) encodes terrain image patch observations v_1 and v_2 of the same location s into visual representations ϕ_{v_1} and ϕ_{v_2}, respectively. Similarly, an inertial-proprio-tactile encoder (4-layer MLP with 0.25 million parameters) encodes frequency domain IPT observations of the robot at that location to an inertial-proprio-tactile representation ϕ_i. We follow the framework of prior self-supervised representation learning algorithms from the computer vision community such as VICReg [26], and utilize a parameterized projector network (2-layer MLP with 0.25 million parameters) that maps encoded visual and non-visual representations independently to a higher-dimensional feature space $\psi_{v_{1,2}}$ and ψ_i respectively, over which the self-supervision objectives are computed. The STERLING objective composed of the multi-modal correlation $\mathcal{L}_{MM}(\psi_{v_{1,2}}, \psi_i)$ and viewpoint-invariance $\mathcal{L}_{VI}(\psi_{v_1}, \psi_{v_2})$ objectives are defined as:

$$
\begin{aligned}
\mathcal{L}_{\text{STERLING}} &= \mathcal{L}_{VI}(\psi_{v_1}, \psi_{v_2}) + \mathcal{L}_{MM}(\psi_{v_{1,2}}, \psi_i) \\
\mathcal{L}_{VI}(\psi_{v_1}, \psi_{v_2}) &= \mathcal{L}_{\text{VICR}eg}(\psi_{v_1}, \psi_{v_2}) \\
\mathcal{L}_{MM}(\psi_{v_{1,2}}, \psi_i) &= [\mathcal{L}_{\text{VICR}eg}(\psi_{v_1}, \psi_i) + \mathcal{L}_{\text{VICR}eg}(\psi_{v_2}, \psi_i)]/2
\end{aligned} \tag{1}
$$

$\mathcal{L}_{\text{VICR}eg}$ is the VICReg loss that is composed of variance-invariance-covariance representation learning objectives, as proposed by Bardes et al. [26]. Given two alternate projected representations Z and Z' of a data sample (in STERLING, Z and Z' are projected representations of the visual and non-visual sensor modalities), the VICReg loss is defined as $\mathcal{L}_{\text{VICR}eg}(Z, Z') = \lambda s(Z, Z') + \mu[v(Z) + v(Z')] + \nu[c(Z) + c(Z')]$. Note that while Bardes et al. [26] use artificial data augmentations, we utilize different sensor modalities as the input for representation learning in this work. This requires careful data curation to extract observations from different paired modalities, as explained in Sec. 3.1. λ, μ, and ν are hyper-parameters and the functions v, s, and c are the variance, invariance, and covariance terms computed on a mini-batch of projected features. We refer the reader to Bardes et al. [26] for more details regarding the individual terms in the loss function. On a mini-batch of data containing paired terrain image patches and IPT observations, we compute the $\mathcal{L}_{\text{STERLING}}$ loss and update parameters of the two encoder networks and the shared projector network together using Adam optimizer.

In all experiments, we use a legged Boston Dynamics Spot robot and collect robot experiences on eight different types of terrain around the university campus that we labeled as `mulch`, `pebble sidewalk`, `cement sidewalk`, `grass`, `bushes`, `marbled rock`, `yellow bricks`, and `red bricks`. We refer the reader to Supplemental Sec. 7 for more details on data collection for this work.

3.3 Preference-Aligned Off-Road Navigation

In this subsection, we describe the problem formulation of terrain-aware navigation and how STERLING features can be utilized within a planner to navigate in a preference-aligned manner. While the main novelty of our work is on learning terrain representations from unconstrained robot experience, we use preference-aligned terrain-aware navigation as a real-world task to evaluate STERLING features.

Preliminaries: We formulate the task of preference-aligned terrain-aware navigation as a local path-planning problem, where the robot operates within a state space $\mathcal{S}$, action space $\mathcal{A}$, and a deterministic transition function $\mathcal{T} : \mathcal{S} \times \mathcal{A} \longrightarrow \mathcal{S}$ in the environment. The state space consists of $s = [x, y, \theta, \phi_v]$, where $[x, y, \theta]$ denote the robot's position in $SE(2)$ space, and ϕ_v denotes the visual features of the terrain at this location. Given a goal location G, the preference-aligned navigation task is to reach this goal while adhering to operator preferences over terrains.

Sampling-based planning: We assume access to a receding horizon sampling-based motion planner with a fixed set of constant-curvature arcs $\{\Gamma_0, \Gamma_1, \ldots, \Gamma_{ns}\}$, $\Gamma \in \mathcal{S}^N$ which solves for the optimal arc $\Gamma^* = \arg\min_{\Gamma}[\mathcal{J}(\Gamma, G)]$, minimizing the objective function $\mathcal{J}(\Gamma, G)$, $\mathcal{J} : (\Gamma, G) \longrightarrow \mathbb{R}^+$. For the task of preference-aligned off-road navigation, we assume the objective function is composed of two components $\mathcal{J}_{geom}(\Gamma, G)$ and $\mathcal{J}_{terrain}(\Gamma)$, and can be defined as $\mathcal{J}(\Gamma, G) = \alpha\mathcal{J}_{geom}(\Gamma, G) + (1 - \alpha)\mathcal{J}_{terrain}(\Gamma)$. $\mathcal{J}_{geom}(\Gamma, G)$ is the geometric cost that deals with progress towards the goal G and avoiding geometric obstacles, whereas $\mathcal{J}_{terrain}(\Gamma)$ is the terrain cost associated with preference-alignment. We utilize the geometric cost as defined in the open-source graph navigation stack.[1] The multiplier $\alpha \in [0, 1]$ trades off relative contributions of the geometric and terrain preference components of the path planning objective. A 1D time-optimal controller translates the sequence of states in the optimal trajectory Γ^* to a sequence of receding horizon actions $(a_0, a_1, \ldots, a_N)$. For a given arc $\Gamma = \{s_0, s_1, \ldots, s_N\}$, such that state s_0 is closest to the robot, the

[1] https://github.com/ut-amrl/graph_navigation

terrain-preference cost can be computed as follows.

$$\mathcal{J}_{terrain}(\Gamma) = \sum_{v_i \sim \Gamma, i=0}^{N} \frac{\gamma^i C(u(f_v(v_i)))}{N+1} \tag{2}$$

The function $f_v(.)$ maps from RGB space of a visual patch of terrain v_i at a specific state s_i, to its visual representation $\phi_v \in \Phi_v$. For instance, f_v can be the visual encoder learned using STERLING, as described in Section 3.2. The utility function $u(.)$ maps the visual representation ϕ_v of a patch of terrain to a real-valued utility of preferences. We follow the utility function formulation of Zucker et al. [29] and assume the terrain preference cost follows a multiplicative formulation such that given a utility value $x \in \mathbb{R}^+$, the traversability cost is $C(x) = e^{-x}$. The discount factor γ weighs the terrain cost proportional to its proximity to the robot. We set γ to 0.8, which we find to work well in practice. More details about the planner is provided in the Supplementary Sec. 8

Learning the preference utility: Following Zucker et al. [29], we learn the utility function $u : \Phi_v \to \mathbb{R}^+$ using human queries. From the predicted terrain features on data samples in our training set, we cluster the terrain representations using k-means with silhouette-score elbow criterion, and sample candidate terrain patches from each cluster, which is presented to the human operator using a GUI. The human operator then provides a full-order ranking of terrain preferences over clusters, which is utilized to learn the utility function $u(.)$, represented by a 2-layer MLP. While recovering absolute cost values from ranked preference orders is an under-constrained problem, we find that this approximation provided by Zucker et al. [29] works well in practice.

4 Experiments

In this section, we describe the experiments performed to evaluate STERLING. Specifically, the experiments presented in this section are tailored to address the following questions:

(Q_1) How effective are STERLING features in comparison to baseline approaches at enabling terrain awareness in off-road navigation?

(Q_2) In a large-scale real-world off-road setting, how reliable are STERLING-features when used for autonomous off-road navigation?

We investigate both questions Q_1 and Q_2 through physical robot experiments on the task of preference-aligned off-road navigation. We perform quantitative evaluations in six different outdoor environments, and then further perform a large-scale qualitative evaluation by autonomously hiking a 3-mile long off-road trail using preference costs learned using STERLING features. To compare various methods, we use the success rate of preference alignment as a metric. If a trajectory followed by any algorithm fails to reach the goal, or at any time traverses over any terrain that is less preferred than any traversed by the operator-demonstrated trajectory, we classify the trial as a failure. We additionally evaluate STERLING against other unsupervised terrain representation learning methods and perform an ablation study on the two STERLING objectives, provided in Supplementary Sec. 9.2.

Baselines: To perform quantitative evaluations for Q_1, we compare STERLING with SE-R [5], RCA [16], GANav [9], geometric-only planning [30], and a fully-supervised baseline. SE-R and RCA perform self-supervised learning from unconstrained robot experience to learn terrain representations and traversability costs respectively, making them relevant baselines for this problem. Since there is no open-source implementation of RCA, we replicate it to the best of our abilities. The geometric-only approach ignores terrain costs ($\mathcal{L}_{terrain}$) and plans with geometric cost ($\mathcal{L}_{geom}$) only, making it a relevant ablation on the cost formulation for preference-aware planning. GANav[2] [9] is a segmentation-based approach trained on the RUGD [6] dataset. We additionally train the fully-supervised baseline in which the terrain cost function is learned end-to-end using supervised learning from linear extrapolation of operator preferences. GANav and the fully-supervised baseline

[2] https://github.com/rayguan97/GANav-offroad

require supervision via terrain labels to learn and hence serve as references for comparison. We normalize the terrain cost predicted by all methods to be between 0 and 1 for a fair comparison.

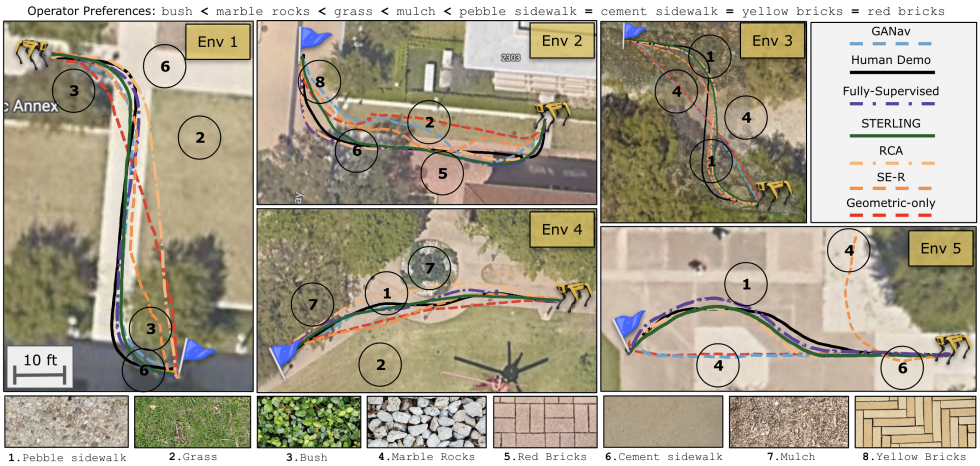

Figure 3: Trajectories traced by different approaches in 5 environments containing 8 different terrains. The operator preferences are shown above. We see that STERLING navigates in an operator-preference aligned manner, by preferring `cement sidewalk`, `red bricks`, `pebble sidewalk`, and `yellow bricks` over `mulch`, `grass`, `marble rocks`, and `bush`, outperforming other baselines and performing on-par with the Fully-Supervised approach.

4.1 Evaluating Terrain-Awareness via Robot Experiments

In this subsection, we report on experiments to investigate the effectiveness of STERLING features in enabling terrain awareness during off-road navigation. We quantitatively compare the performance of STERLING with baselines RCA [16], GANav [9], SE-R [5] and the fully-supervised baseline, on the task of preference-aligned navigation. We identify six environments within the university campus, with eight different terrain types, as shown in Fig. 3. For this study, we use the same data collected on the robot to train RCA, SE-R, fully-supervised baseline, and STERLING, and the operator provides the same rankings for all methods during training.

Fig. 3 shows the operator's (first author) terrain preferences for all Envs. 1 to 5, and the performance of baseline approaches, including an operator-demonstrated trajectory for reference. In all environments, we see that STERLING navigates in a terrain-aware manner while adhering to the operator's preferences. Note that although Fully-Supervised also completes the task successfully, it requires privileged information such as terrain labels during training, whereas STERLING does not require such supervision, and can potentially be used on large datasets containing unlabelled, unconstrained robot experiences. RCA uses inertial spectral features to learn terrain traversability costs and hence does not adhere to operator preference. SE-R does not address viewpoint invariance which is a significant problem in vision-based off-road navigation and hence performs poorly in Envs. 1 and 2. We perform additional experiments in an outdoor environment (Env. 6) to study adherence to operator preferences, detailed in Supplementary Sec. 9.1.

Table 1 shows the success rate of preference alignment for all approaches in all environments, over five different trials. STERLING outperforms other self-supervised baselines and performs on par with the fully-supervised approach. In summary, the physical experiments conducted in six environments quantitatively demonstrate the effectiveness of STERLING features in enabling terrain awareness during off-road navigation.

4.2 Large-Scale Qualitative Evaluation

In this subsection, we address Q_2 by reporting on a large-scale evaluation of STERLING by autonomously hiking a 3-mile-long off-road trail.

Figure 4: A large-scale qualitative evaluation of STERLING on a 3-mile outdoor trail. STERLING features successfully complete the trail with only two manual interventions (shown in red).

We train STERLING using unconstrained robot experience collected within the university campus and train the preference utility function using operator-provided preferences: `marble rocks < grass < dirt = cement`. The task is to navigate the trail without a global map while adhering to operator preferences at all times. Since we do not use a global map, visual terrain awareness is necessary to navigate within the trail and avoid catastrophic events such as falling into the river next to the trail. We set a moving goal of six meters in front of the robot, updated every second. While the robot navigates autonomously, the operator walks behind the robot and takes manual control only to correct the robot's path during forks, or to yield to incoming pedestrians and pets. The attached supplementary video shows the robot navigating the trail successfully while avoiding less preferred terrains. The robot needed two manual interventions while traversing along the trail. Fig. 4 shows the 3-mile trajectory traced by the robot and the two failure cases that required manual intervention. This large-scale qualitative experiment addresses Q_2 by demonstrating the reliability of STERLING during real-world off-road deployments.

Approach	**Environment**						
	1	**2**	**3**	**4**	**5**	**6 (a)**	**6 (b)**
Geometric-only	0/5	0/5	0/5	0/5	0/5	0/5	5/5
RCA[16]	2/5	4/5	2/5	0/5	1/5	5/5	0/5
GANav[9]	5/5	0/5	0/5	5/5	0/5	4/5	5/5
SE-R[5]	1/5	0/5	5/5	1/5	3/5	5/5	4/5
Fully-Supervised	5/5	5/5	5/5	5/5	5/5	5/5	5/5
STERLING (Ours)	5/5	5/5	5/5	5/5	5/5	5/5	5/5

Table 1: Success rates of different algorithms on the task of preference-aligned off-road navigation

5 Limitations and Future Work

STERLING requires traversing over terrains in order to learn representations, which may be unsafe in certain situations. Uncertainty-aware safe exploration for data collection is a promising direction for future work. Extending STERLING features to work on unstructured non-flat environments such as stairs [31] and boulders [18] is another promising direction for future work.

6 Conclusion

In this paper, we introduce *Self-supervised TErrain Representation LearnING* (STERLING), a novel framework for learning terrain representations from easy-to-collect, unconstrained (e.g., non-expert), and unlabelled robot experience. STERLING utilizes non-contrastive representation learning through viewpoint invariance and multi-modal correlation self-supervision objectives to learn relevant terrain representations for visual navigation. We show how features learned through STERLING can be utilized to learn operator preferences over terrains and integrated within a planner for preference-aligned navigation. We evaluate STERLING against state-of-the-art alternatives on the task of preference-aligned visual navigation on a Spot robot and find that STERLING outperforms other methods and performs on par with a fully-supervised baseline. We additionally perform a qualitative large-scale experiment by successfully hiking a 3-mile-long trail using STERLING, demonstrating its robustness to off-road conditions in the real world.

References

[1] T. Wang, B. Chen, Z. Zhang, H. Li, and M. Zhang. Applications of machine vision in agricultural robot navigation: A review. *Computers and Electronics in Agriculture*, 198:107085, 2022.

[2] Agility robotics launches next generation of digit: World's first human-centric, multi-purpose robot made for logistics work, 2023. URL https://agilityrobotics.com/news/2022/future-robotics-l3mjh.

[3] X. Xiao, J. Dufek, and R. R. Murphy. Autonomous visual assistance for robot operations using a tethered uav, 2019.

[4] H. Karnan, K. S. Sikand, P. Atreya, S. Rabiee, X. Xiao, G. Warnell, P. Stone, and J. Biswas. Vi-ikd: High-speed accurate off-road navigation using learned visual-inertial inverse kinodynamics. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3294–3301. IEEE, 2022.

[5] J. Zürn, W. Burgard, and A. Valada. Self-supervised visual terrain classification from unsupervised acoustic feature learning. *IEEE Transactions on Robotics*, 37(2):466–481, 2020.

[6] M. Wigness, S. Eum, J. G. Rogers, D. Han, and H. Kwon. A rugd dataset for autonomous navigation and visual perception in unstructured outdoor environments. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5000–5007. IEEE, 2019.

[7] P. Jiang, P. Osteen, M. Wigness, and S. Saripalli. Rellis-3d dataset: Data, benchmarks and analysis, 2020.

[8] S. Sharma, L. Dabbiru, T. Hannis, G. Mason, D. W. Carruth, M. Doude, C. Goodin, C. Hudson, S. Ozier, J. E. Ball, et al. Cat: Cavs traversability dataset for off-road autonomous driving. *IEEE Access*, 10:24759–24768, 2022.

[9] T. Guan, D. Kothandaraman, R. Chandra, A. J. Sathyamoorthy, K. Weerakoon, and D. Manocha. Ga-nav: Efficient terrain segmentation for robot navigation in unstructured outdoor environments. *IEEE Robotics and Automation Letters*, 7(3):8138–8145, 2022.

[10] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba. End to end learning for self-driving cars, 2016.

[11] Y. LeCun, U. Muller, J. Ben, E. Cosatto, and B. Flepp. Off-road obstacle avoidance through end-to-end learning. In *Proceedings of the 18th International Conference on Neural Information Processing Systems*, NIPS'05, page 739–746, Cambridge, MA, USA, 2005. MIT Press.

[12] M. Wulfmeier, P. Ondruska, and I. Posner. Maximum entropy deep inverse reinforcement learning. *arXiv preprint arXiv:1507.04888*, 2015.

[13] Y. Pan, C.-A. Cheng, K. Saigol, K. Lee, X. Yan, E. A. Theodorou, and B. Boots. Imitation learning for agile autonomous driving. *The International Journal of Robotics Research*, 39 (2-3):286–302, 2020.

[14] G. Kahn, P. Abbeel, and S. Levine. Land: Learning to navigate from disengagements. *IEEE Robotics and Automation Letters*, 6(2):1872–1879, 2021.

[15] G. Kahn, P. Abbeel, and S. Levine. Badgr: An autonomous self-supervised learning-based navigation system. *IEEE Robotics and Automation Letters*, 6(2):1312–1319, 2021.

[16] X. Yao, J. Zhang, and J. Oh. Rca: Ride comfort-aware visual navigation via self-supervised learning. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7847–7852. IEEE, 2022.

[17] X. Meng, N. Hatch, A. Lambert, A. Li, N. Wagener, M. Schmittle, J. Lee, W. Yuan, Z. Chen, S. Deng, G. Okopal, D. Fox, B. Boots, and A. Shaban. Terrainnet: Visual modeling of complex terrain for high-speed, off-road navigation, 2023.

[18] A. Datar, C. Pan, M. Nazeri, and X. Xiao. Toward wheeled mobility on vertically challenging terrain: Platforms, datasets, and algorithms, 2023.

[19] T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter. Learning robust perceptive locomotion for quadrupedal robots in the wild. *Science Robotics*, 7(62):eabk2822, 2022.

[20] K. S. Sikand, S. Rabiee, A. Uccello, X. Xiao, G. Warnell, and J. Biswas. Visual representation learning for preference-aware path planning. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 11303–11309. IEEE, 2022.

[21] Y. LeCun, U. Muller, J. Ben, E. Cosatto, and B. Flepp. Off-road obstacle avoidance through end-to-end learning. In *Proceedings of the 18th International Conference on Neural Information Processing Systems*, NIPS'05, page 739–746, Cambridge, MA, USA, 2005. MIT Press.

[22] X. Xiao, B. Liu, G. Warnell, and P. Stone. Motion planning and control for mobile robot navigation using machine learning: a survey, 2020.

[23] F. Schilling, X. Chen, J. Folkesson, and P. Jensfelt. Geometric and visual terrain classification for autonomous mobile navigation. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2678–2684. IEEE, 2017.

[24] M. Wigness, J. G. Rogers, and L. E. Navarro-Serment. Robot navigation from human demonstration: Learning control behaviors. In *2018 IEEE international conference on robotics and automation (ICRA)*, pages 1150–1157. IEEE, 2018.

[25] C. A. Brooks and K. D. Iagnemma. Self-supervised classification for planetary rover terrain sensing. In *2007 IEEE aerospace conference*, pages 1–9. IEEE, 2007.

[26] A. Bardes, J. Ponce, and Y. LeCun. Vicreg: Variance-invariance-covariance regularization for self-supervised learning. *arXiv preprint arXiv:2105.04906*, 2021.

[27] A. J. Sathyamoorthy, K. Weerakoon, T. Guan, J. Liang, and D. Manocha. Terrapn: Unstructured terrain navigation using online self-supervised learning, 2022.

[28] L. Wellhausen, A. Dosovitskiy, R. Ranftl, K. Walas, C. Cadena, and M. Hutter. Where should i walk? predicting terrain properties from images via self-supervised learning. *IEEE Robotics and Automation Letters*, 4(2):1509–1516, 2019.

[29] M. Zucker, N. Ratliff, M. Stolle, J. Chestnutt, J. A. Bagnell, C. G. Atkeson, and J. Kuffner. Optimization and learning for rough terrain legged locomotion. *The International Journal of Robotics Research*, 30(2):175–191, 2011.

[30] J. Biswas. Amrl autonomy stack. https://github.com/ut-amrl/graph_navigation, 2013.

[31] T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter. Learning robust perceptive locomotion for quadrupedal robots in the wild. *Science Robotics*, 7(62), jan 2022. doi:10.1126/scirobotics.abk2822.