Literature Review

A Literary Review of Comparative Learning Strategies for Off-Road Robot Traversability Estimation

Mrinank Sivakumar [1]

[1] Master of AI & ML, University of Adelaide, Australia; mrinank.sivakumar@student.adelaide.edu.au

1. Introduction

Autonomous ground robots that are deployed in off-road environments face the unique challenge of determining which parts of the terrain around them are safe to traverse. This is what is known as traversability estimation and is vital to preventing navigation failure across a variety of contexts from agriculture to extraterrestrial exploration. The core challenge of traversability estimation comes from unstructured and/or un-mapped terrain that could present features such as mud, vegetation, snow, and uneven surfaces to name a few. Recent research has turned to using machine learning to solve this problem, as opposed to the traditional rule-based perception. Several learning strategies have emerged through this research including supervised models that are trained on large, annotated datasets, semi-supervised models that use partially labelled or pseudo-labelled data, self-supervised models trained on robot motion outcomes, and hybrid models that run off human-in-the-loop feedback or vision-language reasoning. Unfortunately, these methods are generally evaluated separately from each other, using different datasets and evaluation metrics. This literature review aims to examine a selection of papers proposing differing methods, and exploring the trade-offs between accuracy, adaptability, and data efficiency amongst them. By identifying key themes in domain shift, label scarcity, and continual learning we can effectively simplify the direct comparison and adoption of these methods. This review will serve as a foundation of knowledge for our research project, that aims to evaluate and compare these various methodologies under shared conditions and ultimately contribute to improving reliable, vision-based autonomy.

2. Literature Search Process

The literature search process for this review began with a broad, exploratory search using Google Scholar, arXiv, and IEEE Xplore. Key search phrases such as "traversability estimation", "terrain classification", "robot vision", "positive-unlabeled learning", and "inverse reinforcement learning" were used in various combinations. In order to ensure our work is relevant, only papers published in the last decade were considered, with the exception of frequently referenced, foundational works. The search for papers was otherwise indiscriminate, so as to gain a broader understanding of the topic of traversability estimation. Papers that were later selected for review prioritized research that tackled the trade-offs between supervision level, data annotation costs, and generalization across environments. The filtering of papers also considered practicality towards the research project. This means papers had to provide access to code or well-described models, as well as use datasets that were publicly available or reproducible through simulation. From a folder of more than 20 academic research papers, nine were chosen to examine in this review for their representation of different methodologies, depth of detail, and availability of supporting resources.

3. Academic Literature Review

Current traversability estimation is moving in the direction of using data-driven machine learning approaches over the heuristic rule-based systems of the past. These methods are able to make better use of the increasing amounts of sensory robotic data we are able to collect and train traversability estimation models that generalize across foreign environments. In order to better understand these methodologies this review explores representative works and evaluates each for its methodology, efficiency, adaptability, and relevance.

Supervised Learning

The first paper from Palazzo et al. (2020) outlines a classical approach to vision-based traversability estimation. This approach, built on a deep segmentation backbone, predicts traversability distances over multiple directions from a single RGB image. An original safety-preserving loss function enhances the model by penalizing overestimations more heavily than underestimations to avoid collisions. In this way, the model maximizes its interpretability and modularity, allowing its outputs to be used directly by navigation systems.

When examining the proposed architecture, it is important to note the incorporation of unsupervised domain adaptation (UDA) through a gradient reversal layer (GRL). The model is pretrained on a labeled dataset and then fine-tuned after using unlabeled target domain data. This might seem like a semi-supervised approach, but it is just a unique combination of supervised learning and domain adaptation.

The approach seemed to perform well in generalizing to unseen environments. However, it still relies on manual annotations and is limited to using monocular RGB input. The custom dataset used in the paper is not available to the public, which hinders the reproducibility of this method. Despite all this, the method proposed by Palazzo et al. (2020) acts as a strong baseline model and provides a good example of safety-aware vision design for robotics.

Semi-Supervised Learning

GONet is a semi-supervised GAN-based model for conducting traversability estimation with minimal labeled data. This method, proposed by Hirose et al. (2018), is mainly trained on positive traversable images, with only a small set of negatives used for fine-tuning the system. The generative adversarial network will learn what safe terrain looks like and will subsequently flag anything that looks different, during unsupervised use, as potentially untraversable.

This specific approach is able to achieve real-time inference and operates on 360-degree fisheye imagery of outdoor scenes. GONet can bridge the gap between generalization and label efficiency, which makes it optimal for robots potentially operating in foreign terrains. In addition to this, GONet's accessibility is enhanced by the inclusion of public indoor datasets that are newly released and its compatibility with simple sensors.

Unfortunately, GONet's weakness comes from its binary classification of traversability, lack of uncertainty gradient, and lack of directional context. Nonetheless, GONet is still a relevant example of semi-supervised traversability estimation.

Self-Supervised Learning

The first model we look at in this group is from a paper called Follow the Footprints, proposed by Jeon et al. (2024). This is a self-supervised approach that labels terrain based on robot motion outcomes. This means if the robot successfully traverses any given path, the corresponding image frames are marked as safe, while failed traversals are marked as unsafe. This method removes the need for human annotation entirely and allows for very efficient label generation in various conditions. This model was evaluated on RELLIS-3D and ORFD and proved to generalize well across several different domains (lighting conditions, weather, and surface data).

The second model we take a look at is proposed by Seo et al. (2023) and is called ScaTE (Scalable Traversability Estimation). This method uses contrastive self-supervised learning on clustered trajectory outcomes. What this means is that the robot's traversed paths are grouped as success or failure, and a contrastive loss is applied to help the model record the differences. This method also uses Positive-Unlabeled (PU) learning to filter the confident training signals and reduce false supervision. In this way, ScaTE is able to achieve a strong temporal consistency and spatial generalization after being validated on RELLIS-3D.

Both of the models we looked at learn directly from sensor feedback, which makes them much more attractive for long-term and/or minimal interference deployments. The downsides being, their performance depends on the quality of the robot's trajectory monitoring, where inaccurate odometry can propagate label errors. Not to mention they lack the safety guarantees of fully supervised models. However, they are both much more scalable and adaptive.

Online and Continual Learning

One of our papers talks about ARTE (Adaptive Robot Traversability Estimation), a real-world online learning system, developed by Yoon et al. (2024), that continually updates traversability predictions as the robot traverses through unfamiliar terrain. It has an episodic memory to record the terrain, a resilience unit to prevent forgetting, and a self-supervised loss to aid in adaptation. The ARTE system was tested in various outdoor locations after being deployed on a Clearpath Husky robot. From these tests we see that the strength of ARTE lies in its real-time adaptability and its absence of catastrophic forgetting. These essential features make it perfect for long-duration missions in evolving environments.

This concept was built on further by Mattamala et al. (2024) in their Wild Visual Navigation (WVN) system. WVN uses vision transformers and proprioceptive feedback to quickly learn about the terrain. A robot with this system can fine-tune a segmentation model using live sensor data within only five minutes of being deployed. WVN also contains modules that detect anomalies and conduct temporal smoothing to adapt faster without initial training. While the datasets used in this paper are all proprietary, the described system is both modular and generalizable.

Despite having the same challenges in regards to model stability and feedback alignment as the self-supervised methods, the ARTE and WVN systems show the feasibility of real-time traversability estimation. Approaches like these are constantly pushing the boundary of on-the-fly learning in unseen environments.

Hybrid and Vision-Language Model

One research paper proposed an interesting hybrid system that combined vision-language models (VLMs) with human-in-the-loop verification. This system, called AnyTraverse and proposed by Sahu et al. (2025), uses CLIP to capture terrain images and detect unfamiliar features. When the model has a high uncertainty, it prompts a human operator with natural language queries, the response for which is used to update the model's internal representation.

AnyTraverse was tested on RELLIS-3D, Freiburg Forest, and RUGD datasets, and showed strong performance in zero-shot and high-risk scenarios. The VLM component lets the model generalize across different types of scenes, while the human guidance fallback ensures a safety protocol.

Unfortunately, the model is very computationally intensive and sensitive to the quality of prompt engineering. The model's reliance on human operators also presents as a potential limitation in autonomous missions. However, these issues aside, AnyTraverse represents the future where visual learning and semantic reasoning combine to form interpretable, collaborative vision systems.

Survey and Analysis

Sevastopoulos and Konstantopoulos (2022) have put together a comprehensive survey of traversability estimation methods for mobile robots. They have organized these methods by learning paradigms, sensor modality, and evaluation framework. The authors bring to light the critical issue in this field, that is a lack of consistent benchmarking across the various models. They also highlight the increasing focus being placed on multimodal sensor fusion and domain adaptation.

This survey ultimately helps us contextualize the papers in this literature review and affirms our need for a standard comparative analysis. Through our research project we will seek to evaluate the different learning strategies under shared datasets and parameters.

4. Existing Project Review

When looking at existing projects, the survey conducted by Sevastopoulos and Konstantopoulos (2022) is likely the most comprehensive prior effort related to our research. Their work records the state of traversability estimation across methodologies, sensor modalities, and application settings. It then categorizes the methods into supervised, unsupervised, and hybrid approaches and highlights the importance of label-efficient learning as well as generalization across terrains. Most importantly, the survey points out the lack of a standard benchmark and how that leads to variability in evaluation metrics, which in turn makes it difficult to compare models across datasets. This observation is what forms the basis of our research. Despite giving us valuable taxonomic and theoretical insights, the survey does not conduct its own experimental evaluations or assess the reproducibility of its reviewed models. Our research project builds on the foundation of the survey's meta-analysis with hands-on benchmarking.

In addition to the survey, many of the papers we reviewed offer public repositories or datasets that describe the practical implementation. Follow the Footprints for example, has its model and training pipeline published on GitHub to enable replication and public validation (Jeon et al. 2024). Many of these implementations also include wrappers for datasets like RELLIS-3D. Having access to these projects facilitates the direct experimentation and comparison that we are aiming to conduct with our research.

As such, these papers have helped build the methodological foundation for our research project. The current research landscape in this field comes across as fragmented through the papers we have reviewed. Our project will ideally bridge these gaps by evaluating methods across consistent datasets and metrics, ultimately validating or challenging the claims made in literature for traversability estimation, under practical constraints. This approach will hopefully extend beyond simple theoretical comparison and will provide insights for future traversability estimation method development.

5. Conclusion

Throughout this literature review we have examined the various machine learning methodologies used in robot traversability estimation. Supervised models, like the one proposed by Palazzo et al. (2020), show us a strong performance in exchange for costly annotation and limited ability to generalize. Semi-supervised methods, such as GONet, have less of a reliance on labels thanks to adversarial learning, while self-supervised strategies, like ScaTE and Follow the Footprints, use robot motion outcomes to generate labels from experience (Hirose et al. 2018, Seo et al. 2023, Jeon et al. 2024). ARTE and Wild Visual Navigation showed us how online learning systems were pushing those concepts further and allowing robots to adapt continually in real-time without intervention (Yoon et al. 2024, Frey et al. 2023). On the other hand, we also looked at hybrid methodologies that are trying to improve the safety and generalization of models in novel terrain, through the use of interactive supervision and language reasoning, with one such method being AnyTraverse (Sahu et al. 2025).

A common theme across all these methodologies is the growing focus on labeling efficiency, cross-domain generalization, and deployment readiness. All the papers reviewed thus far show us a point of convergence between classical robotics and modern machine learning, where data-driven models are expected to meet the strict requirements of real-world autonomy. However, our understanding of which learning strategies are most effective is limited by the lack of unified benchmarks or cross-domain comparisons in this field.

Our proposed research project aims to address this exact gap by conducting a comparative evaluation of learning methodologies for traversability estimation. It will be done using consistent datasets and reproducible training conditions, which will inform future system designs about the trade-offs between label cost, adaptability, and performance. Ultimately, this research aims to contribute towards the development of more scalable, generalizable, and autonomous robotic vision systems for unstructured outdoor environments.

References

1. Frey, J., Patel, M., Atha, D., Nubert, J., Fan, D., Agha, A., Padgett, C., Spieler, P., Hutter, M. and Khattak, S. (2024) 'RoadRunner: Learning traversability estimation for autonomous off-road driving', *arXiv preprint arXiv:2402.19341*.

2. Hirose, N., Sadeghian, A., Vázquez, M., Goebel, P. and Savarese, S. (2018) 'GONet: A semi-supervised deep learning approach for traversability estimation', *arXiv preprint arXiv:1803.03254*.

3. Jeon, Y., Son, E.I. and Seo, S.W. (2024) 'Follow the Footprints: Self-supervised traversability estimation for off-road vehicle navigation based on geometric and visual cues', *arXiv preprint arXiv:2402.15363*.

4. Mattamala, M., Frey, J., Libera, P., Chebrolu, N., Martius, G., Cadena, C., Hutter, M. and Fallon, M. (2024) 'Wild Visual Navigation: Fast Traversability Learning via Pre-Trained Models and Online Self-Supervision', *arXiv preprint arXiv:2404.07110*.

5. Palazzo, S., Guastella, D.C., Cantelli, L., Spadaro, P., Rundo, F., Muscato, G., Giordano, D. and Spampinato, C. (2020) 'Domain adaptation for outdoor robot traversability estimation from RGB data with safety-preserving loss', *arXiv preprint arXiv:2009.07565*.

6. Sahu, S., Singh, A., Nambiar, K., Saripalli, S. and Sujit, P.B. (2025) 'AnyTraverse: An off-road traversability framework with VLM and human operator in the loop', *arXiv preprint arXiv:2506.16826*.

7. Sevastopoulos, C. and Konstantopoulos, S. (2022) 'A survey of traversability estimation for mobile robots', *arXiv preprint arXiv:2204.10883*.

8. Seo, J., Kim, T., Kwak, K., Min, J. and Shim, I. (2023) 'ScaTE: A scalable framework for self-supervised traversability estimation in unstructured environments', IEEE Robotics and Automation Letters, 8(2), pp. 888–895. doi:10.1109/LRA.2023.3234768.

9. Yoon, H.-S., Hwang, J.-H., Kim, C., Son, E.I., Yoo, S.-W. and Seo, S.-W. (2024) 'Adaptive robot traversability estimation based on self-supervised online continual learning in unstructured environments', IEEE Robotics and Automation Letters, 9(6), pp. 4902–4909. doi:10.1109/LRA.2024.3386451.