*Research Proposal*

# Comparative Evaluation of Supervised, Semi-Supervised, and Self-Supervised Methods for Visual Traversability Estimation in Off-Road Environments

**Mrinank Sivakumar** [1]

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

**Abstract**

Robotic ground vehicles that operate in off-road environments are dependent on traversability estimation to ensure safe and autonomous navigation. Several machine learning models have been created to address this problem, utilizing supervised, semi-supervised, and self-supervised methodologies. Comparisons between these methods in consistent settings, however, are lacking. This research project aims to implement and evaluate all three methodologies of traversability estimation on publicly available off-road datasets. The study will benchmark the performance, data requirements, and generalization across terrains of each model, to gain insight into the effectiveness of each strategy.

**Keywords:** traversability estimation; computer vision; supervised learning; semi-supervised learning; self-supervised learning; domain adaptation; robotics

## 1. Introduction and Research Problem

Autonomous ground vehicles must accurately estimate whether their surrounding terrain is traversable or not, especially in off-road situations. Traditionally, supervised learning solutions that require large quantities of human-labelled datasets are used. However, these are costly to make and unreliable when the models face new, unseen environments.

Recently, research has explored several semi-supervised and self-supervised machine learning methods, with the goal of reducing/removing the need for human annotation. Unfortunately, most of these studies test their methods separately from one another with differing datasets and experiment setups. This makes it difficult to effectively compare the methodologies in predicting traversability under real-world conditions or domain shifts.

This gap in comparative evaluation limits our ability to isolate the most effective and practical solutions for robots in a variety of situations. This research aims to address that problem by conducting a systematic comparison of each methodology of traversability estimation (supervised, semi-supervised, and self-supervised), with consistent datasets and evaluation criteria.

## 2. Datasets

The following public datasets will be used as the benchmark for all methods tested:

- RELLIS-3D: A Multi-modal Dataset for Off-Road Robotics (Palanisamy et al. 2020)
  - Outdoor images with varied terrain types (dirt, mud, grass, etc.)
  - High-resolution and RGB
  - Semantic segmentation labels
- Freiburg Forest Dataset (Valada et al. 2016)
  - Forest images across different seasons
  - RGB
  - Semantic labels
- RUGD Dataset (Oliveira et al. 2021)
  - Urban and off-road ground-level images
  - Pixel-level traversability annotations
- CTU Traversability Datasets (Černý et al. 2020)
  - Underground and outdoor environment images
  - Traversability annotations

These datasets provide an excellent foundation for testing model robustness, environmental variability, domain shifts, and cross-domain generalization.

## 3. Task Selection and Benefit

Visual traversability estimation (using RGB images) is the selected task.
Benefits of this investigation:

- Provides a practical impact by allowing robotics users to effectively choose the methods that balance accuracy with training efficiency
- Provides a scientific contribution by filling a gap in literature of comparative evaluation between methodologies under identical conditions
- Provides readiness by informing of safer and more reliable deployment configurations of autonomous robots in a variety of environmental contexts

## 4. Research Questions

This research aims to answer the following:

1. Performance Comparison:
   How do supervised, semi-supervised, and self-supervised methodologies compare in predicting traversability, based on visual data of off-road environments?

2. Data Efficiency:
   How much labelled data does each method require to achieve satisfactory performance?

3. Domain Generalization:
   Do models trained on one dataset generalize effectively to other datasets with different terrain/environments?

4. Claim Verification:
   Do the performance results reported in the original research papers hold true when methods are implemented under consistent experimental settings?

## 5. Proposed Methodology

1. **Literature Review and Method Selection**
   Three papers have been chosen that represent each methodology:
   - **Supervised**
     - *Domain Adaptation for Outdoor Robot Traversability Estimation from RGB data with Safety-Preserving Loss* (Palazzo et al. 2020)
       - Predicts traversable distance in different directions based on a single RGB image
       - Introduces a safety-preserving loss to avoid overestimations
   - **Semi-Supervised**
     - *Learning Off-Road Terrain Traversability with Self-Supervisions Only* (Seo et al. 2023)
       - Uses Positive-Unlabeled (PU) learning
       - Generates pseudo-labels based on vehicle trajectory outcomes
   - **Self-Supervised**
     - *RoadRunner - Learning Traversability Estimation for Autonomous Off-road Driving* (Frey et al. 2024)
       - Makes use of hindsight supervision to connect future navigation outcomes with current observations
       - Can be adapted to RGB-only inputs for more consistent comparisons
2. **Data Preparation**
   - Review and preprocess selected datasets
   - Standardize image properties for all experiments
   - Define consistent dataset splits
3. **Model Implementation**
   - Implement the models and pipelines described by the chosen papers
   - Adapt official code to standard datasets
   - Document the details of ach implementation to ensure reproducibility
4. **Evaluation Metrics**
   - Prediction performance
     - Accuracy, Precision, Recall, F1-score, errors, etc.
   - Robustness under domain shifts
     - Training on one dataset and testing on another dataset
   - Data efficiency
     - Evaluating how the performance scales with different amounts of labelled data
   - Reproducibility
     - Comparing the results achieved to the results reported in original papers
5. **Analysis**
   - Compare our findings for each methodology
   - Assess the trade-offs for each methodology
   - Identify limitations or unexpected outcomes for each method's implementation

## 6. Ethics Statement

This research project will exclusively use publicly available datasets that are collected in outdoor environments. There are no human subjects involved, and no personally identifiable information exists in the data that will be used to train each model. All work done will comply with the University of Adelaide's research integrity policy (University of Adelaide 2018).

## 7. Timeline

| Week | Activity |
| --- | --- |
| 1 - 5 | Research exploration and proposal |
| 6 - 8 | Literature review and dataset curation |
| 9 - 11 | Objective method comparison and preparation for implementation |
| 12 | Analysis presentation |
| 13+ (Research Project B) | Method implementation, evaluation and analysis, research report |

## 8. Expected Outcome

This project expects to deliver:

- A comparative evaluation of each, supervised, semi-supervised, and self-supervised, traversability estimation method
- Verification of the published claims for each methodology tested under controlled conditions
- Practical recommendations for method selection based on:
    - Performance
    - Cross-domain generalization
    - Data annotation cost

The results are expected to fill a current gap in robotics literature and provide a perspective on the strengths and weaknesses of each machine learning methodology in the field of traversability estimation.'

## References

1. Černý, T, Kocmanová, P & Přeučil, L 2020, Traversability dataset, Czech Technical University in Prague, viewed 7 June 2025, http://subtdata.felk.cvut.cz/robingas/data/.

2. Frey, L, Garcia, E, Drozdov, V, Vora, S, Mohamed, R & Scherer, S 2024, 'RoadRunner: Hindsight Supervision for Off-Road Traversability Estimation', arXiv preprint, viewed 7 June 2025, https://arxiv.org/abs/2402.19341.

3. Palanisamy, P, Babu, R, Lee, Y, Hensley, J, Bian, A & Jo, K 2020, RELLIS-3D: A dataset for multimodal perception in off-road driving, Texas A&M University, viewed 7 June 2025, https://www.unmannedlab.org/research/RELLIS-3D.

4. Palazzo, S, Spampinato, C, Sillitti, S & Brizzi, V 2020, 'Domain Adaptation for Outdoor Robot Traversability Estimation from RGB data with Safety-Preserving Loss', arXiv preprint, viewed 7 June 2025, https://arxiv.org/abs/2009.07565.

5. Seo, J, Sim, S & Shim, I 2023, 'Learning Off-Road Terrain Traversability with Self-Supervisions Only', arXiv preprint, viewed 7 June 2025, https://arxiv.org/abs/2305.18896.

6. University of Adelaide 2018, Research integrity policy, viewed 7 June 2025, https://www.adelaide.edu.au/staff/research/ethics-compliance-integrity/research-integrity/.

7. Wigness, M, Eum, S, Rogers, JG, Han, D & Kwon, H 2019, 'A RUGD dataset for autonomous navigation and visual perception in unstructured outdoor environments', in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China.