Problem Statement

Design mapless adaptive robots that can operate safely in unknown environments while detecting and mitigating adversarial attacks during operation to ensure human safety and trust.

Hypothesis

Incorporating adversarial input detection into reinforcement learning training will result in lower collision rates during mapless navigation.

Research Questions

1. What is the most effective way to reduce the collision rate in mapless navigation without interfering with exploration and adaptability?
2. How can mitigating adversarial attacks during reinforcement learning reduce the risk of external manipulation during operation?

Literature Review 1

Title: A Deep Safe Reinforcement Learning Approach for Mapless Navigation

Authors: Shaohua Lv, Yanjie Li, Qi Liu, Jianqi Gao, Xizheng Pang & Meiling Chen

Publication Year: 2021

Key Ideas

The paper focuses on using deep reinforcement learning (DRL), specifically for mapless robot navigation in unknown environments. It discusses how mapped robots use Simultaneous Localization and Mapping (SLAM), which becomes computationally expensive and time-consuming, making it difficult to use robots in unknown environments due to the constant need to rebuild the map. DRL allows robots to learn through trial and error, navigating through environments and learning through internal representations. However, while it seems more beneficial, there is still a safety risk where real robots may collide during training. The paper discusses ways to address safety concerns using Constrained Policy Optimization (CPO) and an Actor-Critic Safety (ACS) architecture to improve mapless navigation while avoiding collisions with static or dynamic objects.

Contributions

The paper introduces a safe RL framework that focuses on modifying punishments into negative rewards to improve performance in collision avoidance. By adjusting the reward and cost structure, it helps reduce risk-taking and risk-seeking behaviors during navigation training. The models are trained in a static environment and then transferred to dynamic settings with

movement. The model was able to achieve high success rates and fewer collisions during simulation.

Limitations

Though the results were promising, the models and training were conducted in simulations like Gazebo, not in real-world environments. The models are limited to lidar inputs and do not incorporate sensor fusion. While the paper discusses safety during RL training, it does not address adversarial attacks.

Extend/Improvement

For my project, I will build on the idea of mapless robots navigating unknown environments, but also simulate adversarial inputs during training. I will extend the article's focus on safe navigation by applying it to real-world environments with sensor fusion.

Literature Review 2

Title: Collaborative Assembly in Hybrid Manufacturing Cells: An Integrated Framework for Human-Robot Interaction

Authors: Behzad Sadrfaridpour and Yue Wang

Publication Year: 2017

Key Ideas

The paper focuses on human-robot collaboration (HRC) that helps improve teamwork, specifically in manufacturing assembly lines. The framework combines physical human-robot interaction (pHRI) and social human-robot interaction (sHRI) to make the partnership more trustworthy, efficient, and usable for repeatable tasks. It discusses how trust is the main focus to have a successful collaboration, as it can affect coordination, workload amount, and performance. It proposes an intelligent system that helps the robot adjust its movement based on the human pace and displays emotional facial expressions to signal the robot's state during interaction.

Contributions

The paper discusses integrating cognitive and emotional aspects of HRI into motion control systems. This allows the robot to detect and adapt to human movement using sensory devices and a trust model with facial expressions for trust feedback. This approach improved communication between human and robot, reducing the human's physical burden. In the

real-world training, there was a 44% decrease in workload and a 60% increase in robot usability when trust and emotion were combined in the model.

Limitations

While it did focus on real-world environments, it was in a structured environment where the action was repetitive and predictable. Though the emotion showed some increase in collaboration, it was limited to only three emotions and wouldn't be able to work for more complex interactions and situations. The robot was also static, unable to adapt or evolve during its learning.

Extend/Improvement

The overall focus of my paper is on designing adaptive robotic systems that will operate in dynamic and unpredictable environments. The article demonstrated the benefits of emotional and trust-aware interaction, and I will extend those ideas into learning systems that allow robots to adjust trust metrics and behavior policies during changing conditions.

Literature Review 3

Title: Securing Cyber-physical Robotic Systems for Enhanced Data Security and Real-time Threat Mitigation

Authors: Akashdeep Bhardwaj, Salil Bharany, Ateeq Ur Rehman, Ghanshyam G. Tejani and Seada Hussen

Publication Year: 2025

Key Ideas

This article focuses on securing cyber-physical robotic systems (CPRS) by looking at the vulnerabilities in both computational and physical components. It proposes a smart cybersecurity framework that compares real-time sensor readings to detect abnormalities and adversarial attacks. It also introduces a tree-based algorithm to simulate known and unknown vulnerabilities to differentiate classes of intruders that can compromise the systems.

Contributions

It provides a layered taxonomy of robotic cybersecurity challenges using the Confidentiality, Integrity, Availability, Authentication, and Privacy (CIAAP) model, detailing the vulnerabilities and discussing the different failures. It also demonstrates real-time simulations of live attacks on robotic control servers that affect robotic safety and trust. It proves that using

physical nodes and sensors allows the system to detect intrusions and mitigate them more effectively.

Limitations

It focuses on detection and threshold-based defense methods rather than an adaptive learning approach. During the simulations, it assumes zero-day vulnerabilities that are known by the attacker but not publicly disclosed, limiting it to predefined attacks rather than evolving, unpredictable ones.

Extend/Improvement

My project will focus on machine learning models that can mitigate unknown adversarial attacks during operation. The article categorized robotic vulnerabilities and used a rule-based detection system with sensors and attack tree simulations, which I will build on by integrating learning-based detection systems for unknown and evolving threats.

High-Level Proposed Approach

AI Models

- Reinforcement Learning (RL) + Proximal Policy Optimization (PPO)
    - Training in a mappless environment
- Imitation Learning (IL)
    - Learning from demonstrations
- Autoencoders
    - Adversarial input detection during training
- LLaVA (LLM + Vision)
    - Explaining robot behavior during failure cases or abnormal navigation patterns

Tools and Libraries

- Gymnasium + RoboTHOR + AI2-THOR
    - robot simulation environment for safe policy testing
- PyTorch
    - Flexible deep learning framework
- Stable-Baselines3
    - Pre-built reinforcement learning algorithms
- Hugging Face Transformers
    - LLM-based explainability for robot decisions
- Qiskit
    - Quantum decision making

Cybersecurity Context

- Adversarial input simulation
    - Simulate test-time attacks to RL agent inputs
- Trust quantification
    - Track metrics related to transparency, performance, and failure recovery to build trust scores

Experimental Design

Inputs

- Gymnasium
    - Simulated robot sensor data
- Adversarial inputs
    - Altered sensor values
- Failure logs
    - Collision flags
    - Robot actions

Evaluation Metrics

- Collision rate
    - Safety metric
- F1 Score
    - Detecting adversarial inputs
- Path distance
    - Navigation efficiency
- Success Rate
    - Reached the goal without issues
- Trust score
    - Robot responsiveness using human feedback

Baselines

- Proximal Policy Optimization (PPO)
    - Reinforcement learning
- Constrained Policy Optimization (CPO)
    - Train in unknown environment (mapless)
- Deep Autoencoder
    - Adversarial input detection
- Graph Neural Network (GNN)
    - Understanding relationship between entities

Experiment Table - Baseline Comparisons Based on Literature Reviews

Model/Setup	Dataset	Metric	Result	Notes / Source
PPO (Baseline)	Gym	Collision Rate	~28% (est.)	RL baseline for mapless navigation, no safety constraints (Lit Review 4)
CPO (Safe RL)	Gazebo	Task Success Rate	92%	Cost function to avoid unsafe actions during training (Lit Review 4)
		Collision Rate	<10%	Reinforcement learning with safety constraints
Deep Autoencoder	IoT-23	F1 Score (Anomaly)	86.7%	Detects inputs anomalies (Lit Review 6)
GNN (Proposed)	IoT-23	F1 Score (Anomaly)	91.2%	Model dependencies between events over time (Lit Review 6)

SOTA Baseline

Model Name: Constrained Policy Optimization (CPO)

Paper: A Deep Safe Reinforcement Learning Approach for Mapless Navigation

Baseline Role: SOTA safe reinforcement learning algorithm used to improve robot navigation by minimizing unsafe actions like collisions while maximizing task success

In Progress Implementation Plan

- Framework
    - Gymnasium + Pytorch using Stable-Baselines3
- Changes
    - Modify PPO reward structure tp punish for unsafe risks
    - Add custom cost functions to simulate safety constraints
        - Distance to obstacle
    - Simulate adversarial or unpredictable environments
- Environment
    - Gym or RoboTHOR navigation environment
    - Obstacle layout
    - Start and goal positions
- Evaluation:
    - Comparing performance metrics with PPO baseline
        - Collision rate

- Success rate

Comparison

Metric	PPO (Baseline)	CPO (from Paper)	Notes
Task Success Rate	~70% (est.)	**92%**	CPO uses safety-aware reward shaping
Collision Rate	~28% (est.)	**<10%**	Lowered through constraint-based policy optimization
Adaptability	Moderate	High	Generalizes better to dynamic settings

* Results are based on findings from *A Deep Safe Reinforcement Learning Approach for Mapless Navigation*

References

Bhardwaj, Akashdeep, Salil Bharany, Ateeq Ur Rehman, Ghanshyam G. Tejani, and Seada Hussen. “Securing Cyber-Physical Robotic Systems for Enhanced Data Security and Real-Time Threat Mitigation.” EURASIP Journal on Information Security, vol. 2025, no. 1, Jan. 2025, https://doi.org/10.1186/s13635-025-00186-7.

Lv, Shaohua, Yanjie Li, Qi Liu, Jianqi Gao, Xizheng Pang, and Meiling Chen. “A Deep Safe Reinforcement Learning Approach for Mapless Navigation.” 2021 IEEE International Conference on Robotics and Biomimetics (ROBIO), Dec. 2021, pp. 1520–25. IEEE, https://doi.org/10.1109/ROBIO54168.2021.9739251.

Sadrfaridpour, Behzad, and Yue Wang. “Collaborative Assembly in Hybrid Manufacturing Cells: An Integrated Framework for Human–Robot Interaction.” IEEE Transactions on Automation Science and Engineering, vol. 15, no. 3, July 2017, pp. 1178–92, https://doi.org/10.1109/tase.2017.2748386.