Integrated Vision-Physics-Reinforcement Learning Framework for Dynamic Industrial Robot Navigation

Benyamain Yacoob, **Jingyuan Wang**, **Xinyang Zhang**

Machine Intelligence Enthusiasts

ELEE 5350: Machine Learning

University of Detroit Mercy

February 23$^{\text{rd}}$, 2025

Contents

1 Contributions of Team Members

Member	Contributions
Benyamain Yacoob	Primary reader for Papers 1, 2, and 7; collaborative research and paper selection; active participation in group discussions and paper evaluations; cross-team proofreading
Jingyuan Wang	Primary reader for Papers 5 and 6; collaborative research and paper selection; active participation in group discussions and paper evaluations; cross-team proofreading
Xinyang Zhang	Primary reader for Papers 3 and 4; collaborative research and paper selection; active participation in group discussions and paper evaluations; cross-team proofreading

2 Summary of Related Works

Recent advancements in autonomous robot navigation can be grouped into three main categories: reinforcement learning-based navigation, vision and sensor fusion, and physics-informed methods. Each approach offers distinct strengths and limitations, with similarities and differences evident across the reviewed studies.

2.1 Reinforcement Learning-Based Navigation

Studies by (Taheri and Hosseini, 2024), (Ma et al., 2023), (Zhang et al., 2022), and (Tsuruta and Morioka, 2024) explore reinforcement learning (RL) to enhance navigation adaptability. These works highlight real-time decision-making and collision-free pathfinding as key strengths, utilizing algorithms like Proximal Policy Optimization (PPO) and Twin Delayed DDPG (TD3). However, challenges include lengthy training periods and difficulties transitioning from simulation to real-world settings. For instance, (Taheri and Hosseini, 2024) enhances PPO with Residual Blocks, while (Ma et al., 2023) integrates LSTM with TD3 for memory-augmented navigation, showing superior performance in dynamic environments.

2.2 Vision and Sensor Fusion

Research from (Tsuruta and Morioka, 2024), (Schneider and Stemmer, 2024), and (Wang, 2025) focuses on integrating visual and sensory data for comprehensive environmental perception. These studies excel in detailed scene understanding, using techniques like semantic segmentation (e.g., U-Net in (Tsuruta and Morioka, 2024)) and CNNs with LiDAR data (e.g., (Schneider and Stemmer, 2024)). Limitations include high computational demands and sensor synchronization issues. (Wang, 2025) emphasizes multimodal fusion with deep learning, achieving effective obstacle recognition, whereas (Schneider and Stemmer, 2024) leverages synthetic datasets to overcome real-world data scarcity.

2.3 Physics-Informed Methods

The approach by (Liu and Wang, 2021) introduces physics-informed RL, integrating physical laws into a Dyna-style MBRL framework. This method ensures physical consistency and improves generalization, outperforming data-driven RL in sample efficiency. However, it faces increased model complexity and computational needs. Unlike purely data-driven approaches, it reduces real-world interaction demands, as seen in its application to ODE/PDE-governed systems.

Table 1: Comparative Analysis of Reviewed Papers

Paper	**Year**	**ML Method**	**Dataset (Samples)**	**Key Contributions**	**Metrics**	**Main Findings**
Taheri et al.	2024	Enhanced PPO	ROS/Gazebo (10K episodes)	Safe navigation	Success rate	100% success
Wang	2025	Sensor Fusion	Indoor (unspecified)	Multi-modal perception	Recognition rate	Unspecified
Tsuruta et al.	2024	PPO+U-Net	ZED2 (191 pairs)	Monocular navigation	IoU, success rate	0.9147 IoU, 99.4% success
Ma et al.	2023	LTD3+LSTM	Gazebo (unspecified)	Mapless navigation	Success rate	96.5% static success
Schneider et al.	2024	CNN+LiDAR	Gazebo (10,593 samples)	3D detection	mAP	High precision
Zhang et al.	2022	Dist. PPO+LSTM	Gazebo (unspecified)	Distributed navigation	Success rate	84% success
Liu et al.	2021	Physics-MBRL	Simulated (unspecified)	Dynamic control	Sample efficiency	30.2% fewer steps

State-of-the-art methods blend RL with sensory or physical constraints, as seen in (Ma et al., 2023) and (Liu and Wang, 2021). The authors find

RL-based approaches most effective for adaptability, with (Ma et al., 2023) excelling in dynamic navigation due to its memory integration, though (Liu and Wang, 2021) offers superior efficiency for physics-constrained tasks ((Ma et al., 2023), (Liu and Wang, 2021)).

3 Paper Summaries

3.1 Paper 1

3.1.1 Paper Information

- **Authors**: Hamid Taheri, Seyed Rasoul Hosseini, Mohammad Ali Nekoui
- **Title**: Deep Reinforcement Learning with Enhanced PPO for Safe Mobile Robot Navigation
- **Type**: Conference
- **Publication Date**: 2024
- **Number of Citations**: 4

3.1.2 Contributors

- **Primary Reader**: Benyamain Yacoob
- **Proofreader**: Jingyuan Wang, Xinyang Zhang

3.1.3 Research Focus

This study concentrates on applying deep reinforcement learning (DRL) to guide a mobile robot in navigating complex, unmapped environments, a process known as mapless navigation. The authors seek to create a system that directs robots to specific targets while steering clear of obstacles, using real-time sensor data instead of preconstructed maps. Such an approach holds potential for scenarios like search and rescue or exploration in unpredictable settings where maps are absent or impractical.

The research explores several central questions: How can DRL support effective mapless navigation for mobile robots? What adjustments to the Proximal Policy Optimization (PPO) algorithm can elevate its performance in these tasks? Additionally, how do different reward function designs affect the robot's learning and navigation abilities across environments of varying complexity? These inquiries steer the investigation toward practical, reliable navigation solutions.

Notable contributions include an upgraded PPO algorithm with Residual Blocks (ResBlocks) integrated into its neural network structure, enhancing navigation efficiency. The work also presents a thorough observation setup that combines diverse sensory inputs for a complete environmental picture. Lastly, it examines creative reward functions tailored to both simple and obstacle-heavy settings, refining the robot's decision-making process (Taheri and Hosseini, 2024).

3.1.4 Technical Details

The investigation utilizes DRL as its primary method, focusing on a modified PPO algorithm enhanced with ResBlocks in both actor and critic networks, improving gradient flow and convergence speed. For comparison, the authors evaluate their method against the Deep Deterministic Policy Gradient (DDPG) algorithm.

Data comes from simulated environments developed with the Robot Operating System (ROS) and Gazebo simulator. Two setups are tested: an obstacle-free indoor space and a challenging one with strategically placed obstacles, both within a 10x10 square meter area. The observation space includes 30-dimensional LiDAR readings (scaled to 0-1), previous linear and angular velocities, target position in polar coordinates, yaw angle, and target-facing orientation, totaling 16 dimensions. Around 161 training episodes are used, though precise sample numbers beyond this are not detailed.

The setup relies on ROS for sensor and actuator coordination and Gazebo for realistic simulation. The action space consists of linear velocities (0 to 0.25 m/s, via sigmoid) and angular velocities (-1 to 1 rad/s, via tanh), matching the Turtlebot robot's capabilities. Performance is measured through average reward, success rate (percentage of successful navigations), average steps per episode, and cumulative reward, offering a broad evaluation (Taheri and Hosseini, 2024).

3.1.5 Outcomes

Results highlight distinct performance patterns across scenarios. In a simple, obstacle-free environment with a basic reward function (favoring target closeness and penalizing collisions), the modified PPO with ResBlocks surpasses DDPG (100% to 98.46%), achieving quicker navigation and faster adaptation due to its robust design. In contrast, in a complex environment with the same basic reward, DDPG demonstrates greater accuracy, managing intricate obstacle arrangements more effectively despite slower operation.

When an advanced reward function, discouraging wall proximity and rewarding target approach exponentially, is introduced in the complex setting, the modified PPO's performance rises notably, though DDPG maintains a higher success rate. The authors conclude that their PPO upgrades, paired with customized reward designs, significantly advance autonomous navigation capabilities. The study emphasizes the need to align algorithm selection and reward structure with environmental demands, suggesting valuable applications in industrial, commercial, and rescue robotics (Taheri and Hosseini, 2024).

3.2 Paper 2

3.2.1 Paper Information

- **Authors**: X. Wang
- **Title**: Mobile Robot Environment Perception System Based on Multimodal Sensor Fusion
- **Type**: Conference Paper (Applied and Computational Engineering)
- **Publication Date**: January 2025
- **Number of Citations**: Unspecified

3.2.2 Contributors

- **Primary Reader**: Benyamain Yacoob
- **Proofreader**: Jingyuan Wang, Xinyang Zhang

3.2.3 Research Focus

This study investigates the application of multimodal sensor fusion to enhance environmental perception in mobile robots. The primary goal is to develop a system that improves a robot's ability to interpret its surroundings by combining data from multiple sensors, laying a groundwork for future advancements in autonomous navigation and related fields. The research targets applications in areas like intelligent manufacturing, autonomous driving, and disaster rescue, where robust perception is critical.

The paper addresses several key questions: How can integrating data from various sensors improve the completeness and precision of environmental understanding? How does this approach overcome limitations of individual sensors? What role do data fusion algorithms and real-time processing play in boosting adaptability in dynamic settings? Additionally, it explores whether neural network-based models can enhance scene comprehension and navigation autonomy. These inquiries guide the exploration of practical solutions.

Significant contributions include a detailed analysis of multimodal sensor fusion's potential, highlighting the synergy between sensors like LiDAR and cameras. The study also examines data fusion techniques and real-time processing frameworks to refine adaptability, alongside the use of deep learning and computer vision for precise obstacle recognition and path planning (Wang, 2025).

3.2.4 Technical Details

The research utilizes neural network-based feature extraction models and deep learning algorithms to process sensor data, with machine learning techniques applied for pattern recognition to accelerate responses to anomalies. Specific methods include the Extended Kalman Filter (EKF) for real-time state estimation and the Particle Filter (PF) for handling nonlinear, non-Gaussian noise in dynamic environments. Additional tools like Kalman filtering, median filtering, and visual SLAM (Simultaneous Localization and Mapping) support data preprocessing and environmental modeling. A Stream Processing Model facilitates rapid data analysis.

The dataset specifics (e.g., name, sample count) are not detailed, though it involves multimodal inputs from sensors like LiDAR, cameras, and infrared collected in varied environments. Implementation details are also limited, but the study likely leverages standard robotic platforms. Evaluation metrics are not explicitly listed, though outcomes suggest a focus on accuracy, adaptability, and processing speed (Wang, 2025).

3.2.5 Outcomes

Findings indicate that multimodal sensor fusion significantly enhances a robot's perception and decision-making by creating detailed environmental models, improving navigation accuracy and obstacle avoidance. The synergy of sensors, LiDAR for distance, cameras for visual detail, yields robust results. Effective fusion algorithms and appropriate sensor selection further strengthen system reliability. The study concludes that this technology paves the way for advancements in autonomous navigation and adaptive decision-making, offering broad prospects for intelligent systems in dynamic contexts (Wang, 2025).

3.3 Paper 3

3.3.1 Paper Information

- **Authors**: Ryuto Tsuruta, Kazuyuki Morioka
- **Title**: Autonomous Navigation of a Mobile Robot with a Monocular Camera using Deep Reinforcement Learning and Semantic Image Segmentation
- **Type**: Conference
- **Publication Date**: 2024
- **Number of Citations**: 2

3.3.2 Contributors

- **Primary Reader**: Xinyang Zhang
- **Proofreader**: Jingyuan Wang, Benyamain Yacoob

3.3.3 Research Focus

This study explores autonomous navigation for mobile robots using a monocular camera as the primary sensor, diverging from traditional methods reliant on pre-mapped environments and LiDAR. The primary goal is to enhance navigation performance in dynamic settings by integrating deep reinforcement learning (DRL) with semantic image segmentation, aiming to bridge the gap between simulation and real-world applications.

The research addresses key questions: How can a monocular camera, paired with DRL and segmentation, enable effective navigation in unpredictable environments? What improvements can this approach offer over conventional sensor-based methods? The study seeks to develop a practical navigation system adaptable to real-world challenges.

Key contributions include a navigation model combining Proximal Policy Optimization (PPO) with a U-Net segmentation model (using ResNet50 as the encoder), enabling the robot to interpret its surroundings by distinguishing floor areas from obstacles. This integration supports real-time decision-making and obstacle avoidance, validated through both simulation and real-robot testing (Tsuruta and Morioka, 2024).

3.3.4 Technical Details

The study utilizes PPO for DRL, paired with a U-Net segmentation model featuring a ResNet50 encoder to process monocular camera images. This setup allows the robot to classify floor regions and obstacles. Training occurs in a simulated indoor corridor (5-8 meters) using the Unity ML-Agents toolkit, with a dataset collected via a ZED2 camera comprising 191 image pairs (150 training, 16 validation, 25 testing) for automatic floor segmentation. The ROS framework supports real-world testing.

Evaluation metrics include navigation success rate (via Dice loss), Intersection over Union (IoU) for segmentation accuracy, motion smoothness (linear and angular velocity stability), and performance across varying camera heights. These metrics assess the system's effectiveness and adaptability (Tsuruta and Morioka, 2024).

3.3.5 Outcomes

Findings show the PPO model completed 994 out of 1000 episodes (99.4%) in simulation, demonstrating high reliability, with successful generalization to a real robot via ROS. The segmentation model achieved a 0.9147 IoU score, indicating precise floor-obstacle differentiation. Camera height impacted performance, with optimal results at 33 cm, while 44 cm and 71 cm led to unstable velocities and navigation errors. The robot effectively avoids obstacles by slowing down and adjusts turns smoothly using continuous angular velocity. The study concludes that this approach enhances navigation in dynamic environments, offering practical potential for real-world deployment (Tsuruta and Morioka, 2024).

3.4 Paper 4

3.4.1 Paper Information

- **Authors**: Haiyue Ma, Siqi Wang, Shouwu Zhang, Song Ren, Heng Wang
- **Title**: Map-less End-to-end Navigation of Mobile Robots via Deep Reinforcement Learning
- **Type**: Conference
- **Publication Date**: 2023
- **Number of Citations**: 2

3.4.2 Contributors

- **Primary Reader**: Xinyang Zhang
- **Proofreader**: Jingyuan Wang, Benyamain Yacoob

3.4.3 Research Focus

This study investigates mapless navigation for mobile robots using deep reinforcement learning (DRL), diverging from traditional map-based approaches. The primary objective is to develop an end-to-end navigation system that enhances a robot's capability to avoid obstacles and make real-time decisions in unknown, dynamic environments without relying on pre-existing maps.

The research addresses key questions: How can DRL enable effective navigation in unmapped settings? What improvements can be made to enhance obstacle avoidance and decision-making in unpredictable conditions? The study aims to evaluate and refine the robot's navigation performance under such challenges.

Significant contributions include the development of a Long-Term TD3 (LTD3) model, integrating Long Short-Term Memory (LSTM) with Twin Delayed Deep Deterministic Policy Gradient (TD3) to leverage past observations for improved motion prediction. This approach tackles sparse reward issues in RL and supports navigation in complex scenarios (Ma et al., 2023).

3.4.4 Technical Details

The research utilizes the LTD3 model, combining LSTM and TD3 to process sequential data and optimize continuous actions. The framework operates on the TurtleBot3-Waffle-Pi robot, using a 2D LiDAR sensor for state representation. The 76-dimensional state space includes three time-stepped LiDAR frames (24 dimensions each), target coordinates, and the robot's linear and angular velocities. A novel reward mechanism accounts for multiple navigation factors to address sparsity.

Training occurs in the Gazebo simulator, comparing LTD3 against DDPG, PPO, and TD3. Evaluation metrics include training efficiency (learning speed), navigation success rate (task completion percentage), and trajectory smoothness (path stability). Specific sample counts are not detailed, but the focus is on simulation-based validation (Ma et al., 2023).

3.4.5 Outcomes

Findings indicate that LTD3 outperforms other DRL methods in mapless navigation. In Gazebo simulations, it achieves a 96.5% success rate in static environments, surpassing TD3 (95%), PPO (77.5%), and DDPG (70.5%), and an 83.5% success rate in dynamic settings, exceeding comparative algorithms. LTD3 produces smoother, more efficient trajectories, minimizing detours and enhancing obstacle avoidance. The study concludes that integrating long-term memory into DRL significantly improves navigation effectiveness in dynamic, unmapped conditions, offering practical implications for real-world robotics (Ma et al., 2023).

3.5 Paper 5

3.5.1 Paper Information

- **Authors**: Danilo Giacomin Schneider, Marcelo Ricardo Stemmer
- **Title**: CNN-based Multi-Object Detection and Segmentation in 3D LiDAR Data for Dynamic Industrial Environments
- **Type**: Journal (MDPI)
- **Publication Date**: October 2024
- **Number of Citations**: Unspecified

3.5.2 Contributors

- **Primary Reader**: Jingyuan Wang
- **Proofreader**: Benyamain Yacoob, Xinyang Zhang

3.5.3 Research Focus

This study aims to develop a convolutional neural network (CNN) system for multi-object detection and segmentation using 3D LiDAR data, enhancing autonomous navigation for mobile robots in dynamic industrial settings. The focus is on generating a semantic 2D map from Bird's Eye View (BEV) representations to support path planning and navigation by aggregating data from multiple sources.

The research addresses two primary questions: How can deep learning, synthetic data, and a ROS2 communication framework create an effective perception system for navigation in dynamic environments using 3D LiDAR? How can BEV representations enable rapid and precise detection, tracking, and mapping in such contexts? These questions guide the development of a practical solution.

Key contributions include a CNN-based model within a ROS2 framework for data integration and map creation, a synthetic dataset generated via simulation to train the model, and a centralized algorithm operating at 10Hz to synchronize detections and maintain an instance dictionary with class, pose, and occupancy details (Schneider and Stemmer, 2024).

3.5.4 Technical Details

The study utilizes CNNs with a ResNet-50 backbone and KeyPoint Feature Pyramid Network (FPN) for 3D object detection, alongside a Fully Convolutional Network (FCN) for semantic segmentation. BEV representations cover a 10x10 meter area at 608x608 pixels, derived from 3D LiDAR data.

Datasets include a synthetic set from Gazebo Simulator (10,593 samples: 75.52% training, 5.93% validation, 18.9% test) with 14 classes (e.g., person, robots, objects), and a subset of the NVIDIA r2b dataset (annotated sequences r2b_storage and r2b_hallway) for real-world person class evaluation. Tools include ROS2 for communication, Gazebo for simulation, and PyTorch for CNN training. Metrics encompass per-pixel precision, recall, and Intersection over Union (IoU) for segmentation, Average Precision (AP) and Mean Average Precision (mAP) for detection, and absolute errors in position and yaw for localization (Schneider and Stemmer, 2024).

3.5.5 Outcomes

Findings show the model achieves high precision in simulation, accurately predicting pixel labels with minimal errors, and supports effective data aggregation for tracking and mapping. In real-world tests, it partially detects and segments the person class despite synthetic-only training, with performance gaps attributed to domain differences. The study concludes that this approach holds promise for industrial applications, though further real-world validation is needed (Schneider and Stemmer, 2024).

3.6 Paper 6

3.6.1 Paper Information

- **Authors**: Yi Zhang, Zhile Yang, Zihan Zhu, Wei Feng, Zhaokun Zhou, Weijun Wang
- **Title**: Visual Navigation of Mobile Robots in Complex Environments Based on Distributed Deep Reinforcement Learning
- **Type**: Conference (6th Asian Conference on Artificial Intelligence Technology)
- **Publication Date**: 2022
- **Number of Citations**: 3

3.6.2 Contributors

- **Primary Reader**: Jingyuan Wang

- **Proofreader**: Benyamain Yacoob, Xinyang Zhang

3.6.3 Research Focus

This study aims to develop and validate a visual navigation model for mobile robots in complex, mapless environments using distributed deep reinforcement learning (DRL). The primary goal is to enable robots to generate precise actions and find collision-free paths in large-scale, intricate scenes without prior maps or human guidance, relying solely on visual inputs.

The research addresses several key questions: How can DRL support autonomous navigation in unknown, complex settings using visual sensors? How can dividing a complex navigation task into simpler subtasks reduce training time? How does a distributed DRL approach, integrating LSTM and PPO algorithms, enhance navigation accuracy in expansive scenes? These inquiries guide the exploration of an efficient navigation strategy.

Significant contributions include a distributed DRL-based model that segments complex environments into regions, each with a local visual navigation model using LSTM and PPO. The study also introduces a novel reward function based on robot actions, target distance, and runtime, enhancing training effectiveness across these regions (Zhang et al., 2022).

3.6.4 Technical Details

The research utilizes DRL, combining Proximal Policy Optimization (PPO) and Long Short-Term Memory (LSTM) for local navigation models within each region. A Convolutional Neural Network (CNN) processes RGB-D images into embedded vectors for input. The simulation environment, built in Gazebo on Ubuntu 16.04, features four regions (production line, office, outdoor, restaurant) and provides real-time RGB-D images from a robot's first-person perspective, alongside target polar coordinates. Sample counts are not specified.

Implementation involves Gazebo for simulation and Ubuntu 16.04 as the operating system. The model outputs continuous actions (linear and angular velocities). Evaluation metrics include success rates for local region tasks (new goals at 60%, old targets over 75%), rewards during training (continuously rising), and overall navigation success rate (84% in the full environment) (Zhang et al., 2022).

3.6.5 Outcomes

Findings demonstrate that the distributed DRL model enhances navigation performance in large-scale complex environments, achieving an 84% success rate in Gazebo. Segmenting scenes into regions and training local models proves effective, with the model adeptly selecting appropriate regional strategies to reach targets. The study concludes that this approach significantly improves existing DRL navigation frameworks, offering a practical solution for autonomous navigation in intricate settings (Zhang et al., 2022).

3.7 Paper 7

3.7.1 Paper Information

- **Authors**: Xin-Yang Liu, Jian-Xun Wang
- **Title**: Physics-informed Dyna-Style Model-Based Deep Reinforcement Learning for Dynamic Control
- **Type**: Journal (Proceedings of The Royal Society A)
- **Publication Date**: 2021
- **Number of Citations**: 49

3.7.2 Contributors

- **Primary Reader**: Benyamain Yacoob
- **Proofreader**: Jingyuan Wang, Xinyang Zhang

3.7.3 Research Focus

This study focuses on improving the efficiency and accuracy of model-based reinforcement learning (MBRL) for dynamic control by integrating known physical laws into a Dyna-style framework. The primary aim is to leverage the environment's governing equations to refine the learned model, reducing reliance on extensive real-world interactions. This approach, termed Physics-informed MBRL (PiMBRL), targets complex systems governed by ordinary or partial differential equations (ODEs/PDEs).

The research investigates several key questions: How can prior physical knowledge enhance MBRL performance? Can PiMBRL decrease the number of real-world interactions required while boosting control accuracy? How does this method compare to model-free and data-driven MBRL approaches in terms of precision and sample efficiency? These questions drive the exploration of a physics-guided learning strategy.

Key contributions include the PiMBRL framework, which embeds physical constraints into model learning and policy optimization. The study introduces a novel autoencoding recurrent network architecture, combining convolutional encoders, MLP decoders, and LSTM blocks to capture system dynamics. It also demonstrates superior performance across control tasks like Cart-Pole, Pendulum, Burgers' equation, and the Kuramoto-Sivashinsky (KS) equation, showcasing applicability to chaotic systems (Liu and Wang, 2021).

3.7.4 Technical Details

The research utilizes MBRL, incorporating a Dyna-style algorithm that blends model-based and model-free learning. Physics-Informed Neural Networks (PINNs) guide model training by embedding physical laws into the loss function. A recurrent neural network with LSTM units, paired with an encoder-decoder setup (convolutional encoder, MLP decoder), models spatio-temporal dynamics. The Twin Delayed Deep Deterministic Policy Gradients (TD3) algorithm optimizes policies off-policy.

Datasets include simulated environments: Cart-Pole (4D state vector, discrete actions), Pendulum (2D state vector), Burgers' equation (150-grid spatial mesh), and KS equation (64-grid mesh). Sample counts are not specified, but dynamics are detailed. Tools are not explicitly listed, though numerical solvers (e.g., finite difference methods) are implied. Metrics include return (total reward), model prediction error, and sample efficiency (interactions needed) (Liu and Wang, 2021).

3.7.5 Outcomes

Findings reveal PiMBRL outperforms model-free and data-driven MBRL in sample efficiency and accuracy. For Burgers' equation, PiMBRL achieves a 0.1 return in 800 steps versus 1300 for MBRL. In the KS environment, it uses 30.2% of MFRL's steps to reach a -55 return. Physics integration reduces model bias and enhances generalization. The study concludes that PiMBRL offers a robust approach for dynamic control, with potential for broader applications in complex physical systems (Liu and Wang, 2021).

References

Liu, X.-Y., & Wang, J.-X. (2021). Physics-informed dyna-style model-based deep reinforcement learning for dynamic control. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, *477*(2255). https://doi.org/10.1098/RSPA.2021.0618

Ma, H., Wang, S., Zhang, S., Ren, S., & Wang, H. (2023). Map-less end-to-end navigation of mobile robots via deep reinforcement learning. *IEEE 13th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, 48–52. https://doi.org/10.1109/ICEIEC58029.2023.10200799

Schneider, D., & Stemmer, M. R. (2024). Cnn-based multi-object detection and segmentation in 3d lidar data. *Preprints*. https://doi.org/10.20944/preprints202410.0496.v1

Taheri, H., & Hosseini, S. H. (2024). Deep reinforcement learning with enhanced ppo for safe mobile robot navigation. *arXiv preprint arXiv:2405.16266*. https://doi.org/10.48550/arxiv.2405.16266

Tsuruta, R., & Morioka, K. (2024). Autonomous navigation with monocular camera using drl and segmentation. *IEEE SII*, 1107–1112. https://doi.org/10.1109/sii58957.2024.10417188

Wang, X. (2025). Mobile robot environment perception system based on multimodal sensor fusion. *Applied and Computational Engineering*, *127*(1), 42–49. https://doi.org/10.54254/2755-2721/2025.20215

Zhang, Y., Feng, W., Yang, Z., Zhou, Z., Zhu, Z., & Wang, W. (2022). Visual navigation of mobile robots in complex environments based on distributed deep reinforcement learning. *6th Asian Conference on Artificial Intelligence Technology (ACAIT)*, 1–5. https://doi.org/10.1109/ACAIT56212.2022.10137974