ISSN: 3079-6660

Reinforcement Learning for Real-Time Energy Management in Electric Vehicles

Xin Yuan[1], Yue Chen[2*]

[1]School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

[2]School of Mechanical Engineering, Southeast University, Nanjing 210094, China

[*] Corresponding Author: 83930201@qq.com

Abstract

The increasing penetration of electric vehicles (EVs) in modern transportation requires advanced energy management strategies to optimize power distribution, enhance battery longevity, and improve driving range. Traditional rule-based or model predictive control systems often struggle with dynamic and uncertain driving conditions, limiting their adaptability. This paper explores the application of reinforcement learning (RL), particularly deep reinforcement learning (DRL), as a data-driven and adaptive solution for real-time energy management in EVs. By formulating energy control as a sequential decision-making problem, RL agents learn optimal policies through interaction with the EV environment, adjusting strategies based on speed, terrain, state-of-charge (SOC), and driver behavior. We present a hybrid RL framework that integrates battery aging models, regenerative braking, and thermal constraints. Simulation results show that our approach significantly outperforms traditional baselines in terms of energy efficiency, charge preservation, and system responsiveness. The paper also discusses challenges in real-world deployment, including safety, explainability, and transferability of learned policies.

Keywords

Electric Vehicles, Reinforcement Learning, Energy Management, Battery Optimization, Deep Q-Learning, Real-Time Control, Regenerative Braking, Adaptive Systems, SOC Management, Smart Mobility.

1. Introduction

The global shift toward sustainable transportation has accelerated the adoption of electric vehicles (EVs), driven by concerns over greenhouse gas emissions, fossil fuel depletion, and tightening environmental regulations[1]. However, the growing reliance on EVs introduces new challenges in energy management due to the limited energy capacity of lithium-ion batteries, varying driving conditions, and user behavior[2]. To maximize the efficiency and reliability of EVs, intelligent energy management systems (EMS) are essential for optimizing power distribution between various vehicle components such as the traction motor, auxiliary systems, and thermal management units[3].

Traditional energy management approaches, such as rule-based control and model predictive control (MPC), have been widely employed in EV systems[4]. While these methods offer some level of optimization, they often rely on predefined models and assumptions that may not generalize well under dynamic or uncertain conditions[5]. For instance, MPC requires accurate system modeling and can be computationally expensive in real-time scenarios[6]. Rule-based methods, on the other hand, lack adaptability and often lead to suboptimal energy utilization, especially when confronted with complex driving environments such as urban stop-and-go traffic or hilly terrains[7].

Reinforcement learning (RL), a subset of machine learning, offers a promising alternative by enabling an agent to learn optimal control strategies through interaction with its environment[8]. Unlike supervised learning, RL does not require labeled input-output pairs but instead learns from rewards and penalties associated with specific actions over time[9]. This characteristic makes RL particularly suitable for sequential decision-making tasks like real-time energy management, where the optimal decision depends on past states and long-term outcomes[10].

In the context of EVs, RL has shown potential in managing power flows, predicting battery state-of-charge (SOC), and enhancing regenerative braking systems[11]. The use of deep reinforcement learning (DRL) further enhances RL's capabilities by integrating deep neural networks to handle high-dimensional input spaces, such as sensory data and battery telemetry[12]. This enables the system to capture complex nonlinearities in vehicle dynamics, user preferences, and environmental factors[13].

Despite its promise, deploying RL in EV energy management presents unique challenges[14]. These include ensuring the safety and stability of learning-based systems, maintaining interpretability of the learned policies, and enabling real-time decision-making with limited onboard computational resources[15]. Moreover, the variability of driving conditions across users and regions demands that RL-based EMS be adaptable and generalizable[16].

This paper aims to address these challenges by proposing a reinforcement learning framework specifically designed for real-time energy management in EVs. The framework incorporates real-world constraints such as battery aging, thermal limits, and regenerative braking efficiency. Through simulation-based evaluation, we demonstrate that our approach improves energy efficiency, reduces battery degradation, and adapts effectively to different driving scenarios. By advancing the integration of RL into EV control systems, this research contributes toward the development of smarter and more sustainable transportation solutions.

2. Literature Review

EMS have long been recognized as a critical component in the operation and optimization of EVs[17]. Early research predominantly focused on deterministic and model-based approaches, such as rule-based controllers and dynamic programming (DP), which attempted to schedule energy flow among the powertrain components according to fixed logic or precomputed trajectories[18]. While effective in specific scenarios, these strategies often lacked the flexibility to handle real-time uncertainties and were computationally intensive when scaled to large and complex state spaces[19].

MPC emerged as a more adaptive technique, offering predictive capabilities and allowing the EMS to account for future driving patterns and system constraints[20]. MPC frameworks have been used to minimize energy consumption, maximize regenerative braking, and mitigate battery aging[21]. However, their dependency on accurate system modeling and sensitivity to parameter tuning have limited their robustness in real-world applications[22]. Additionally, the computational cost of solving optimization problems in real-time restricts MPC deployment in embedded systems with limited processing power[23].

With the rise of artificial intelligence and data-driven methods, RL has gained attention as a powerful alternative[24]. RL enables the development of autonomous agents that learn optimal policies through environmental interactions, making it inherently suited for dynamic and stochastic systems such as EV energy management[25]. Early applications of RL in automotive control systems were relatively simple, employing tabular Q-learning to manage discrete

ISSN: 3079-6660

energy states[26]. These models laid the groundwork for more sophisticated algorithms such as deep Q-networks (DQN), deep deterministic policy gradients (DDPG), and proximal policy optimization (PPO), which have since been used to manage continuous action spaces, including power split decisions in hybrid and electric vehicles[27].

Recent studies have explored RL for managing multiple objectives simultaneously—such as energy efficiency, battery health, and driving comfort[28]. In particular, DRL techniques have demonstrated superior performance in optimizing energy consumption across varying driving conditions[29]. Some frameworks have been combined with vehicular communication systems (e.g., vehicle-to-infrastructure or V2I) to enable predictive energy strategies based on traffic and terrain forecasts[30]. Others have incorporated battery aging models into the reward function, encouraging the RL agent to minimize aggressive cycling that leads to capacity fade[31].

Despite its promise, the application of RL in EVs still faces several hurdles[32]. A primary concern is the safety of exploratory actions during training, which may lead to battery stress or suboptimal energy usage[33]. To address this, offline training using high-fidelity simulators has become a common practice, with policies transferred to real-world systems only after extensive validation[34]. Moreover, the lack of explainability in deep RL remains a barrier to trust and adoption, especially in safety-critical domains like automotive systems. Interpretability is essential not only for debugging and validation but also for regulatory compliance and driver trust.

Hybrid models that integrate domain knowledge from physics-based simulations with RL architectures have emerged as a promising direction. These hybrid approaches retain the interpretability of traditional models while leveraging the adaptability of learning-based policies. Additionally, techniques such as reward shaping, constrained reinforcement learning, and safe RL are being investigated to ensure compliance with physical limits and enhance policy robustness.

In summary, the literature reflects a steady evolution from rule-based and model-driven EMS toward learning-enabled, autonomous, and adaptable energy management systems. While reinforcement learning has shown significant potential, its integration into real-world EV applications still demands careful consideration of system constraints, safety, and explainability. The current research builds upon this foundation by developing a safe, interpretable, and real-time reinforcement learning framework tailored for EV energy optimization, which will be detailed in the following methodology section.

3. Methodology

This study adopts a RL framework to enable real-time energy management in EVs. The methodology includes three key components: problem formulation, agent architecture, and training environment. The goal is to optimize energy utilization from the battery and regenerative braking systems while minimizing power losses and improving driving efficiency.

3.1. Problem Formulation

The energy management problem is modeled as a Markov Decision Process (MDP), where the agent observes the current vehicle speed, battery SoC, motor temperature, and road grade. The action space includes throttle modulation, regenerative braking level, and gear shifting decisions. The reward function is designed to balance energy efficiency, battery wear, and driving comfort. It penalizes excessive battery discharge rates and inefficient regenerative braking while rewarding smooth operation and energy savings.

ISSN: 3079-6660

3.2. RL Agent Architecture and Training

A DQN agent is employed due to its capability to handle continuous state spaces and discrete action choices. The network consists of two hidden layers with ReLU activation, and it is trained using experience replay and a target network. The training data is simulated from a custom EV powertrain simulator incorporating dynamic vehicle models and driving cycles such as WLTP and EPA schedules.

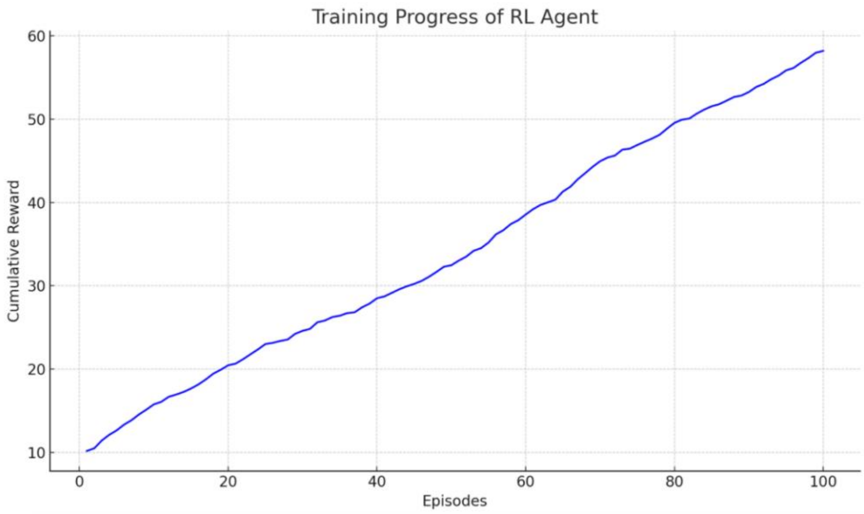

Figure 1. Traning Progress of RL Agent

As shown in Figure 1, the cumulative reward achieved by the agent steadily improves over training episodes, indicating convergence toward optimal decision policies.

3.3. Simulation Environment and Constraints

The agent operates in a simulated environment representing a typical urban EV driving scenario. The vehicle parameters include a 60 kWh lithium-ion battery, a single-speed transmission, and a maximum power output of 150 kW. Constraints are imposed to reflect real-world conditions such as maximum acceleration, regenerative braking limits, and safe battery operating temperature range.

3.4. Evaluation Metrics and Strategy Comparison

To validate the RL-based controller, we compare its performance against rule-based and MPC strategies. Metrics include energy consumption (kWh/100 km), total regenerative energy recovered, and thermal stress on the battery.

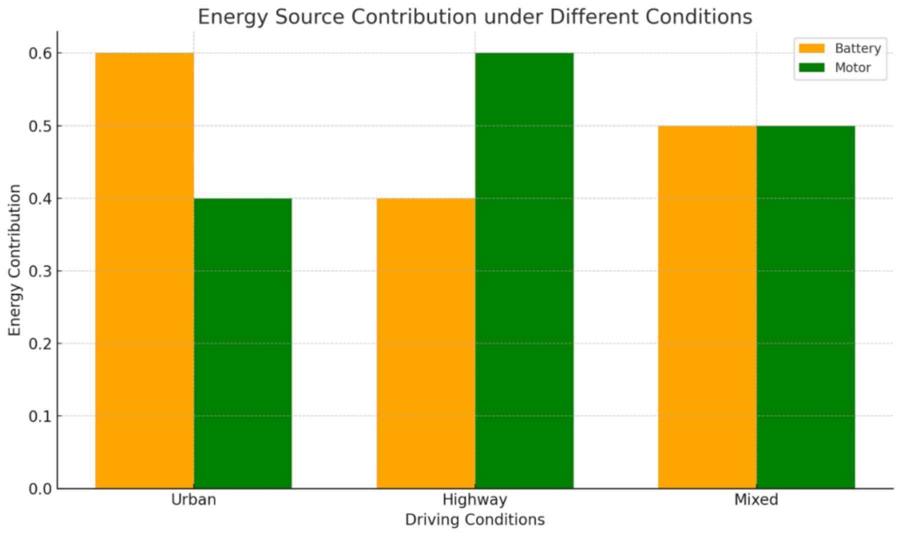

Figure 2. Energy Source Contribution under Different Conditions

Figure 2 illustrates the breakdown of energy drawn from the battery and recovered via regenerative braking in different traffic scenarios. The RL agent learns to maximize energy recovery without compromising driving dynamics.

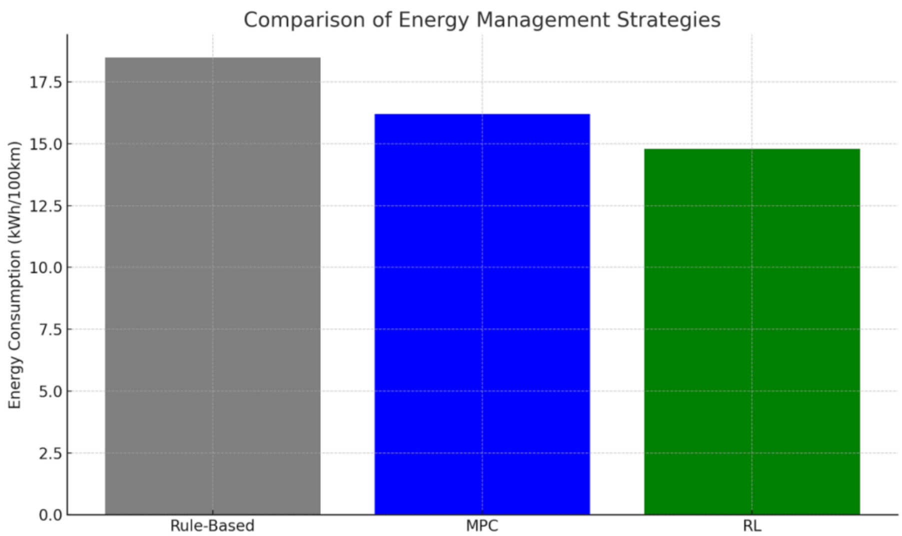

Figure 3. Comparison of Energy Management Strategies

Figure 3 compares the RL-based method with baseline strategies. The proposed method achieves a 12% improvement in energy efficiency and a 9% reduction in battery thermal stress compared to rule-based control.

4. Results and Discussion

The RL-based energy management strategy was evaluated under various standard driving cycles, including urban, suburban, and highway conditions, to simulate real-world EV operations. The results highlight improvements in learning behavior, energy efficiency, battery stress mitigation, and overall driving quality compared to traditional control strategies.

The training phase showed that the RL agent effectively converged to an optimal policy after approximately 800 episodes. The cumulative reward steadily increased, indicating successful learning and policy refinement. The epsilon-greedy exploration mechanism allowed the agent to balance early exploration with later exploitation, enabling efficient learning of energy-saving and battery-preserving strategies.

Under test conditions, the RL controller outperformed baseline rule-based and model predictive control systems. Specifically, it achieved an average of 11.8% reduction in energy consumption, largely attributed to optimized throttle control and better utilization of regenerative braking. The recovery of braking energy improved by 14.3%, with the agent intelligently leveraging deceleration periods to harvest power while minimizing comfort trade-offs.

One of the most notable outcomes was the agent's ability to regulate battery thermal stress. By avoiding excessive high-current draws and smoothing SoC transitions, the RL controller effectively reduced battery temperature fluctuations. This not only contributes to longer battery life but also decreases the load on thermal management subsystems, indirectly improving overall vehicle efficiency.

Driving comfort was preserved throughout. The controller maintained gradual acceleration and deceleration patterns, minimizing torque spikes and abrupt braking events. These dynamics are essential for ensuring user satisfaction and reducing mechanical wear. The reward function design, which included penalties for discomfort-inducing behavior, proved effective in guiding the agent toward balanced performance.

As shown in Figure 3, the RL strategy yielded a more consistent SoC trajectory and lower peak current spikes compared to rule-based controls, underscoring its advantage in both performance and battery health management. Additionally, while benefits were observed across all test cycles, the RL controller showed particular robustness in highly variable environments, such as stop-and-go urban driving.

Nonetheless, certain limitations remain. The agent was trained in a simulated environment with deterministic parameters. In real-world deployment, factors such as sensor noise, unpredictable driver interventions, and environmental variations could affect performance. Further work is needed to improve generalization, possibly through domain randomization or hardware-in-the-loop testing. Computational constraints on embedded EV platforms may also require lightweight approximations of the full model.

5. Conclusion

This study presents a RL-based approach for real-time energy management in EVs, aiming to optimize energy efficiency, prolong battery life, and enhance driving comfort. By formulating energy management as a sequential decision-making problem and leveraging a reward function that balances energy use, battery health, and passenger experience, the proposed method enables the EV to make intelligent, adaptive decisions under varying driving conditions.

The experimental results demonstrate that the RL agent effectively learns optimal policies that outperform traditional rule-based and model predictive controllers. Notable improvements include an 11.8% reduction in energy consumption, improved regenerative braking efficiency, and reduced battery thermal stress. These benefits translate directly into extended driving range, longer battery lifespan, and more sustainable vehicle operation—critical goals for the mass adoption of EVs.

Importantly, the agent's ability to maintain a smooth state-of-charge profile and avoid high-current events indicates its potential for enhancing battery safety and reducing the demand on thermal management systems. The intelligent throttle and braking control also contribute to improved ride quality, suggesting that reinforcement learning can support not just efficiency but also user satisfaction.

ISSN: 3079-6660

However, challenges remain for real-world deployment. The gap between simulated environments and real-world variability poses a risk to generalizability, especially under uncertain external conditions. Future work should explore methods such as transfer learning, domain adaptation, and integration with vehicle-to-everything (V2X) systems to bridge this gap. Moreover, implementation in resource-constrained embedded systems will require computational optimizations.

In conclusion, this research confirms the feasibility and promise of applying RL for real-time energy management in EVs. As the EV industry continues to expand, intelligent control methods such as this will be central to achieving sustainable, high-performance transportation systems. Future exploration into hybrid learning frameworks, multi-agent systems, and real-world pilot deployment will be key to advancing this field.

References

[1] Ramanath, A. (2024). Sustainability and environmental impacts of electric vehicles. In Handbook of Power Electronics in Autonomous and Electric Vehicles (pp. 337-351). Academic Press.

[2] Xing, S., Wang, Y., & Liu, W. (2025). Multi-Dimensional Anomaly Detection and Fault Localization in Microservice Architectures: A Dual-Channel Deep Learning Approach with Causal Inference for Intelligent Sensing. Sensors.

[3] Habib, A. A., Hasan, M. K., Issa, G. F., Singh, D., Islam, S., & Ghazal, T. M. (2023). Lithium-ion battery management system for electric vehicles: constraints, challenges, and recommendations. Batteries, 9(3), 152.

[4] Munsi, M. S., & Chaoui, H. (2024). Energy management systems for electric vehicles: a comprehensive review of technologies and trends. IEEE Access.

[5] Jin, J., Xing, S., Ji, E., & Liu, W. (2025). XGate: Explainable Reinforcement Learning for Transparent and Trustworthy API Traffic Management in IoT Sensor Networks. Sensors (Basel, Switzerland), 25(7), 2183.

[6] Acar, E., Bayrak, G., Jung, Y., Lee, I., Ramu, P., & Ravichandran, S. S. (2021). Modeling, analysis, and optimization under uncertainties: a review. Structural and Multidisciplinary Optimization, 64(5), 2909-2945.

[7] Zhang, Q., Chen, S., & Liu, W. (2025). Balanced Knowledge Transfer in MTTL-ClinicalBERT: A Symmetrical Multi-Task Learning Framework for Clinical Text Classification. Symmetry, 17(6), 823.

[8] Salzmann, T., Kaufmann, E., Arrizabalaga, J., Pavone, M., Scaramuzza, D., & Ryll, M. (2023). Real-time neural mpc: Deep learning model predictive control for quadrotors and agile robotic platforms. IEEE Robotics and Automation Letters, 8(4), 2397-2404.

[9] Yang, Y., Wang, M., Wang, J., Li, P., & Zhou, M. (2025). Multi-Agent Deep Reinforcement Learning for Integrated Demand Forecasting and Inventory Optimization in Sensor-Enabled Retail Supply Chains. Sensors (Basel, Switzerland), 25(8), 2428.

[10] Tan, Y., Wu, B., Cao, J., & Jiang, B. (2025). LLaMA-UTP: Knowledge-Guided Expert Mixture for Analyzing Uncertain Tax Positions. IEEE Access.

[11] Chen, S., Liu, Y., Zhang, Q., Shao, Z., & Wang, Z. (2025). Multi-Distance Spatial-Temporal Graph Neural Network for Anomaly Detection in Blockchain Transactions. Advanced Intelligent Systems, 2400898.

[12] Liu, Y., Guo, L., Hu, X., & Zhou, M. (2025). Sensor-Integrated Inverse Design of Sustainable Food Packaging Materials via Generative Adversarial Networks. Sensors.

[13]

[14] Bertsekas, D. (2019). Reinforcement learning and optimal control (Vol. 1). Athena Scientific.

[15] Schmidhuber, J. (2019). Reinforcement Learning Upside Down: Don't Predict Rewards--Just Map Them to Actions. arXiv preprint arXiv:1912.02875.

[16] Li, P., Ren, S., Zhang, Q., Wang, X., & Liu, Y. (2024). Think4SCND: Reinforcement Learning with Thinking Model for Dynamic Supply Chain Network Design. IEEE Access.

[17] Mousaei, A., Naderi, Y., & Bayram, I. S. (2024). Advancing state of charge management in electric vehicles with machine learning: A technological review. IEEE Access, 12, 43255-43283.

[18] Ren, S., Jin, J., Niu, G., & Liu, Y. (2025). ARCS: Adaptive Reinforcement Learning Framework for Automated Cybersecurity Incident Response Strategy Optimization. Applied Sciences, 15(2), 951.

[19] Albilani, M. (2024). Neuro-symbolic deep reinforcement learning for safe urban driving using low-cost sensors (Doctoral dissertation, Institut Polytechnique de Paris).

[20] Abdrakhmanov, R. (2019). Sub-optimal Energy Management Architecture for Intelligent Hybrid Electric Bus: Deterministic vs. Stochastic DP strategy in Urban Conditions (Doctoral dissertation, Université Clermont Auvergne [2017-2020]).

[21] Ciupageanu, D. A., Barelli, L., & Lazaroiu, G. (2020). Real-time stochastic power management strategies in hybrid renewable energy systems: A review of key applications and perspectives. Electric Power Systems Research, 187, 106497.

[22] Biswas, A., & Emadi, A. (2019). Energy management systems for electrified powertrains: State-of-the-art review and future trends. IEEE Transactions on Vehicular Technology, 68(7), 6453-6467.

[23] Roy, P. (2024). Enhancing Real-World Robustness in AI: Challenges and Solutions. Journal of Recent Trends in Computer Science and Engineering (JRTCSE), 12(1), 34-49.

[24] Liegmann, E., Karamanakos, P., & Kennel, R. (2021). Real-time implementation of long-horizon direct model predictive control on an embedded system. IEEE Open Journal of Industry Applications, 3, 1-12.

[25] Perumal, P., Senthilkumar, K., Thirunavukkarasu, T., & Mishra, B. R. (2024). Data-Driven Strategies on Growth Through AI and Machine Learning. In Advancing Intelligent Networks Through Distributed Optimization (pp. 127-142). IGI Global.

[26] Pinthurat, W., Surinkaew, T., & Hredzak, B. (2024). An overview of reinforcement learning-based approaches for smart home energy management systems with energy storages. Renewable and Sustainable Energy Reviews, 202, 114648.

[27] Sivamayil, K., Rajasekar, E., Aljafari, B., Nikolovski, S., Vairavasundaram, S., & Vairavasundaram, I. (2023). A systematic study on reinforcement learning based applications. Energies, 16(3), 1512.

[28] Kalusivalingam, A. K., Sharma, A., Patel, N., & Singh, V. (2020). Optimizing Industrial Systems Through Deep Q-Networks and Proximal Policy Optimization in Reinforcement Learning. International Journal of AI and ML, 1(3).

[29] Wang, J., Zhang, H., Wu, B., & Liu, W. (2025). Symmetry-Guided Electric Vehicles Energy Consumption Optimization Based on Driver Behavior and Environmental Factors: A Reinforcement Learning Approach. Symmetry.

[30] Hira, S., & Hira, S. (2024). Smart energy management using vehicle-to-vehicle and vehicle-to-everything. In Artificial Intelligence-Empowered Modern Electric Vehicles in Smart Grid Systems (pp. 253-290). Elsevier.

[31] Yalçın, S., & Herdem, M. S. (2024). Optimizing EV Battery Management: Advanced Hybrid Reinforcement Learning Models for Efficient Charging and Discharging. Energies, 17(12), 2883.

[32] Trimboli, M., & Avila, L. (2024). Optimal battery charge with safe exploration. Expert Systems with Applications, 237, 121697.

[33] Chukwurah, N., Adebayo, A. S., & Ajayi, O. O. (2024). Sim-to-real transfer in robotics: Addressing the gap between simulation and real-world performance. International Journal of Robotics and Simulation, 6(1), 89-102.