JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Evolution of Deep Reinforcement Autonomous Racing: A Study

Mrs. P. Navya[1], Cheripally Jitendra Kumara Prasad[2], Arisela Meghana[3], Toomu Sahithi[4]
Assistant Professor[1], Student/Research Scholar[2], Student/Research Scholar [3], Student/Research Scholar [4]
Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning)
Vignana Bharathi Institute of Technology, Aushapur(V), Ghatkesar (M), Medchal Dist. 501301, Telangana, India

Abstract: Car racing, a popular sport since the early twentieth century, has grown alongside the automotive industry, entertaining fans in prominent championships such as F1, WRC, and NASCAR. Success in these competitions' hinges on two critical factors: engineering the most potent and durable machines and training expert pilots capable of maximizing their potential on the track. With artificial intelligence achieving significant advancements in a variety of fields in recent years, a pressing question arises: Can AI outperform human drivers in motor racing? This paper delves into this inquiry by surveying the landscape of autonomous racing agents developed through deep reinforcement learning algorithms. By analysing these algorithms, which enable agents to autonomously learn and refine their racing strategies without external guidance, this study explores the potential for AI to achieve unparalleled consistency and efficiency on the racetrack. This review opens new possibilities for future developments in autonomous racing by providing insightful knowledge on how reinforcement learning methods are applied in this groundbreaking domain.

Index Terms **- Car Racing, Deep Reinforcement Learning, Autonomous Racing, Artificial Intelligence (AI), Self-Racing Cars, Autonomous Vehicles, Racing games, Racing Simulations.**

I. Introduction

Artificial intelligence has become a major area of interest worldwide and has demonstrated its potential to excel in specific tasks through training. Within AI, Reinforcement Learning (RL) stands out as a subdomain that enables computer algorithms to learn autonomously, like human learning processes. RL works on the principle of reinforcing desirable actions through rewards and discouraging undesirable ones through punishments. This reward-based system helps computers understand their environment, the tasks at hand, and the goals they want to achieve. RL algorithms are ubiquitous in various everyday applications, including OpenAI's ChatGPT, which uses Human Feedback Reinforcement Learning (HFRL), and Google's ad recommendation systems, among others.

Driving and racing present people with complex challenges that require practice and improvement of their skills. Similarly, AI can use reinforcement learning to gain proficiency in complex tasks like racing, iterate through practice sessions, and gain experience through rewards. Scaling reinforcement learning to high-dimensional tasks is possible with neural networks, leading to the development of deep reinforcement learning algorithms. These models combine the ability of RL to handle complicated tasks with the robustness of deep learning architectures in understanding complex environments. Despite these advantages, deep reinforcement learning algorithms come with certain disadvantages, notably the need for careful fine-tuning of reward functions and model parameters, as well as longer training times compared to traditional deep learning models.

II. Reinforcement Learning

Reinforcement learning is a machine learning paradigm that facilitates learning by obtaining reinforcements that are dependent on the actions the model performs within a given task. Essentially, it is like a trial-and-error method in which the model, called an agent, performs successive trials to explore and refine its actions. The agent interacts with an environment that represents the domain of the task it wants to accomplish. Within RL, environments evaluate the actions taken by the agent in the context of its current state, determine whether those actions are beneficial or harmful, and subsequently provide appropriate rewards. This iterative process of interaction between the agent and the environment, characterized by action selection and reward acceptance, continues until the agent recognizes the optimal course of action to maximize cumulative rewards, given its current state - a fundamental feature of reinforcement learning.

Deep Reinforcement Learning extends RL by equipping the agent with neural networks that serve as a cognitive architecture, capable of recognizing intricate patterns in complex environments characterized by an extensive set of state-action pairs, such as they are typically found in racing environments. This expansion allows the agent to process and interpret a variety of environmental stimuli, leveraging the neural network's ability to recognize and respond to composite patterns. However, the use of neural networks in deep reinforcement learning introduces additional complexities, particularly the challenge of selecting suboptimal actions in certain states to optimize long-term rewards - a task exacerbated by the inherent complexity of racing environments. This requirement for strategic action selection highlights the nuanced nature of Deep RL, where optimal decision making involves balancing immediate benefits with long-term goals in the context of complex environmental dynamics.

III. Literature Survey

Recent years have seen tremendous progress in the field of autonomous racing, primarily because to the astonishing breakthroughs in deep reinforcement learning techniques, which allow agents to interact with simulated environments and learn and refine their racing strategies on their own.

3.1 2D racing game with reinforcement learning and supervised learning [1]

Authored by Henry Teigar, Miron Storožev and Janar Saks from the University of Tartu, explores the application of machine learning techniques, particularly RL and supervised learning, in the context of autonomous racing. The authors address the challenges associated with training models in real-world environments due to the possibility of errors, particularly crashes, requiring the use of virtual simulators. Their main goal is to develop a 2D racing simulator using Pygame and OpenAI Gym environments to train models that can complete courses flawlessly using RL and supervised learning approaches.
The literature review section provides a comprehensive overview of relevant work in the field. The wide application of RL is discussed and its success in various fields such as robotic obstacle avoidance and visual navigation is highlighted. In addition, the emergence of deep reinforcement learning as a powerful methodology is highlighted and examples are given from well-known works such as “Deep Reinforcement Learning: Pong from Pixels” by Andrej Karpathy and “Write an AI to win at Pong from Scratch with Reinforcement Learning” by Dhruv Parthasarathy. The review also draws inspiration from DeepMind's proposed architecture, as described in "Demystifying Deep Reinforcement Learning" by Tambet Matiisen.

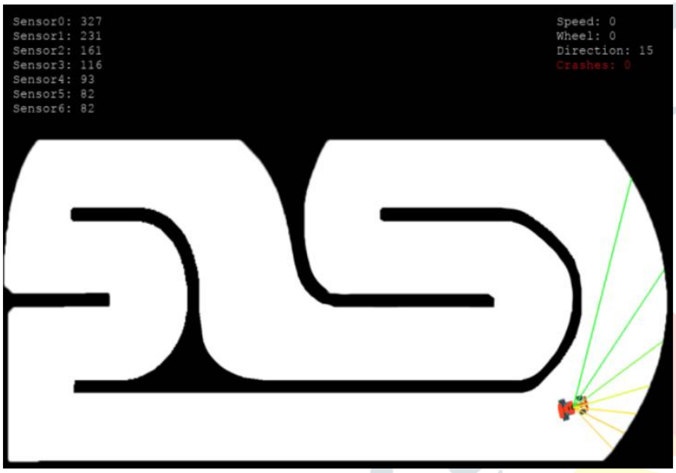


Fig-1: Gameplay in Pygame environment

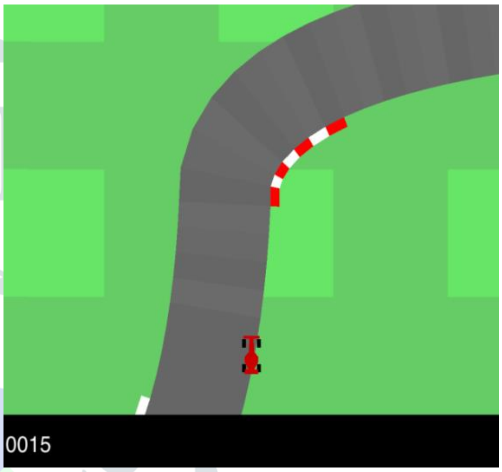


Fig-2: Gameplay in CarRacing-v0 environment

The authors delve into the fundamental concepts of RL and explain the interaction between an agent and its environment in the context of a Markov Decision Process (MDP). They discuss the challenges associated with RL, including the credit allocation problem and the exploration-exploit dilemma, which complicate the learning process. Furthermore, the review introduces policy gradients as a popular approach for RL implementation and contrasts it with the traditional Q-learning method. In their study of simulation environments, the authors detail the development of both homegrown and OpenAI Gym environments for training autonomous racing models. They provide insights into the design and reward systems of these environments and illustrate the differences in input modalities and training methods. The article also discusses the challenges encountered in training, such as the need for image preprocessing and exploring reward systems to encourage desired behaviour.

The experimental results presented in the paper provide a comprehensive analysis of the performance of RL and supervised learning models in autonomous racing tasks. The authors discuss the effectiveness of various training approaches, including the use of custom reward systems and image preprocessing techniques. They critically evaluate the strengths and limitations of each approach and provide valuable insights for future research in the field. Finally, the paper contributes to the evolving landscape of autonomous racing by demonstrating the effectiveness of RL and supervised learning techniques in training models for navigation and control tasks. By combining theoretical insights with practical experiments, the authors provide a valuable resource for researchers and practitioners seeking to harness the power of machine learning in autonomous systems.

3.2 End-to-End Race Driving with Deep Reinforcement Learning [2]

The article by Jaritz and team presents an end-to-end deep reinforcement learning approach to autonomous racing using only forward-facing camera images and speed data. The key contribution is the use of the Asynchronous Advantage Actor-Critic (A3C) algorithm combined with tailored reward shaping, control discretization and agent initialization strategies to achieve faster convergence and more robust driving compared to previous deep reinforcement learning racing methods. The authors demonstrate their approach in the more realistic and complex WRC6 racing simulator, which significantly outperforms simpler environments such as The Open Racing Car Simulation (TORCS). This complex world with stochastic elements such as changing weather and lighting more closely approximates the challenges of the real world. Despite this difficulty, the agent will learn to drive through sharp curves, downhill roads and winding sections at high speeds of 73 km/h on 29.6 km of varied training routes.

The authors also demonstrate some cross-domain generalizations by testing real-world driving videos. This type of simulation-to-reality transfer has not been demonstrated for end-to-end deep reinforcement learning driving before. Their ablation studies continue to demonstrate the value of the proposed learning improvements over baseline approaches. This work clearly advances autonomous racing capabilities. The remaining limitations provide opportunities for further development. For example, while speeds exceeding human levels have been achieved, it still lacks optimal racing suitability in terms of trajectories and collision avoidance and despite the encouraging real-world demonstrations, sim-to-real transfer remains a major challenge.

Subsequent research could pursue adapting this approach to real vehicles, leveraging Sim2Real techniques such as domain randomization or applying this method to newer racing simulators with more advanced physics and graphics. Expanding the perception stack beyond the front-facing camera, such as surround views or spatial sensors like LiDAR could also lead to faster driving. By further optimizing reward functions and network architectures, robustness can be further advanced. By demonstrating A3C racing in realistic stochastic simulators and first real-world deployments, this article pushes autonomous driving research in exciting new directions. However, many open research questions remain to achieve human/superhuman racing on real tracks. This paper lays the foundation for researchers to build on as they address the next phases of this grand challenge.

3.3 Formula RL: Deep Reinforcement Learning for Autonomous Racing using Telemetry Data [3]

Remonda et al. (2019) investigates the application of Deep RL techniques to the problem of autonomous racing. The authors acknowledge that autonomous driving for racing cars differs from typical self-driving scenarios as the primary goal is to minimize lap times rather than prioritizing safety. Traditional approaches to autonomous racing rely on control theory and heuristics and often require manual tuning and domain knowledge for each new track or vehicle configuration. In contrast, RL offers the potential to learn driving policies directly from experience, without the need for explicit modelling or manual parameter tuning.

The authors describe the autonomous racing problem as a reinforcement learning task in which the agent receives a multidimensional input that includes vehicle telemetry data and a continuous action space representing steering, acceleration and braking. They explore various Deep RL algorithms and architectures, including Deep Deterministic Policy Gradient (DDPG), Long Short-Term Memory (LSTM) networks, multi-stage objectives, and prioritized experience replay. To address the challenges of continuous action spaces with mutually exclusive outputs (e.g., throttle and brake), the authors propose a technique called “brake exploration,” which controls the exploration process to avoid conflicting actions.

Through a series of experiments conducted in the open-source racing simulator TORCS, Remonda et al. Evaluate the performance of the proposed RL algorithms and architectures. In the first study, they investigate whether RL models can effectively learn to drive a race car using telemetry data and compare their performance with state-of-the-art hand-crafted bots. The results show that RL models, particularly those trained with prioritized experience replay and a novel “look-ahead curvature” feature, outperform baseline bots in terms of lap times on complex tracks.

In the second study, the authors evaluate the generalization abilities of the learned RL models by evaluating their performance on unknown routes. The results suggest that models trained on relatively complex routes can generalize relatively well to unknown routes, although their performance is worse than that of models trained specifically on these routes. This behaviour is compared to the practice of professional human drivers who familiarize themselves with each new route to exploit their full potential.

Overall, the work of Remonda et al. makes significant contributions to the field of autonomous racing with Deep RL. The authors propose novel techniques such as brake exploration and look-ahead curvature to address the challenges of continuous action spaces and integrate route information. Their experiments demonstrate the feasibility and potential of using RL models trained solely on telemetry data to achieve competitive lap times and transfer them to new scenarios. The paper also highlights the importance of considering generalization capabilities when evaluating autonomous driving models, as real-world deployment often involves encountering unseen environments.

3.4 Deep Reinforcement Learning for Autonomous Driving [4]

The research explores the use of deep reinforcement learning, specifically the Deep Deterministic Policy Gradient (DDPG) algorithm, to train policies for autonomous driving. While previous work has applied Deep RL to simpler racing games, the key contribution here is the adoption for the much more realistic and complex TORCS simulator. The authors design an appropriate DDPG model architecture and a sensor/reward configuration tailored to TORCS. Through training in hundreds of episodes, your agent will learn to drive at high speeds of more than 100 km/h and successfully complete laps by learning steering, acceleration and braking controls. The agent generalizes well to compete with built-in AI drivers.

The performance is promising, but not yet superhuman. Your racer has learned basic driving concepts but is not optimizing racing lines or avoiding collisions in multi-agent environments. This is likely due to only providing training in single agent scenarios. The videos also reveal some instability in fully mastering the vehicle dynamics. Nonetheless, this work advances Deep RL's autonomous driving capabilities and demonstrates them in a far more realistic environment than previous attempts. The next steps could focus on improving Sim2Real transmission through domain randomization. The DDPG approach could also integrate more sensors such as surround cameras or spatial inputs from LiDAR, rather than just front-facing data. Additionally, multi-agent environments could improve robustness by learning to negotiate with other vehicles. The main limitations arise from the gap between simulation and reality and a lack of perceptual competence. However, this creates a strong Deep RL baseline in TORCS. Follow-up work can build on these foundations by bridging the reality gap – for example, by applying Sim2Real techniques or back-translating policies into real-world datasets and perception could be improved by incorporating auxiliary goals such as semantic segmentation or depth estimation.

(a) Map Aalborg (b) Training Mode: shaky at beginning of training

(c) Compete Mode: falling behind at beginning (d) Compete Mode: Overtake other cars at corner

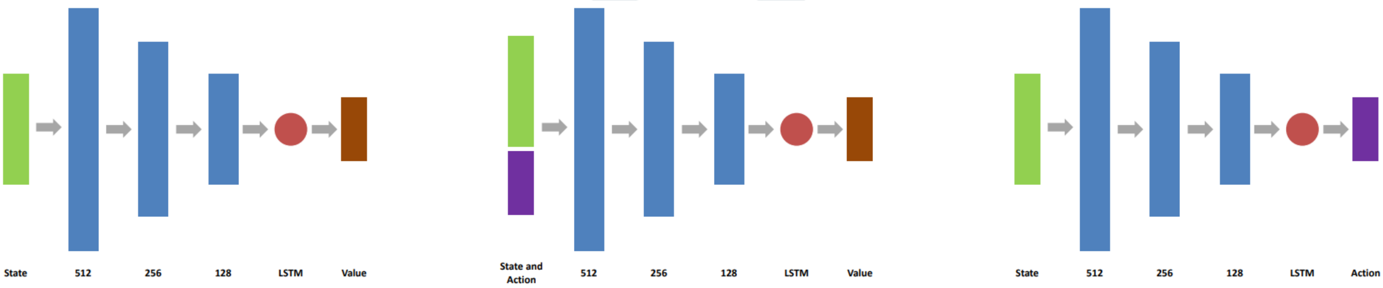
Fig-3: Training and Evaluation in TORCS environment

In summary, this paper provides a practical blueprint for realistic simulator based Deep RL for autonomous racing. While far from solved, it lays the foundation for subsequent research to improve robustness, multi-agent interactivity, and adaptation to real-world autonomous driving environments. This opens an exciting new direction beyond simpler game-based environments.

3.5 Autonomous Car Racing in Simulation Environment Using Deep Reinforcement Learning [5]

Kıvanç Güçkıran and Bülent Bolat present a comprehensive study on the application of Deep Reinforcement Learning techniques to develop autonomous agents for car racing in a simulation environment. The authors examine TORCS platform, which serves as a highly portable and open-source car racing simulation environment enabling the development and testing of artificial intelligence agents. The study delves into the complex area of reinforcement learning, a goal-oriented machine learning practice that aims to maximize an agent's cumulative rewards through interactions with the environment. The authors provide a detailed overview of the theoretical foundations of reinforcement learning, including MDPs, value-based and policy-based methods, as well as hybrid approaches such as actor-critic methods.

One of the main contributions of this paper is the implementation and evaluation of two state-of-the-art Deep RL algorithms: Soft Actor-Critic (SAC) and Rainbow DQN. The authors present a detailed description of the SAC algorithm, which uses a modified RL objective function using maximum entropy formulation, stimulates exploration and promotes stability by dual Q networks, objective value networks and the reparameterization trick. The Rainbow DQN algorithm, on the other hand, combines various improvements over traditional DQN, such as Double Q-Learning, Priority Experience Replay, Dueling Networks and Noisy Networks. These improvements aim to address challenges such as overestimation, efficient exploration and value function estimation, thereby improving agent performance. A key aspect of the study is the authors' approach to reward design and termination conditions, which are critical to effective learning in the TORCS environment. They explore various reward functions, including those that consider track position, lateral speed, and angle penalty, to mitigate undesirable behaviour such as slalom riding. Additionally, they introduce a “try-brake” mechanism to encourage the agent to learn when to brake effectively.



(a) SAC V Network (b) SAC Q Network (c) SAC Policy Network

Fig-4: Soft Actor-Critic Architecture

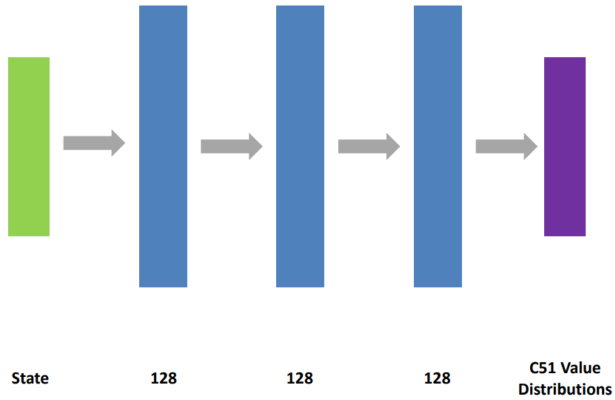


Fig-5: Deep Q Network Architecture

The paper also addresses the issue of generalization, where the authors train and test the agents on multiple road routes to prevent overfitting and promote transferability to unfamiliar environments. They provide detailed insights into their exploration strategies, action space representations, and network architectures used for the SAC and Rainbow DQN algorithms. The results presented in the paper demonstrate the successful implementation of both algorithms, with the SAC-LSTM agent exhibiting superior data efficiency and trace consistency. The authors provide comprehensive performance metrics, including episode values, speeds, and lane-specific ratings, highlighting the agents' ability to complete routes at high average and maximum speeds while also transferring well to unfamiliar environments.

Overall, this paper makes a significant contribution to this field of autonomous racing and the application of deep reinforcement learning techniques in simulation environments. The authors' rigorous approach, in-depth analysis, and insightful discussions pave the way for further advances in this field and demonstrate the potential of Deep RL methods for self-driving car development and real-world robotics applications.

3.6 Super-Human Performance in Gran Turismo Sport Using Deep Reinforcement Learning [6]

Fuchs et al. (2021) presents a novel approach to autonomous car racing using deep reinforcement learning. Autonomous car racing is a challenging problem in robotics because it requires planning trajectories in the shortest possible time under uncertain dynamics and controlling the car at its driving limits. Classic approaches such as model predictive control (MPC) and model predictive path integral control (MPPI) have shown impressive results in controlling physical vehicles at high speeds. However, these approaches have limitations, such as: lack of flexibility in cost function design and the need for highly parallel calculations. To overcome these limitations, the authors propose a learning-based system that utilizes high-precision physical vehicle simulation, course-progress proxy reward, and deep reinforcement learning. The system is used in Gran Turismo Sport, a world-leading car simulator known for its realistic physics simulation of various racing cars and tracks and is even used to recruit human racers.

The authors formulate the minimum time race problem as maximizing route progress using an exponentially discounted future reward as a proxy. This approach allows for a compromise between an easily attributable but biased reward and a reward that is closer to the overall lap time goal. In addition, a wall contact penalty is introduced to incentivize collision avoidance. Driving policy is represented by a deep neural network that maps input features such as linear velocity, acceleration, odometer measurements and curvature measurements to control commands for steering angle, throttle and brake. The network parameters are optimized using the SAC algorithm, state-of-the-art off-policy algorithm with maximum entropy and deep RL.

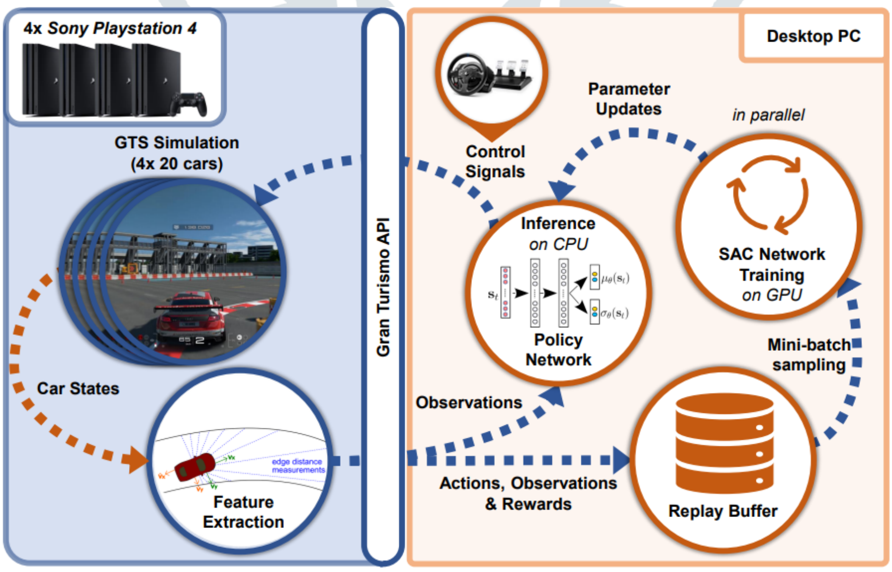


Fig-6: System Architecture for Gran Turismo

The authors evaluate their approach in three racing environments with different vehicles and tracks of varying difficulty and compare the lap times achieved by their approach with those of the built-in AI and over 50,000 human drivers. The results show that their approach outperforms the best human lap time in all three reference settings and overcomes the limitations of the built-in AI. Furthermore, the authors analyse the learned driving behaviour and find that their approach learns to drive trajectories that are qualitatively like those of the best human players, while maintaining slightly higher average speeds in corners.

The work also investigates the robustness of their trained agent by modifying the test environment, e.g. B. by applying the policy to a new car or track, changing tire friction, adding noise to observations and delaying conclusions. The results suggest that the agent can make corrections to its path to some extent, but some parts of the route require stronger corrections to accommodate changes in dynamics. Overall, this paper represents a significant achievement in autonomous car racing and demonstrates the first super-human performance in time trial environments using deep reinforcement learning. The authors' approach overcomes the limitations of classical approaches and imitation learning and paves the way for further investigations of reinforcement learning and neural network controllers to control autonomous vehicles at their driving limits.

3.7 Autonomous Racing using a Hybrid Imitation-Reinforcement Learning Architecture [7]

To minimize lap times in a time attack racing event, this paper presents a strict end-to-end control strategy for autonomous vehicles. The AutoRACE Simulator is a simulation system that was created especially for this research project and allows the authors to simulate realistic audio-visual effects along with accurate dynamics of vehicles and environments. In less than 20 hours, a deep neural network policy will have been trained to drive an autonomous car via imitation learning and to race it using reinforcement learning, thanks to the study's adoption of a hybrid imitation-reinforcement learning architecture and innovative reward function. The environment, the agent, the actor (vehicle), and the academy (which oversees training and decision-making) make up the four essential components of the suggested learning architecture, which combines imitation and reinforcement learning techniques.

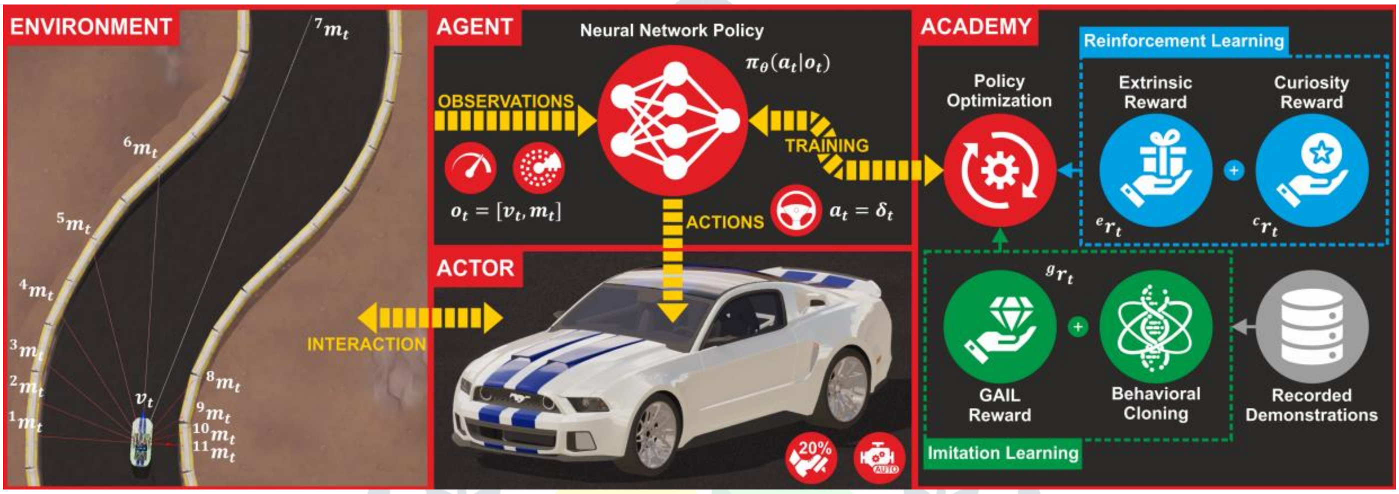


Fig-7: Hybrid Imitation-Reinforcement Learning Architecture

After letting a human driver operate the car for 11 laps on the racetrack, the authors first recorded a demonstration dataset. Using imitation learning, this dataset was used to teach the autonomous agent the fundamentals of driving. Later, a hybrid learning architecture was used, in which the agent's driving ability was prioritized during exploration thanks to imitation learning, and its racing performance was enhanced recursively through reinforcement learning. Utilizing a deep neural network policy, the agent translated environmental observations, such as velocity and range measurements, into discrete steering commands. Based on four signals, the policy was optimized: a novel extrinsic reward function (minimizing lap times and enforcing motion constraints), curiosity reward (promoting exploration), generative adversarial imitation learning (GAIL) reward, and behavioural cloning (imitation learning).

The experiment involved training an autonomous agent and testing its performance against ten different human players in a time attack racing scenario. The outcomes showed that the autonomous agent was, on average, 1 point 46 seconds faster than the human players and outperformed all the human players in nine out of ten courses. Additionally, by a margin of 0 points 96 seconds, the autonomous agent outperformed the top human player in these kinds of races, a major advantage. Because it can optimize trajectories and prevent continuous drifting during cornering maneuvers, the autonomous agent is superior, according to the authors. Furthermore, by minimising undesired drifting, the autonomous agent's increased control frequency helped it to dominate human players in terms of velocity. The open, realistic simulation system for autonomous racing research presented in this work is ideal for training an autonomous agent that can outperform a human expert in a simulated racing environment. Additionally, the hybrid learning approach that this work employs combines the power of imitation and reinforcement learning.

3.8 Memory based neural networks for end-to-end autonomous driving [8]

The article presents a novel, memory-based, end-to-end neural architecture tailored for steering and throttle control in autonomous driving and addresses the increasing interest in this area, driven by its potential benefits such as improved traffic safety and Mobility optimization is required. Autonomous driving methods are divided into modular and end-to-end pipelines, with the latter receiving particular attention due to its ability to derive decisions directly from input data, predominantly images.

The literature is reviewed with an emphasis on the importance of taking temporal dependencies in input data into account, as demonstrated by earlier research using memory-based deep neural networks like Long Short-Term Memory (LSTM) and Convolutional LSTM (ConvLSTM). Examples of previous work in vision-based end-to-end autonomous driving include NVIDIA's PilotNet framework. The authors provide two categories of metrics for evaluating various end-to-end solutions: external metrics derived from models' performance on simulated test circuits, enabled by an open-source application called Behavior Metrics, and internal metrics (mean average error and mean square error).

The authors present their proposed deep learning architecture DeepestConvLSTMConv3DPilotnet (memDCCP), which merges ConvLSTM and 3D convolutional layers (Conv3D) to extract temporal information and improve control command prediction. A comparative analysis is performed with other memoryless models such as PilotNet and DeepestLSTMTinyPilotNet, as well as a memory-based model called PilotNet x3. Three experiments were conducted to evaluate the performance of the proposed and existing models and to demonstrate the superior generalization and robustness of memDCCP due to its ability to learn temporal dependencies from image sequences. Additionally, the authors advocate behavioural metrics as a valuable tool for evaluating end-to-end autonomous driving solutions, highlight the importance of external metrics alongside internal metrics, and provide project materials as open-source resources for reproducibility and further research.

3.9 Continuous Control of Autonomous Vehicles using Plan-assisted Deep Reinforcement Learning [9]

Dwivedi et al. proposes a novel plan-based deep reinforcement learning framework that utilizes a “trajectory space” alongside the typical state space to learn optimal control. The authors acknowledge the limitations of existing end-to-end approaches that recreate reflexive responses like those observed in living organisms and that are characterized by an involuntary, unplanned sequence of actions for a given stimulus. While reflexive responses have their advantages, the authors argue that conscious responses based on logical plans and behaviours are more effective and fail-safe for operating in complex real-world environments. They assume that for better generalization of end-to-end architectures, the agent must be able to derive logical plans and adaptive maneuvering strategies, thereby improving not only the performance but also the interpretability and reliability of the system.

The proposed framework includes a local planner that navigates the agent on the correct trajectory and intrinsically captures the agent's action plan. The authors define a cumulative reward function and a value function that considers the expected cumulative return for following the plan within a given time horizon. An imagination-trained network of actors and critics uses latent features from world models to predict the politics and function of state value while minimizing deviation from planned development. The paper presents a detailed description of the model architecture built on DreamerV2 and Racing Dreamer. It includes a world model learning component, a trajectory encoder, reward shaping mechanisms and an actor-critic network. The authors provide a comprehensive mathematical formulation of the plan-based RL framework and highlight the optimization problems and loss functions used to train the model.

Experimental results are presented that demonstrate the improved performance, stability and generalization of the proposed model architecture. The authors first train the agent in a simulation environment and then test it on the F1TENTH racing car. The model's generalization ability is determined by evaluating its performance on a variety of unknown racetracks. Comparative analyses are conducted with other approaches, such as parameter-tuned controllers, model-free methods, and existing learning-based methods, demonstrating the superior performance and lower collision frequency of the plan-based framework. The authors conclude by acknowledging the complexity of the autonomous driving problem and the impracticality of exploring all possible scenarios while driving. They propose that subspace estimation algorithms where predicting actions is relatively easy can allow the agent to learn actions based on existing frameworks. The trajectory-based approach presented in this article limits the directions the agent must explore, thereby minimizing the search space.

Overall, this article makes a significant contribution to the field of autonomous driving by proposing a novel plan-based deep reinforcement learning framework that effectively captures behaviour in latent space. The authors provide a comprehensive theoretical foundation, detailed model architecture and experimental validation, demonstrating the improved generalization, stability and interpretability achieved by incorporating planning information. However, the paper acknowledges the potential for future research directions, such as extending the work to multi-agent, head-to-head autonomous racing scenarios or urban autonomous driving scenarios, where additional modules may be required to control appropriate behaviour could.

3.10 Safe Reinforcement Learning for High-speed Autonomous Racing [10]

One of the earliest and most influential works in this area is that of Sadigh et al. conducted a study that deals with the application of deep neural networks for autonomous racing scenarios. Their research reveals the remarkable potential of these computational models, inspired by the neural architecture of the human brain, in processing complex sensory data and making informed decisions in real time. By leveraging massive amounts of training data and sophisticated algorithms, the authors show how Deep RL agents can acquire the ability to navigate complicated racing environments, anticipate obstacles, and optimize racing strategies. Building on this foundation, Shalev-Shwartz et al. Explore the complex dynamics of autonomous racing, encompassing various aspects such as perception, decision-making and control. Their work highlights the critical role of perception as autonomous vehicles must accurately interpret their surroundings, including lane boundaries, obstacles and the positions of other vehicles. This information is then processed by the Deep RL agent, which uses reinforcement learning techniques to make intelligent decisions regarding acceleration, braking and steering.

The groundbreaking work of Mnih et al. deals with one of the biggest challenges in autonomous racing. examines the trade-off between exploration and exploitation. Exploration refers to the agent's ability to discover new strategies and adapt to unknown situations, while exploitation involves the use of known successful strategies. The authors emphasize the importance of finding the right balance between these two components to achieve optimal performance in autonomous racing scenarios. The role of simulation environments in training and evaluating Deep RL agents is discussed by Dosovitskiy et al. highlighted. Their research highlights the importance of high-fidelity simulations that provide agents with a safe and controlled environment to learn and refine their racing skills, thereby reducing the risks and costs associated with real-world testing. These simulations can accurately replicate various racing conditions, including track layouts, weather conditions and vehicle dynamics, allowing agents to develop robust and generalizable skills.

Furthermore, Bojarski et al. the integration of Deep RL techniques with other cutting-edge technologies such as computer vision and sensor fusion. By combining visual data from cameras with information from sensors such as LiDAR and radar, autonomous vehicles can gain a more comprehensive understanding of their surroundings, improving their decision-making ability and overall performance. Safety and robustness in autonomous racing systems are also highlighted as crucial aspects by Seshia et al. Their work highlights the importance of ensuring the reliability and fail-safe operation of these systems, as even minor errors or system failures could have catastrophic consequences in high-speed racing environments.

In summary, the literature examined in this review provides a comprehensive overview of the development of deep gain autonomous racing, highlighting the notable advances that have been made in this area and the potential for further advances. As Deep RL techniques continue to evolve and integration with other emerging technologies increases, the future of autonomous racing offers exciting opportunities that pave the way for safer, more efficient and intelligent transportation systems.

3.11 High-speed Autonomous Racing using Trajectory-aided Deep Reinforcement Learning [11]

Benjamin Evans, Herman A. Engelbrecht and Hendrik W. Jordaan present a novel approach to training deep reinforcement learning agents for high-performance autonomous racing. The authors fill a significant gap in the literature by developing a method that allows Deep RL agents to perform high-speed racing using only raw LiDAR scans as input. The article begins with a comprehensive literature review comparing classical methods and end-to-end learning approaches for autonomous racing. The authors highlight the limitations of classical methods that rely on explicit state estimation and localization, thereby limiting their flexibility and adaptability to environmental changes. Conversely, end-to-end Deep RL agents have shown success in controlling vehicles at low speeds using raw sensor data such as LiDAR scans but have shown limited performance in high-speed racing scenarios.

The authors present a novel learning formulation called Trajectory-Aided Learning (TAL), which incorporates an optimal trajectory (race line) into the reinforcement learning process. This approach is motivated by literature suggesting that classical solutions leveraging trajectory optimization and path tracing approaches achieve powerful racing results. By integrating the optimal trajectory into the reward signal, the authors aim to train Deep RL agents to select feasible speed profiles and roughly track the ideal line, enabling high-speed autonomous racing. The paper provides a comprehensive evaluation of the proposed TAL approach using open source F1Tenth Simulator on four different maps. The authors conduct a series of experiments to examine the impact of top speed on agent performance, compare lap times and completion rates of agents trained with TAL and a baseline formulation, analyse the trajectories and speed profiles selected by the agents, and their approach to compare classical methods and related works in literature.

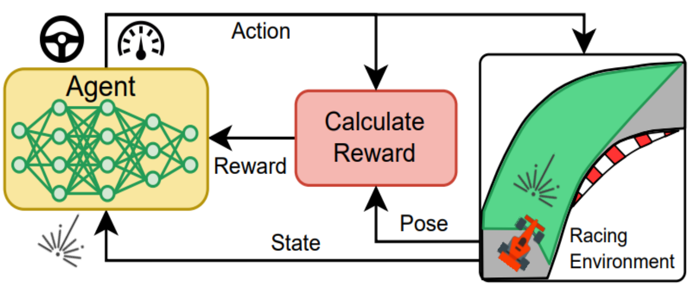


Fig-8: Processing Pipeline

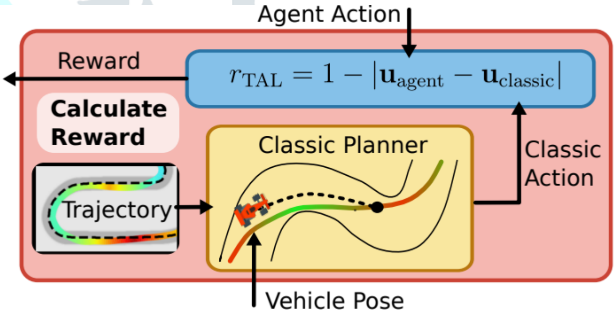


Fig-9: The Trajectory-Aided Learning Reward Function

The results show that the TAL formulation significantly improves the round completion rate of agents at high speeds compared to the baseline formulation. By analysing trajectories and speed profiles, the authors show that TAL agents learn to select speed profiles similar to the optimal trajectory by slowing down on curves and speeding up on straights. This behaviour contrasts with the baseline agents, which tend to maintain near-constant high speeds, resulting in higher slip angles and increased accident probability. Furthermore, the authors compare the performance of TAL agents with a maximum speed of 8 m/s with that of classical schedulers and related approaches in the literature. The results suggest that the TAL agents achieve faster lap times than previous methods, demonstrating the effectiveness of their approach in training agents for high-performance autonomous racing.

The authors' contributions to the field of autonomous racing and the application of Deep RL techniques are significant. By leveraging domain knowledge and incorporating classical components into the learning formulation, they have demonstrated a successful approach to training agents for high-performance control tasks. The paper paves the way for further research in this area, including transferring learned policies to physical vehicles and exploring the optimal integration of trajectories into other applications, such as drone control.

3.12 Improving Environment Robustness of Deep Reinforcement Learning Approaches for Autonomous Racing Using Bayesian Optimization-based Curriculum Learning [12]

Banerjee et al. Explore a novel application of Bayesian optimization to automatically generate curricula to improve the robustness of reinforcement learning (RL) policies for an autonomous racing task with obstacles. Deep RL approaches have shown promising results in various robotic tasks, including autonomous driving and robot manipulation (Zeng et al., 2018; Amini et al., 2020). However, a key challenge is to develop RL agents that generalize well to novel environments beyond their training distribution, which is critical for deploying such systems in unstructured, real-world environments. One approach to improving generalization is curricular learning, in which the agent is exposed to a curated sequence of environments during training (Bengio et al., 2009). Previous work examined curriculum learning for navigation tasks such as maze environments (Dennis et al., 2020), the LunarLander domain (Song & Schneider, 2022), and street driving scenarios (Anzalone et al., 2021; Ozturk et al., 2021). However, curriculum learning for autonomous racing with static obstacle avoidance has not been extensively studied.

Authors establish baselines by training a Proximal Policy Optimization (PPO) agent in a standard environment and using a manually designed curriculum progression from easy to difficult scenarios. They then formulate curriculum learning as a probabilistic derivation of the relationship between curriculum and reward function using a Gaussian process model within the Bayesian optimization framework. The key idea is to model the mapping of curricula to environmental rewards as an unknown function and use Bayesian optimization to iteratively discover high-performing curricula. Empirical results in a modified Car Racing environment show that the Bayesian optimization approach can discover curricula that achieve higher rewards across different levels of difficulty compared to baselines. Notably, the learned curriculum outperforms both the standard curriculum and the manually created curriculum on metrics such as average reward, collision rate, and distance travelled in challenging evaluation environments with high road curvature and obstacle density.

A strength is the generality of the Bayesian optimization formulation, making it potentially applicable to curriculum learning for other RL domains beyond racing. However, the work is limited to the specific environmental parameters of road curvature and obstacle density. Future work could include additional variables such as dynamic obstacles or weather conditions. Another limitation is using a single random seed for training, which risks overfitting the curriculum search for stochasticity effects. Averaging across multiple seeds could lead to more reliable curriculum estimates. Alternative reward formulations in the curriculum could also be explored, such as considering rewards for midterm assessment.

Overall, the paper presents a promising direction for developing autonomous racing agents that can robustly generalize across different environmental conditions. The Bayesian optimization framework provides an automated way to discover sophisticated curricula that could improve real-world safety and performance. With the introduction of autonomous racing systems, such generalization capabilities are becoming increasingly important.

IV. Future Enhancements

The advancement of deep reinforcement learning in autonomous racing provides fertile ground for further exploration and improvement. Based on the comprehensive analysis of the literature, several key areas for future improvement emerge:

4.1 Bridging the reality gap.

A key challenge in Deep RL-based autonomous racing is the reality gap, where models trained in simulation environments struggle to transfer to real-world scenarios. Future research should focus on robust sim-to-real transfer techniques such as domain randomization and transfer learning to enable agents to seamlessly adapt to real-world dynamics. Techniques such as Sim2Real could be further explored to improve the transferability of learned policies from simulation to physical vehicles [2][4].

4.2 Multi-Agent strategies

Realistic racing scenarios often involve interactions between multiple autonomous vehicles. Developing sophisticated multi-agent strategies that facilitate negotiation, collaboration and collision avoidance is critical. Such strategies would enable autonomous vehicles to navigate complex racing environments safely and efficiently, paving the way for more realistic racing simulations [4][9].

4.3 Advanced Perception Modules

Integrating advanced perception modules, including semantic segmentation and object detection, could significantly improve the situational awareness and decision-making ability of autonomous racing agents. By providing more detailed and accurate environmental information, these modules would enable agents to make more informed decisions and control demanding racing scenarios more precisely [5][9].

4.4 Safety-critical applications outside of motorsports

While much of the current focus is on motorsports, the important Principles and techniques developed in autonomous racing find broader application in safety-critical areas such as autonomous driving in the city. Future research should explore adapting Deep RL-based autonomous racing techniques to urban environments where factors such as traffic congestion, pedestrian interactions, and complex road infrastructure present unique challenges [10].

By addressing these key areas of focus, future research in deep reinforcement learning for autonomous racing can drive further advancements in safety, efficiency, and intelligence in both simulated and real-world racing environments.

V. Conclusion

In conclusion, this study has examined the rapidly evolving landscape of deep reinforcement learning for autonomous racing. The literature reviewed demonstrates the remarkable progress in developing autonomous racing agents capable of outperforming human drivers in simulated environments. Key contributions include novel algorithms tailored to racing tasks such as A3C, SAC, and plan based Deep RL frameworks that integrate optimal trajectories to improve performance and generalization. The integration of advanced perception modules such as sensor fusion and 3D convolutional networks has enabled agents to process complex sensory data and operate under realistic racing conditions. Despite the impressive successes, some challenges remain. For real-world use, bridging the gap between simulation and reality and ensuring robust performance in the physical world is critical. The development of multi-agent strategies for negotiating with other vehicles and scaling algorithms for urban scenarios are important future research directions. Furthermore, safety aspects and the interpretability of learned policies are crucial for the widespread adoption of autonomous racing technology.

Ultimately, the development of deep reinforcement autonomous racing holds immense potential for applications beyond motorsports. The techniques developed in this area can influence the development of general autonomous driving systems and lead to safer and more efficient transport solutions. As Deep RL algorithms continue to advance and computing resources become more powerful, we can expect autonomous racing agents to outperform human performance in an ever-widening range of scenarios.

VI. ACKNOWLEDGMENT

We would like to thank Mrs. P. Navya for her valuable comments and suggestions to improve the quality of the paper. We are also grateful to Dr. K. Shirisha Reddy for helping us review our work regularly. We would also like to thank the Department of Computer Science Engineering (AI and ML), VBIT Hyderabad for providing us Jetson Nano Boards and all the help we needed.

REFERENCES

[1] Teigar, Henry, Miron Storožev and Andrej Karpathy. "2 D Racing game using reinforcement learning and supervised learning." (2018).
[2] Jaritz, Maximilian, Raoul de Charette, Marin Toromanoff, Etienne Perot and Fawzi Nashashibi. "End-to-End Race Driving with Deep Reinforcement Learning." 2018 IEEE International Conference on Robotics and Automation (ICRA) (2018): 2070-2075.
[3] Remonda, Adrian, Sarah Krebs, Eduardo Veas, Granit Luzhnica and Roman Kern. "Formula RL: Deep Reinforcement Learning for Autonomous Racing using Telemetry Data." ArXiv abs/2104.11106 (2021): n. pag.
[4] Kiran, Bangalore Ravi, Ibrahim Sobh, Victor Talpaert, Patrick Mannion, Ahmad A. Al Sallab, Senthil Kumar Yogamani and Patrick P'erez. "Deep Reinforcement Learning for Autonomous Driving: A Survey." IEEE Transactions on Intelligent Transportation Systems 23 (2020): 4909-4926.
[5] Guckiran, Kivanc & Bolat, Bulent. (2019). "Autonomous Car Racing in Simulation Environment Using Deep Reinforcement Learning." 1-6. 10.1109/ASYU48272.2019.8946332.
[6] Fuchs, Florian, Yunlong Song, Elia Kaufmann, Davide Scaramuzza and Peter Dürr. "Super-Human Performance in Gran Turismo Sport Using Deep Reinforcement Learning." IEEE Robotics and Automation Letters 6 (2020): 4257-4264.
[7] Samak, Chinmay Vilas, Tanmay Vilas Samak and Sivanathan Kandhasamy. "Autonomous Racing using a Hybrid Imitation-Reinforcement Learning Architecture." ArXiv abs/2110.05437 (2021): n. pag.
[8] Blanco, Sergio Paniego, Sakshay Mahna, Utkarsh Aashu Mishra and José María Cañas. "Memory based neural networks for end-to-end autonomous driving." ArXiv abs/2205.12124 (2022): n. pag.
[9] Dwivedi, Tanay & Betz, Tobias & Sauerbeck, Florian & P.V., Manivannan & Lienkamp, Markus. (2022). "Continuous Control of Autonomous Vehicles using Plan-assisted Deep Reinforcement Learning." 10.23919/ICCAS55662.2022.10003698.
[10] B. D. Evans, H. W. Jordaan, and H. A. Engelbrecht, "Safe reinforcement learning for high-speed autonomous racing," Cognitive Robotics, vol. 3, pp. 107–126, 2023.
[11] Evans, Benjamin David, Herman Arnold Engelbrecht and Hendrik Willem Jordaan. "High-Speed Autonomous Racing Using Trajectory-Aided Deep Reinforcement Learning." *IEEE Robotics and Automation Letters* 8 (2023): 5353-5359.
[12] Banerjee, Rohan, Prishita Ray and Mark Campbell. "Improving Environment Robustness of Deep Reinforcement Learning Approaches for Autonomous Racing Using Bayesian Optimization-based Curriculum Learning." ArXiv abs/2312.10557 (2023): n. pag.
[13] Bradford, Alastair, Grant van Breda and Tobias Fischer. "Racing With ROS 2 A Navigation System for an Autonomous Formula Student Race Car." ArXiv abs/2311.14276 (2023): n. pag.
[14] Salvaji, Aakaash & Taylor, Harry & Valencia, David & Gee, Trevor & Williams, Henry. (2023). "Racing Towards Reinforcement Learning based control of an Autonomous Formula SAE Car."
[15] Cai, Peide & Wang, Hengli & Huang, Huaiyang & Liu, Yuxuan & Liu, Ming. (2021). "Vision-Based Autonomous Car Racing Using Deep Imitative Reinforcement Learning." IEEE Robotics and Automation Letters. 6. 7262-7269. 10.1109/LRA.2021.3097345.
[16] Evans, Benjamin & Betz, Johannes & Zheng, Hongrui & Engelbrecht, Herman & Mangharam, Rahul & Jordaan, Hendrik. (2022). "Accelerating Online Reinforcement Learning via Supervisory Safety Systems." 10.48550/arXiv.2209.11082.
[17] S. Ju, P. van Vliet, O. Arenz and J. Peters, "Digital Twin of a Driver-in-the-Loop Race Car Simulation With Contextual Reinforcement Learning," in *IEEE Robotics and Automation Letters*, vol. 8, no. 7, pp. 4107-4114, July 2023, doi: 10.1109/LRA.2023.3279618.
[18] Agrawal, Avinash. J., Rashmi R. Welekar, Namita Parati, Pravin R. Satav, Uma Patel Thakur, and Archana V. Potnurwar. 2023. "Reinforcement Learning and Advanced Reinforcement Learning to Improve Autonomous Vehicle Planning". *International Journal on Recent and Innovation Trends in Computing and Communication* 11 (7s):652-60. https://doi.org/10.17762/ijritcc.v11i7s.7526.