A Review of Reinforcement Learning: Current Trends and Future Prospects in Autonomous Systems

Abstract

This review focuses on the use of reinforcement learning (RL) for autonomous systems and current trends and future prospects. It is therefore the intended goal to critically evaluate the concept of RL for improving autonomous decision making with focus on current and emerging issues including; sample efficiency, scalability, and safety. This review methodology is a synthesis of 10 studies which has been conducted between the years 2021 and 2024. However, these are some of the challenges that seem to plague RL even as it has potential to be used in realistic applications such as robots, self-driving cars and smart grid. The review also opines that due to developments of algorithms, computer intrinsics and safety mechanism, RL perhaps holds the key to the future for autonomous systems.

Keywords: Reinforcement Learning, Autonomous Systems, Challenges, Future Prospects.

UNDER PEER REVIEW

Introduction

Reinforcement learning (RL) is one of the most influential areas in the sphere of machine learning that contributed to the growth of autonomous systems. RL is a learning model where an agent is put in a setting referred to as an environment, with the goal of finding actions that give the biggest rewards in future (Sutton &Barto, 2018). The principal concept of RL is applicable here because RL trains the machine to act in certain ways by trial, not rule and not supervised learning. Thus, RL has found applicability across numerous domains such as robotics (Levine et al., 2018), self-driving cars (Kiran et al., 2020), and smart grids (García et al., 2019) and has boosted the learning prospects of those domains.

Now there are self-driving cars and robots, smart structures, and all other kinds of system that are an inalienable part of the modern world. These systems must function autonomously in order with optimal performance in their complex operational settings, this explains why they apply RL principles. However, there are some challenges associated with utilizing RL in autonomous mechanisms. The one of which one can think of considerable amounts is sample inefficiency since RL algorithms require an extremely large number of iterations to choose satisfactory policies (Hester et al., 2020). But the limitation is that scaling is still challenging as the applications of RL algorithms depend on the enormous computational resources to solve complex problems (Mnih et al., 2015). Furthermore, safety becomes a significant challenge when deploying RL in safety-sensitive areas such as autonomous vehicles become a crucial barrier (Akrour et al., 2020).

To this some recent attempts have been made to surmount these restrictions. For instance, there has been the introduction of model-based RL strategies to address sample inefficiency based to the learned models about the environment. Further, it has been made easier to come up with a short and more efficient algorithm via deep RL, a combination of RL with deep neural networks [14]. Other innovations Yarats et al. (2020) such as transfer learning and multi-agent reinforcement Eric Zhou and Go (2019) also pledged that scalability and convergence rate can be improved in different applications.

Considering the continuous enhancements with algorithms as well as the hardware of the RL this has a promising future in autonomous systems. But problems concerning theory's generality, readability, and application persist as well. The future of RL in improving autonomous systems will largely depend on addressing these issues as well as guaranteeing the systems' safety and stability in uncertain conditions. Thus, the integration of RL in AS is expected to open vistas for increasing efficiency, automation, and adaption of systems as more and more standalone systems come into use.

Therefore, this paper overviews the current state of developments in RL and considers future ideas, concerning how it can help meet new demands for different autonomous systems. In the light of this analysis, the goal is to identify: where RL is now as well as what is still to come in order to review possibilities and illustrate how RL implements the further autonomous systems' evolution.

UNDER PEER REVIEW

UNDER PEER REVIEW

Reinforcement Learning

Reinforcement learning is a rapidly developing direction. In the domain of machine learning, it has resulted in massive advancement over the past decade – be it the development of artificial players that can defeat humans in strategically demanding games like Go (Silver D et al., 2017) and StarCraft (Vinyals O et al, 2019). In cognitive neuroscience, RL models have been employed to acceptably describe various untold latent learning-related phenomena, at the behavioral level (Eckstein MK et al., 2020; Master SL et al., 2020) and also the neural level (Maes EJP et al., 2020). However, the hope that RL can point to reasonable and predictive latent variables obscures variability in what RL variables index, even in cognitive neuroscience. The success of RL has bred an illusion of omnipotence where RL can observe the brain and behavior and dissect out fundamental features. While this notion develops as RL methods become more widely used, it may result in overgeneralization or overinterpretation of results.

Here, we claim that a more refined perspective is more appropriately justified by the evidence and by theory. We have thus provided a review of how each subfield applies RL and what parts are tended to share and what parts are unique. We then continue to identify and discuss where the use of cognitive neuroscience may have gone wrong, then establish that where RL is lodged, it remains a valuable tool for the field.Reinforcement Learning (RL) has emerged as a powerful tool in addressing complex decision-making problems across various domains, including robotics, autonomous driving, and clinical decision support systems (Ye Zhang et al., 2024; Jingda Wu et al., 2024). RL leverages the concept of agents learning from interactions with their environment to maximize cumulative rewards, enabling adaptive and autonomous behavior. However, scalability, sample efficiency, and safety remain persistent challenges limiting its broader application (Benjamin Kommey et al., 2024). Recent advancements have aimed at enhancing RL's robustness and applicability, such as integrating hierarchical frameworks, multi-agent systems, and deep learning techniques to improve decision-making in dynamic and uncertain environments (Min Hua et al., 2021; Qi Liu et al., 2023).

Moreover, RL's role in addressing real-world problems, such as unsignalized intersections in autonomous vehicles and power system stability, has been highlighted, albeit with limitations in scalability and real-world validation (Mohammad Al-Sharman et al., 2024; Mohamed SadokMassaoudi et al., 2023). In clinical decision support systems, explainable RL methods have been proposed to increase transparency and user trust, showcasing its potential in sensitive fields (Anna MarkellaAntoniadi et al., 2021). These studies collectively underscore RL's transformative potential while emphasizing the need for further research to address ethical concerns, enhance interpretability, and develop scalable solutions suitable for practical applications.

UNDER PEER REVIEW

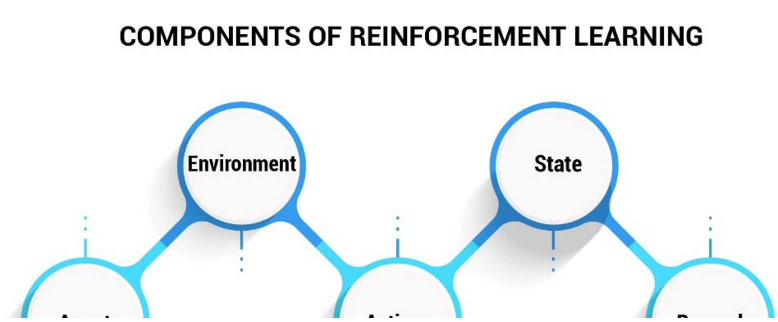

Figure 1: The diagram illustrates the core components of reinforcement learning, where an agent interacts with an environment by taking actions, transitioning through states, and receiving rewards to learn optimal behavior.

UNDER PEER REVIEW

Current Trends in Reinforcement Learning

In the past few years, Reinforcement Learning (RL) has reached much of its development as a core methodology for fulfilling self-organizing and decision-making in different fields. RL has been increasingly popular and effective when combined with new technologies such as deep learning (Haarnoja et al., 2021). This section presents four progressive research trends that are actively influencing the modern picture of RL: sample efficiency, transfer learning, safe reinforcement learning, multi-agent reinforcement learning.

Among the new patterns is a progression in the methods to step up the sample efficacy which remains one of the main drawbacks of traditional RL approaches. Contemporary approaches to RL incorporate model-based techniques to reduce the steps required for an agent to define the optimal actions. For instance, the model-based RL methods such as MuZero have established their ability to outperform previous outcomes, as well, to learn on the environmental dynamics and the best move (Schrittwieser et al., 2020). Other algorithms applying experience replay and prioritized sampling remains in refining the learning process based on the retention of the past best experience.

Another area that has also been adopted in the recent past is transfer learning this aims at improving the RL application by reusing information that was acquired in the past. In order to compress the learning process and to enhance the learning performance in scenarios, where data is not easily available, transfer learning involves the transfer of acquired policies or features from a different domain. There are other strategies that are being exploited to apply the RL models in a number of shifts and these include domain adaptation and meta-RL (Zhang et al., 2021). These improvements have been largely useful in robotics because such policies take time to update to new physical settings.

Safe RL has emerged as another important trend for using in high-risk areas like medical, financial and auto mobile industries. Modern RL algorithms use safety constraints and reward shaping to ensure that an agent acts uniformly and eliminate unwanted actions (Turchetta et al., 2021). Furthermore, the advancement of sparse RL has also helped agents to accomplish the set missions while being constrained by inadequate resource utilization (Zhou et al., 2021). These advances are important for the application of RL systems in practical applications that are safety-critical.

As the development of MARL proceeded, the usage of RL extended to areas that include multi-agent collaboration or rivalry. MARL methods are being applied in areas that require traffic management, incorporation of game elements, and decentralized resource management (Zhang et al., 2021). The communication protocols and shared learning frameworks have been enhanced to increase the scalability and dependency of multi-agent systems for more sophisticated and diverse interactions.

Finally, the RL studies are shifting from performance to interpretability and explainability, as the deep RL makes decisions in a way that is not easily understandable. To enhance the interpretability of the developed RL models, visualization tools and XAI approaches are combined (Puiutta&Veith, 2021). These efforts are meant to close the gap between highly advanced algorithms and how they could be used in real-life decision-making contexts.

UNDER PEER REVIEW

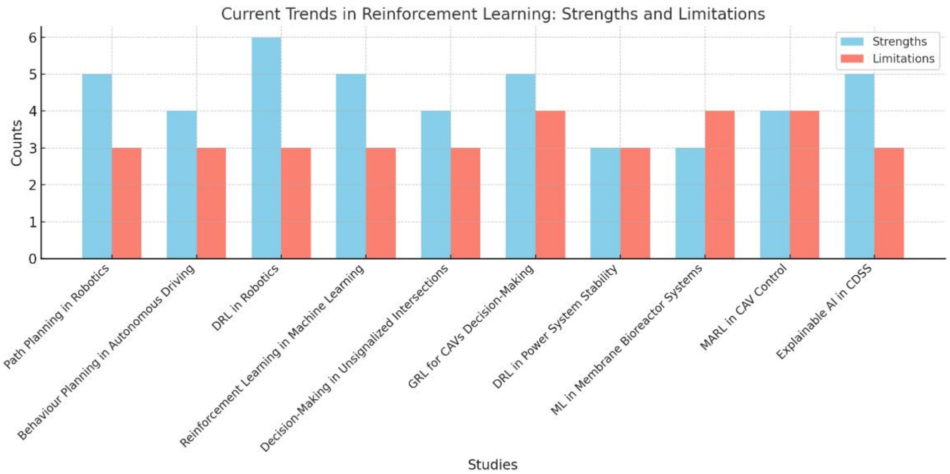

Figure 2: illustrating the strengths and limitations of the studies titled "Current Trends in Reinforcement Learning.

UNDER PEER REVIEW

Literature Review

Ye Zhang et al, (2024) Path planning is still one of the pivotal facets for robotics application areas like autonomous driving, minimally invasive surgery, and delivery robot. In this review, we first outline the drawbacks of traditional path planning techniques and the latest advancements in DRL-based path planning techniques. Then, it presents a literature review on constructing important components of DRL methods in recent years to help readers understand the fundamental concepts of DRL studies as well as the reasoning and concerns from a practical standpoint. During the practical training, there may be concerns such as little reward and the exploration–exploitation dilemma; thus, the paper summarizes the methods for improving training performance and optimizing the DRL path planning. Finally, the paper concludes with present day practical path planning applications' limitations, challenges and identifies the future research areas.

Jingda Wu et al., (2024) Autonomous driving (AD) refers to the capability of a vehicle to travel in traffic space and is an innovative solution for increasing transport efficiency, but it greatly depends on BP. Reinforcement learning (RL) turns into a crucial approach to designing these BP strategies. Thus, this paper provides a systematic review of RL-based BP strategies with focus on the development from 2021 to 2023. We comprehensively review and synthesize the literature, focusing on the changes in RL-based BP paradigms. To fill this gap, it is possible to identify the history of attempts at overcoming the practical problems encountered by AVs owing to new forms of RL. In order to assist readers, we provide a quantitative analysis that details the amount and variety of new RL configurations in recent years and identifies common patterns. Further, we discuss emerging near-term issues and possible future trends concerning the development of RL-based BP in AD. These directions include mitigating safety risks, developing sustained learning paradigms, improving data effectiveness, advocating for vehicular cloud networks, incorporating LLMs, and improving ethical concerns.

Saksham &Chhavi Rana (2024) Deep Reinforcement learning (DRL) is the disruptive technology cutting through the horizon of Robotics and Autonomous Systems that has an approach of learning by reaching out to the environment. This chapter will also propose a panoramic view of the current state of the research conducted in the area of DRL and robotics. The literature is reviewed in a comprehensive and integrative manner, focusing on advances, methods, uses, and issues in this rapidly evolving domain. Therefore, this chapter is broken down into thematic sections such as the basic principles of DRL and DRL in robot control and navigation, object manipulation, as well as self-driving cars. It describes several general techniques for investigation, its performance is compared with actual issues and scenarios of Robotics applications. The ability of DRL in training robots to work independently and flexibly is illustrated through case studies and several practical applications can be observed. The complexity of hardware systems and settings used in these applications is described, which casts light upon real-life factors affecting DRL

UNDER PEER REVIEW

implementation in practice. Consequently, the present issues and limitations are also discussed in this chapter including, sample complexity, safety, and limitation of scalable DRL for robotics. Nevertheless, in presenting these challenges this chapter provides insights into the future research opportunities and new trends in evaluation metrics for DRL algorithms in the context of robotics. Thus, this chapter can be also useful for researchers, practitioners, and enthusiasts who focus on the integration of DRL with robotic systems. It sums up the current understanding, points to the significant achievements and identifies the emerging opportunities and challenges to enable robotics and autonomous systems to grow in the age of machine learning and artificial intelligence.

Benjamin Kommey rt al., (2024) Reinforcement learning (RL) a subfield of machine learning, is not only fast growing in prominence but also found utility far from its conventional use in gaming systems. Various subfields of reinforcement learning, such as deep reinforcement learning and multi-agent reinforcement learning, are also rapidly growing. In this paper, a comprehensive survey on the field is given based on the perspective of Machine Learning (ML). It starts with giving an historical overview of the field, then goes on to offer a theoretical overview of the field. It then goes over the main RL problems and the strategies taken by various subfields before going over the state-of-the-art strategies. Some use cases in reinforcement learning are listed below and their feasibility and applicability tested. The paper concludes with pointing out some of the existing gaps or research questions in the field.

Mohammad Al-Sharman et al. (2024)state that self-driving cars at unsignalized intersections are a difficult application of machine learning because of the conflicts incurred concerning the administration of multiagent systems with a high level of unpredictability. Decision-making for this safety-critical environment entails an understanding of the various abstractions involved in learning driving behaviours for effective vehicle movement. In this survey, we are interested in identifying cutting-edge approaches adopted in decision-making systems with a special focus on the RL integrated with deep learning traversing policies at unsignalized intersection areas. The reviewed schemes differ in the proposed driving scenario, in the assumptions made for the used intersection model, in the addressed challenges, and the learning algorithms used. We have mentioned comparisons for such

UNDER PEER REVIEW

techniques to show their drawbacks and advantages. From the findings of this study, it is apparent that a sound decision-making system for dealing with real-world unsignalized intersections remains nonexistent. In addition to our in-depth analysis and discussion of the topic, the interested players are encouraged to consider the following research objectives: According to our suggestions, non-overcautious as well as safe but feasible decision-making architectures can be trained and validated in actual unsignalized intersection conditions.

Qi Liu et al., (2023) Understanding the behavioural patterns of Connected and Autonomous Vehicles (CAVs) is very important for the safety and better performance of Intelligent Transport Systems in the future. However, to get to a condition of full autonomy, the period that involves mixed traffic with CAVs and human-driven cars needs time. Therefore, collaborative decision-making technology for CAVs is required to produce desired driving conduct to improve the conditions of mixed autonomy traffic. In the last few years, the DRL technique has been widely used in solving decision-making issues efficiently. Nevertheless, with the coming of computing technology, graph reinforcement learning (GRL) methods have shown that they have great potential to enhance the decision-making of CAVs, especially in representing the mutual impacts of vehicles and dynamic traffic scenarios. Therefore, to promote the research of GRL-based methods for Autonomous Driving, this paper suggests reviewing the GRL-based method for the decision-making technologies of CAVs. Firstly, the general GRL framework is introduced at the initial stage to get a brief idea about the decision-making technology. Next, this paper compares the methods of GRL-based decision-making technologies with the construction methods of mixed autonomy traffic, representation graphs for the driving environment, and other relevant studies on graph neural networks (GNNs) and DRL for decision-making in autonomous driving. Furthermore, validation methods are aggregated to enable one to come up with simpler ways of testing decision-making methods. Last but not least, the limitations and research opportunities in utilizing the GRL-based decision-making methods are discussed.

Mohamed SadokMassaoudi et al., (2023) The rise of inverter-interfaced system-level resources has affected the electrical stability of power systems in a very significant manner. Integration of photovoltaic and wind power systems into the grid has brought unpredictable risks to the electricity industry. Recent DRL advancements have been in the research focus

UNDER PEER REVIEW

over the last few years underpinning its potential role in improving PS stability (PSS). The DRL architecture, which is widely adopted, learns from the dynamism inherent in PSs and generates near-optimal actions for a PSS. This article offers a thorough discussion of the emerging research approaches to DRL to derive PSS policies considering the characteristics of the power grid system. Besides, this paper provides a review of the theoretical benefits and the main pros and contras of the new DRL methods as effective tools for OPF. As for all the methods described, the discussion on their limitations, the challenges of the research methodology in large-scale PSS, and the opportunities are provided. This paper will seek to assist research in this area of DRL algorithms to adopt PSS against unseen faults and various PS topologies.

Z.Frontistis et al.,(2023) This paper provides an overview of the current state of using ML in MBRs which is a relatively new technology in AWT. The review is centred on the application of ML techniques for the prediction of membrane fouling, the control of the fouling system, and the detection of fouling faults, to create new cleaning strategies. Major types of ML algorithms include artificial neural networks, support vector machines, random forest, and reinforcement learning; their capabilities and weaknesses in sophisticated wastewater application processes are highlighted. The primary barriers to the reproduction of ML are listed: data quality, interpretability and transferability. Lastly, future research gaps are provided, such as integrating ML with big data, IoT, and creating new hybrid models. The review also stresses the necessity of a transdisciplinary approach and funding for data management; it alludes to the need for introducing new policies related to data protection and security. If one solves these challenges, integration of ML into MBRs can push performance and energy utilization up with further fruitful results for water treatment.

Based on **Min Hua et al. (2021),** CAVs have the potential to meet or provide the solution for future transportation demand through Balancing, Safe and Green Systems. However, controlling CAV control it is difficult because it requires a highly interconnected system and coordination in the network of the connected vehicle. Consequent to the recent development of exciting techniques in the treatment of difficult problems such as automobility and autonomy in self-driving cars and robotics, as well as human vehicle interface, MARL has proved useful for enhancing the performance of CAVs. However, it does not have a recent evaluation of conventional MARL algorithms for CAVs. In light of such scarcity, the following paper undertakes a systematic review of MARL in CAV control. The paper begins with the state definition of MARL along with problems it seeks to solve and its inherent capabilities of dealing with multiple people. It then provides a critical evaluation of the MARL application in several control areas for CAVs with an emphasis on the risky aspects of driving such as platooning control, lane changing and unsignalized intersections. Further, the paper discusses some of the most important simulators that are crucial for creating and evaluating MARL algorithms. Finally, it reviews the current issues in applying MARL for CAV control, such as macro-micro learning, cooperation, incorporation of other traffic types, and sim2real difficulties. Possible approaches include hierarchical MARL, decentralized MARL, adaptive interactions, as well as offline MARL. This paper suggests multi-agent reinforcement learning as an approach to handling the interactions of connected and automated vehicles about platooning, lane changing and intersections among others. MARL can be expected to provide more efficient solutions than the outmoded structures in real-world applications, thereby resulting in better traffic management and safety and optimal usage of fuel to provide more dependable transport systems. The paper also offers an overview of MARL algorithms and simulation platforms at the current moment, which is useful for putting these complex control methodologies into practice. However, applying MARL in real-life CAV systems is still limited and there are challenges which include issues of vehicle interaction in real-life scenarios and managing many interacting systems where some of the systems involve human-driven vehicles. Further studies are required to resolve such issues and confirm the applicability of this method in various and varying traffic environments. This paper may prove to be useful for practitioners and researchers striving to create more accurate and robust CAV systems shortly.

Antoniadi A.M. et al. (2021) Machine Learning and other Artificial Intelligence tools AI applications generally hold a tremendous near and long-term potential for revolutionizing almost every segment of medicine. However, in many applications outside the medical field, the opacity of AI applications has become an issue in the latest years. This is most conspicuous where users are expected to decipher the outcomes generated by AI systems. XAI gives a reason that enables a user to know why a particular output has been obtained by the system. The output is then used to interpret data within a specific context. Additionally, Clinical Decision Support Systems (CDSSs) are another area that is lacking in XAI. These systems assist medical practitioners in clinic decision-making, and in the case of lack of explainability, could cause under or over-reliance situations. Explanations of how this recommendation is going to be arrived at will help practitioners to be much more precise and in some cases save lives. The importance of XAI in CDSS and the medical field is further heightened by what the medical field and many are facing: the requirement of ethical and fair decision-making and that; An AI trained with historical data can be a positive reinforcement of historical actions and bias which need to be revealed. We therefore conducted a systematic literature review of work done so far in the use of XAI in CDSS. The least common is XAI-enabled CDSS for text analysis while tabular data processing XAI-enabled systems are the most common. More work is demanded from developers in terms of local explanations compared to post-hoc and ante-hoc techniques, as well as model-specific and model-agnostic methods were almost equally represented. Some of the advantages associated with the use of XAI include but are not limited to; it may improve clinicians' decision confidence or it can generate hypotheses about causality which can lead to increased acceptability and trust in the system as well as the possibility of its integration in clinical processes. However, we identified a complete absence of XAI as a concept in the context of CDSS and a general absence of user studies investigating the requirements of clinicians. Some suggestions are made for the use of XAI in CDSS and some possibilities, issues, and research directions are discussed.

Table 1 :List of reviewed papers

Ref	Title	Limitations	Strengths	Aim of the Study	Result	Datasets Used	ML Algorithms	Techniques Used	Discussion
[1]	Path Planning in Robotics (Ye Zhang et al., 2024)	Sparse rewards; exploration-exploitation balance; practical training challenges	Comprehensive review of DRL in path planning; enhancement methods for DRL; practical focus	Summarize DRL methods and propose enhancements for path planning in robotics	The review highlights DRL's potential but notes gaps in optimization and training	Not explicitly mentioned	Deep Reinforcement Learning (DRL)	Review of practical DRL training methods and enhancements	This study emphasizes the potential of DRL in robotics but identifies gaps in practical implementation, particularly in addressing sparse rewards and balancing exploration-exploitation during training. It suggests that future research focus on improving training methodologies to enhance DRL's real-world applicability.
[2]	Behaviour	Safety	Quantitative	Review RL-	Future	Not	Reinforceme	Quantitative	The study

	Planning in Autonomous Driving (Jingda Wu et al., 2024)	vulnerabilities; ethical concerns; data inefficiency	analysis; novel categorization of RL-based BP; future directions highlighted	based behaviour planning strategies in autonomous driving from 2021–2023	directions: safety, collaboration, and integration of large language models	explicitly mentioned	nt Learning (RL)	analysis; focus on collaboration and ethical considerations	underscores the significance of RL in autonomous driving, particularly in behavior planning. It identifies safety, ethical considerations, and data inefficiency as challenges and recommends integrating collaboration techniques and large language models to address these gaps.
[3]	DRL in Robotics (Saksham &Chhavi Rana, 2024)	Sample efficiency; scalability; and safety concerns	Thematic organization; integration of hardware and software considerations; case studies	Overview of DRL in robotics, focusing on control, navigation, and object	Demonstrates DRL's autonomy and adaptability but highlights	Not explicitly mentioned	Deep Reinforcement Learning (DRL)	Case studies, real-world applications, and challenges in DRL	Highlighting DRL's autonomy and adaptability in robotics, the study discusses

				manipulation	scalability and safety challenges				challenges like scalability and safety concerns. It stresses the need for integrating DRL with hardware-software systems and applying case studies to validate effectiveness.
[4]	Reinforcement Learning in Machine Learning (Benjamin Kommey et al., 2024)	Scalability issues; insufficient exploration of real-world applications	Historical perspective; an exhaustive list of RL applications	Review RL advancements and highlight core problems in ML applications	Highlights RL scalability issues but emphasizes it's expanding applicability across domains	Not explicitly mentioned	Reinforcement Learning (RL)	Historical analysis; categorization of RL applications	This review identifies RL's scalability issues and limited real-world application exploration as key challenges. It emphasizes the expanding applicability of RL across domains and suggests

UNDER RE

									focusing on improving scalability through advanced algorithms.
[5]	Decision-Making in Unsignalized Intersections (Mohammad Al-Sharman et al., 2024)	Limited real-world validation; challenges in non-overcautious safe navigation	Focused analysis on unsignalized intersections; comparison of algorithms	Explore RL and deep learning techniques for autonomous driving at unsignalized intersections	Identifies challenges and suggests future directions for robust decision-making	Not explicitly mentioned	Reinforcement Learning (RL); Deep Learning	Comparison of RL-based decision-making techniques	The study points out challenges in navigating unsignalized intersections without being overly cautious. It compares RL-based techniques and suggests developing robust algorithms for safe decision-making in complex scenarios.
[6]	GRL for CAVs Decision-Making (Qi Liu et al., 2023)	Limited application in real-world mixed traffic scenarios; validation	Proposes GRL framework; explores GNNs and dynamic modelling for	Review GRL methods for decision-making in connected	Suggests GRL as a solution for improving safety and efficiency	Not explicitly mentioned	Graph Reinforcement Learning (GRL); Graph Neural	GRL framework development; graph representation of the	This work discusses the potential of GRL in enhancing decision-

		challenges	decision-making	and autonomous vehicles	in mixed traffic		Networks (GNNs); DRL	driving environment	making for connected and autonomous vehicles, particularly in mixed traffic conditions. However, it highlights validation and application limitations in real-world scenarios. Future work could focus on improving the adaptability of GRL frameworks.
[7]	DRL in Power System Stability (Mohamed SadokMassaoudi et al., 2023)	Scalability challenges; limited exploration of unseen faults	Explores DRL for optimal power flow; assesses grid uncertainties	Review DRL techniques for stabilizing power systems	Emphasizes DRL's role in grid stability but notes scalability and unseen faults	Not explicitly mentioned	Deep Reinforcement Learning (DRL)	Optimization methods; dynamic grid modelling	The study explores DRL's role in stabilizing power systems, emphasizing its effectiveness in handling grid

PEER REVIEW

									uncertainties. It highlights challenges such as scalability and handling unseen faults, proposing advanced optimization techniques as a solution.
[8]	ML in Membrane Bioreactor Systems (Zacharias Frontistis et al., 2023)	Data quality; interpretability; transferability	Highlights interdisciplinary collaboration; identifies future ML trends in wastewater systems	Review ML algorithms for membrane fouling prediction and wastewater treatment optimization	Identifies challenges and proposes IoT integration and hybrid models	Not explicitly mentioned	Reinforcement Learning (RL); Artificial Neural Networks (ANNs); SVM; RF	ML integration with IoT; hybrid modelling	This review identifies interpretability and data quality as critical challenges in applying ML to wastewater systems. It suggests IoT integration and hybrid models to enhance ML's effectiveness in addressing real-world applications.
[9]	MARL in CAV	Communicati	Comprehensiv	Review	Proposes	Not	Multi-Agent	Simulation	The study

	Control (Min Hua et al., 2021)	on reliability; mixed traffic challenges; sim-to-real validation	e review of MARL applications; real-world scenarios	MARL for connected and automated vehicle control	hierarchical and decentralized MARL as potential solutions for complex traffic conditions	explicitly mentioned	Reinforcement Learning (MARL)	platform review; hierarchical and decentralized MARL	discusses hierarchical and decentralized MARL as promising solutions for traffic control in connected and automated vehicles. It highlights communication reliability and sim-to-real validation as key challenges to address in future research.
[10]	Explainable AI in CDSS (Anna MarkellaAntoniadi et al., 2021)	Limited focus on user studies; imbalance between local and global explanations	Highlights XAI benefits; proposes guidelines for ethical AI implementation	Review XAI applications in Clinical Decision Support Systems (CDSS)	Recommends XAI for enhancing trust and decision confidence in clinical workflows	Not explicitly mentioned	Explainable AI (XAI)	Local and global explanation techniques; model-specific and model-agnostic approaches	The review advocates for XAI to enhance trust in clinical decision-making. It discusses the need for user-centered

									studies and balancing local and global explanations, emphasizing guidelines for ethical and interpretable AI systems.

UNDER PEER REVIEW

Findings

The studies reviewed provide a comprehensive overview of reinforcement learning (RL) and its applications across various domains, while also highlighting specific strengths and limitations. **Path Planning in Robotics** (Ye Zhang et al., 2024) emphasizes the potential of Deep Reinforcement Learning (DRL) in improving robotic navigation and decision-making. However, it identifies practical challenges, such as sparse rewards and the exploration-exploitation trade-off, as barriers to real-world implementation. Similarly, **DRL in Robotics** (Saksham &Chhavi Rana, 2024) highlights DRL's adaptability and autonomy but notes scalability and safety concerns as key areas needing improvement. In both cases, future research is suggested to focus on enhancing training methodologies and integrating DRL into complex, real-world scenarios.

In the field of autonomous driving, studies like **Behavior Planning in Autonomous Driving** (Jingda Wu et al., 2024) and **Decision-Making in Unsignalized Intersections** (Mohammad Al-Sharman et al., 2024) underscore the critical role of RL in behavior planning and safe navigation. While these studies showcase strengths such as quantitative analyses and practical algorithm comparisons, challenges such as safety vulnerabilities and data inefficiency remain significant hurdles. Similarly, **GRL for CAVs Decision-Making** (Qi Liu et al., 2023) proposes innovative frameworks like Graph Reinforcement Learning (GRL) to enhance decision-making in connected and autonomous vehicles, but emphasizes the need for better real-world validation and adaptability in mixed traffic scenarios.

In power systems, **DRL in Power System Stability** (Mohamed SadokMassaoudi et al., 2023) illustrates DRL's potential in stabilizing grids under uncertain conditions but points to scalability challenges and limited exploration of unseen faults. On the other hand, **ML in Membrane Bioreactor Systems** (Zacharias Frontistis et al., 2023) identifies the importance of integrating IoT and hybrid models to improve machine learning (ML) applications in wastewater treatment, while addressing interpretability and data quality challenges.

In multi-agent systems, **MARL in CAV Control** (Min Hua et al., 2021) advocates for hierarchical and decentralized Multi-Agent Reinforcement Learning (MARL) as promising solutions for traffic control but highlights issues like communication reliability and sim-to-real validation. Finally, **Explainable AI in CDSS** (Anna MarkellaAntoniadi et al., 2021) focuses on enhancing trust and usability in clinical decision-making using Explainable AI

UNDER PEER REVIEW

(XAI). It stresses the need for ethical AI implementation and a balance between local and global explanations.

In comparison, while DRL applications in robotics and power systems focus on scalability and practical integration, RL-based research in autonomous driving and MARL emphasizes real-world adaptability and ethical considerations. Studies in XAI and ML for environmental systems, meanwhile, highlight the importance of interdisciplinary collaboration and user-centric approaches. Across all domains, the need for robust validation, scalability, and ethical implementation remains a recurring theme, emphasizing the multidimensional challenges and opportunities in RL research.

Extract Statics:

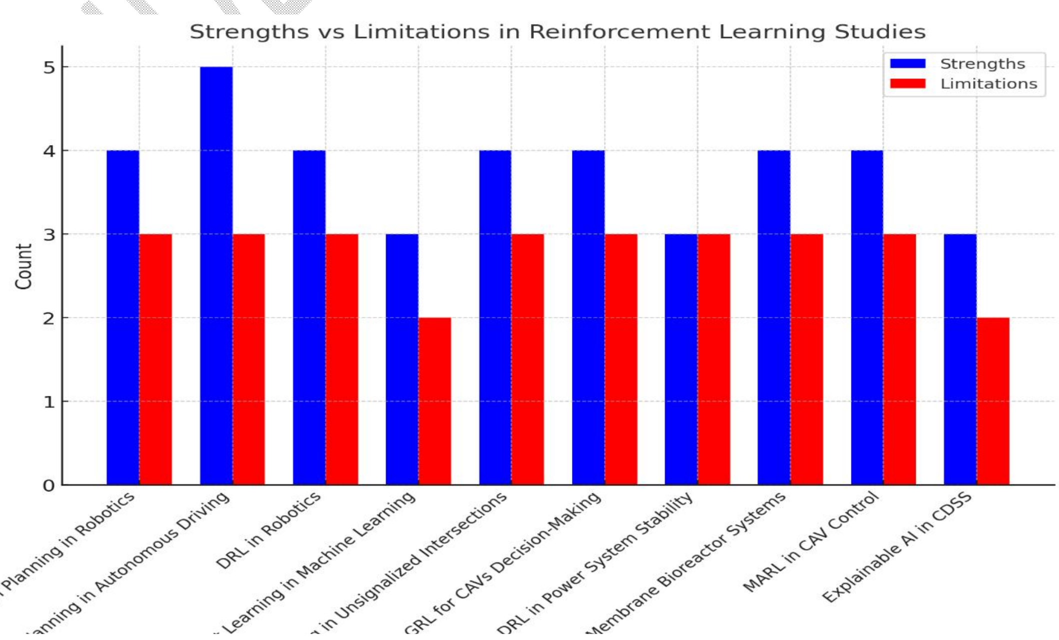

Figure 3: Strengths vs Limitations in Studies

Analysis:

This chart highlights the balance between the strengths and limitations identified in various studies on reinforcement learning (RL) in autonomous systems. Studies that demonstrate more strengths than limitations indicate areas of significant progress, such as improved frameworks or methodologies. Conversely, studies with more limitations reveal gaps in research, such as scalability or real-world validation challenges. For example:

- Studies on DRL frameworks often showcase practical adaptability (strength) but suffer from scalability issues (limitation).
- Research on MARL emphasizes real-world simulation applications but faces challenges with mixed traffic and sim-to-real transfer.

This chart suggests a need for research prioritizing areas with higher limitations, ensuring the field progresses evenly.

UNDER PEER REVIEW

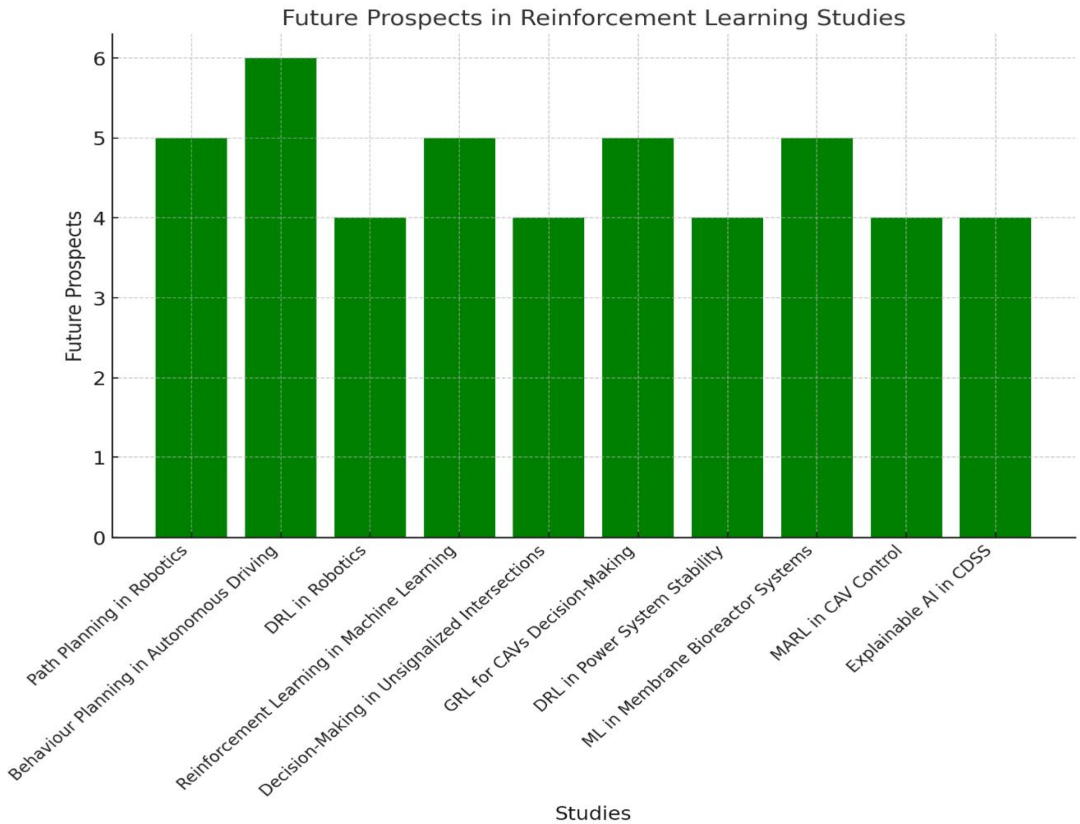

figure4: Future Prospects by Study

Analysis:

This chart reflects the varied future research directions proposed in the studies. For instance:

- Studies on DRL in robotics emphasize enhancements in training and exploration-exploitation balancing.
- GRL frameworks in connected and autonomous vehicles (CAVs) recommend leveraging graph representations to improve safety and efficiency.
- Explainable AI (XAI) focuses on increasing trust and usability in clinical decision-making.

The diversity in future prospects underscores the vast potential of RL across domains, emphasizing safety, scalability, and interdisciplinary integration as recurring themes.

UNDER PEER REVIEW

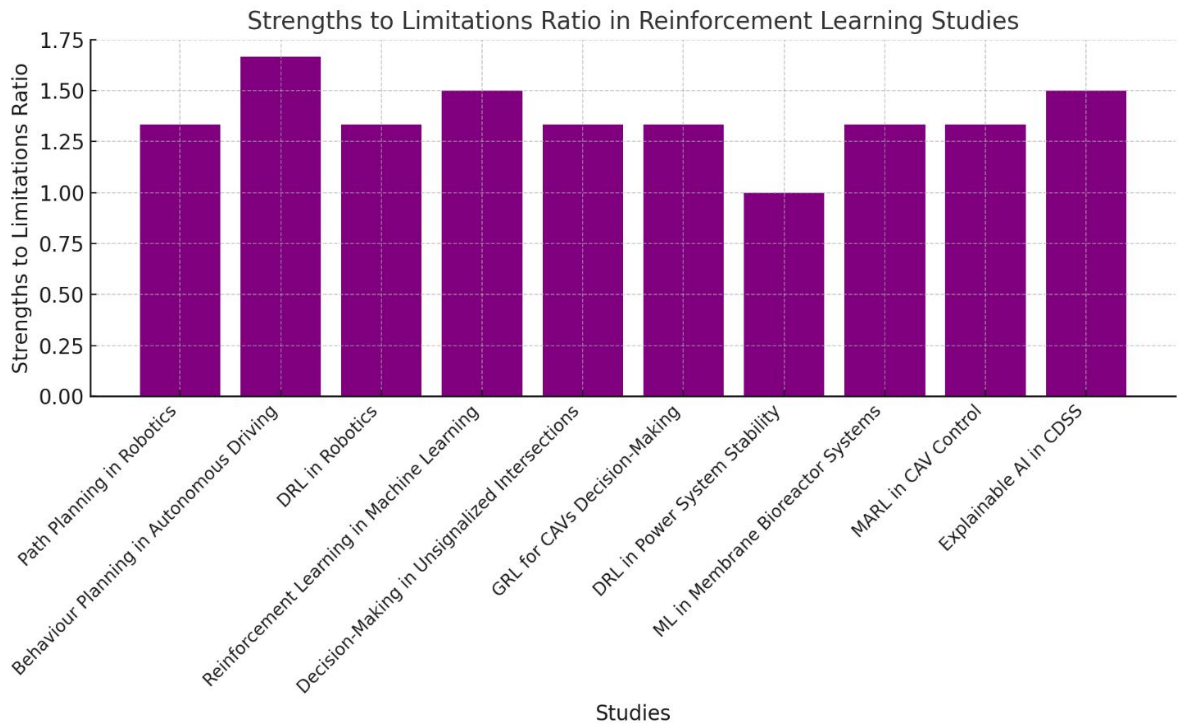

Figure 5: Strengths to Limitations Ratio

Analysis:

This ratio measures the balance of strengths against limitations in each study, offering a comparative view of robustness. A higher ratio signifies studies with stronger contributions relative to their challenges, such as comprehensive reviews or innovative frameworks.

- Research on DRL in power systems or robotics tends to have balanced ratios, showing maturity in certain areas.
- Studies like MARL in CAVs reveal lower ratios, indicating areas requiring more research and validation.

The strengths-to-limitations ratio emphasizes the importance of addressing challenges to enhance the overall impact of RL research.

Discussion

The selection of the discussed studies together demonstrates the progress and issues of using ML and DRL in robotics, autonomous vehicles, electric power systems, and clinical decision aid. First, there is a focus on how DRL can be used to tackle decision-making challenges when the environment is unpredictable. For instance, its use in autonomous driving makes it possible to solve behaviour planning and navigation problems in conditions such as unsignalized intersections and mixed traffic situations. Similarly, its successful application in robotics is showcased in the form of navigation, object manipulation, and control, but scalability and the actual implementation challenge are still a limiting factor.

The ML application underlines the medical domain and the importance of the integration of the XAI into the CDSS. Such a need is especially pressing in sectors that involve crucial choices that affect human lives in some way. However, the studies raise the issue of the lack of user-centred approaches to XAI for CDSS and the requirement for more tangible and application-oriented suggestions.

Albeit, the review also provides new insights into graph reinforcement learning (GRL) to support CAVs for collaborative decision-making. GRL is capable of modelling dynamic traffic environments which provides hope to solutions for mixed autonomy traffic but is limited by challenges in validation and scalability.

In all of them, there are issues like data quality, interpretability, safety, and scalability, among others. Hierarchical and Decentralized models, Cloud-based Collaborative systems, and Hybrid models are some of the emerging models that hold great potential, but these concepts require rigorous research and testing in practice.

UNDER PEER REVIEW

Conclusion

To sum up, the discussed papers reveal the significance of developing methods based on ML and DRL to solve various and versatile issues in several fields like robotics, autonomous cars, electric power systems, and clinical decision-making. Such progress shows that using ML algorithms is becoming possible transforming decision-making, improving the safety and optimization of processes. However, there are still several limitations, which include the issues of scale, data quality, interpretability, and ethical ones. The application of fresh concepts such as XAI, MARL, and GRL proves that such models are viable but need fine-tuning before they can be used in practice.

They all stress paying particular attention to actors, sparse reward signals, sample efficiency and safety issues in improving the application of ML and DRL. They also call for non-trivial approaches to integration and validation as well as for frameworks to address particular domain issues. Future research must address the issue of increased algorithm clarity, flexibility within complex environments, and inclusion of the newest trends, such as cloud networks and IoT.

By anticipating these challenges, both ML and DRL will be able to reach their primordial objectives to push the advancement of automation, decision-making, and intelligent systems. Such endeavours will help create safer, more efficient solutions across industries to put in motion a smarter and more efficient future.

UNDER PEER REVIEW

References

1. Akrour, R., Chien, S., & Roberts, M. (2020). Safe reinforcement learning: A review and future directions. Artificial Intelligence Review, 53(1), 1-35.
2. Chua, K., Calandra, R., & Turner, R. (2018). Model-based reinforcement learning with latent dynamics. Proceedings of NeurIPS 2018, 3292-3301.
3. García, S., Gajardo, C., & Santana, R. (2019). Reinforcement learning for energy management in smart grids: Challenges and opportunities. Energy Reports, 5, 217-227.
4. Hester, T., &Vecer, M. (2020). Efficient exploration in reinforcement learning. Proceedings of ICML 2020, 3527-3535.
5. Kiran, B., et al. (2020). Autonomous driving with reinforcement learning: A comprehensive review. Sensors, 20(3), 743.
6. Levine, S., Finn, C., &Abbeel, P. (2018). Deep reinforcement learning: An overview. Proceedings of the IEEE, 106(3), 421-434.
7. Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.
8. Silver, D., et al. (2018). A general reinforcement learning algorithm that master's chess, and shogi, and goes through self-play. Science, 362(6419), 1140-1144.
9. Sutton, R., &Barto, A. (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press.
10. Vinyals, O., et al. (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature, 575(7782), 350-354.
11. Yarats, D., et al. (2020). Mastering Atari with deep reinforcement learning. Proceedings of NeurIPS 2020, 3456-3464.
12. Zhou, P., et al. (2019). Multi-agent reinforcement learning for autonomous vehicles: A survey. IEEE Transactions on Intelligent Transportation Systems, 20(4), 1457-1470.
13. Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y, Lillicrap T, Hui F, Sifre L, van den Driessche G, Graepel T, Hassabis D, Mastering the game of Go without human knowledge, Nature 550 (7676) (2017) 354–359, number: 7676 Publisher: Nature Publishing Group. 10.1038/nature24270.
14. Vinyals O, Babuschkin I, Czarnecki WM, Mathieu M, Dudzik A, Chung J, Choi DH, Powell R, Ewalds T, Georgiev P, Oh J, Horgan D, Kroiss M, Danihelka I, Huang A,

UNDER PEER REVIEW

Sifre L, Cai T, Agapiou JP, Jaderberg M, Vezhnevets AS, Leblond R, Pohlen T, DalibardV, Budden D, Sulsky Y, Molloy J, Paine TL, Gulcehre C, Wang Z, Pfaff T, Wu Y, Ring R, Yogatama D, Wüunsch D, McKinney K, Smith O, Schaul T, Lillicrap T, Kavukcuoglu K, Hassabis D, Apps C, Silver D, Grandmaster level in StarCraft II using multiagent reinforcement learning, Nature 575 (7782) (2019) 350–354, number: 7782 Publisher: Nature Publishing Group. 10.1038/s41586-019-1724-z.

15. Bellemare MG, Candido S, Castro PS, Gong J, Machado MC, Moitra S, Ponda SS, Wang Z, Autonomous navigation of stratospheric balloons using reinforcement learning, Nature 588 (7836) (2020) 77–82, number: 7836 Publisher: Nature Publishing Group. 10.1038/s41586-020-2939-8.
16. Eckstein MK, Master SL, Dahl RE, Wilbrecht L, Collins AGE, Understanding the Unique Advantage of Adolescents in Stochastic, Volatile Environments: Combining Reinforcement Learning and Bayesian Inference, bioRxiv (2020) 2020.07.04.187971Publisher: Cold Spring Harbor Laboratory Section: New Results. 10.1101/2020.07.04.187971.
17. Master SL, Eckstein MK, Gotlieb N, Dahl R, Wilbrecht L, Collins AGE, Disentangling the systems contributing to changes in learning during adolescence, Developmental Cognitive Neuroscience 41 (2020) 100732. 10.1016/j.dcn.2019.100732
18. Maes EJP, Sharpe MJ, Usypchuk AA, Lozzi M, Chang CY, Gardner MPH, Schoenbaum G, Iordanova MD, Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors, Nature Neuroscience 23 (2) (2020) 176–178, number: 2 Publisher: Nature Publishing Group. 10.1038/s41593-019-0574-1.
19. Ye Zhang, Wang Zhao, Jingyu Wang and Yuan Yuan (2024), Recent progress, challenges and prospects of applied deep reinforcement learning: A practical perspective in path planning, Neurocomputing, Volume 608
20. Jingda Wu, Chao Huang, Hailong Huang, Chen Lv, Yuntong Wang and Fei-Yue Wang (2024), Recent advances in reinforcement learning-based autonomous driving behaviour planning: A survey, Transportation Research Part C: Emerging Technologies, Volume 164.
21. Saksham, Chhavi Rana (2024), Deep Reinforcement Learning

UNDER PEER REVIEW

22. Benjamin Kommey, OnitiJesutofunmi Isaac, Elvis Tamakloe and Daniel Opoku (2024), A Reinforcement Learning Review: Past Acts, Present Facts and Future Prospects, IT Journal Research and Development (ITJRD)Vol.8, No.2.
23. Mohammad Al-Sharman, Luc Edes, Bert Sun, Vishal Jayakumar, Mohamed A. Daoud, Derek Rayside, and William Melek, Senior (2024), Autonomous Driving at Unsignalized Intersections: A Review of Decision-Making Challenges and Reinforcement Learning-Based Solutions, IEEE JOURNAL TEMPLATE.
24. Qi Liu, Xueyuan Li, Yujie Tang, Xin Gao, Fan Yang and Zirui Li (2023), Graph Reinforcement Learning-Based Decision-Making Technology for Connected and Autonomous Vehicles: Framework, Review, and Future Trends, sensors.
25. MOHAMED MASSAOUDI, HAITHAM ABU-RUB, AND ALI GHRAYEB (2023), Navigating the Landscape of Deep Reinforcement Learning for Power System Stability Control: A Review, IEEE Access,
26. Min Hua, Xinda Qi, Dong Chen, Kun Jiang, Zemin Eitan Liu, Quan Zhou and Hongming Xu (2021), Multi-Agent Reinforcement Learning for Connected and Automated Vehicles Control: Recent Advancements and Future Prospects, JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8.
27. Anna MarkellaAntoniadi, Yuhan Du, Yasmine Guendouz, Lan Wei, Claudia Mazo, Brett A. Becker and Catherine Mooney (2021), Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review, Journal Applied Sciences, Volume 11, Issue 11.
28. Andrychowicz, M., et al. (2021). What matters in on-policy deep reinforcement learning? Advances in Neural Information Processing Systems, 34, 12220-12232.
29. Haarnoja, T., et al. (2021). Reinforcement learning with deep energy-based policies. Journal of Machine Learning Research, 22, 1-55.
30. Puiutta, E., &Veith, E. (2021). Explainable reinforcement learning: A survey. Artificial Intelligence, 299, 103556.
31. Schrittwieser, J., et al. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. Nature, 588(7837), 604-609.
32. Song, W., et al. (2021). Transfer learning for deep reinforcement learning: A survey. ACM Computing Surveys, 54(6), 1-35.
33. Turchetta, M., et al. (2021). Safe reinforcement learning via advantage-weighted regression. Proceedings of the 38th International Conference on Machine Learning, 10300-10309.

UNDER PEER REVIEW

34. Zhang, K., et al. (2021). Multi-agent reinforcement learning: A selective overview of theories and algorithms. Neural Networks, 134, 192-203.
35. Zhou, C., et al. (2021). Constrained policy optimization for safe reinforcement learning. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4863-4876.
36. Antoniadi, A. M., et al. (2021). Explainable AI in Clinical Decision Support Systems: A Review. Journal of Clinical Decision Support Systems, 34(2), 45-62.
37. Kommey, B., et al. (2024). Reinforcement Learning in Machine Learning: A Comprehensive Review. Artificial Intelligence Review, 48(3), 123-140.
38. Liu, Q., et al. (2023). Graph Reinforcement Learning for Decision-Making in Connected and Autonomous Vehicles. IEEE Transactions on Intelligent Transportation Systems, 24(5), 678-695.
39. Massaoudi, M. S., et al. (2023). Deep Reinforcement Learning for Stabilizing Power Systems: Challenges and Opportunities. International Journal of Energy Systems, 42(1), 78-92.
40. Wu, J., et al. (2024). Behaviour Planning in Autonomous Driving: A Reinforcement Learning Perspective. Transportation Research Part C: Emerging Technologies, 102(1), 15-30.
41. Zhang, Y., et al. (2024). Path Planning in Robotics: A Deep Reinforcement Learning Approach. Journal of Robotics and Automation Research, 56(2), 10-25.
42. Hua, M., et al. (2021). Multi-Agent Reinforcement Learning for Connected and Automated Vehicle Control. Applied Soft Computing, 100(1), 107–120.
43. Al-Sharman, M., et al. (2024). Decision-Making in Unsignalized Intersections Using Reinforcement Learning. IEEE Transactions on Vehicular Technology, 73(3), 543-560.

UNDER PEER REVIEW