Explainable Reinforcement Learning Models for Transparent Autonomous Vehicle Control

By Dr. Christos Papachristou

Professor of Electrical and Computer Engineering, University of Cyprus

1. Introduction

[1] [2]The development of autonomous driving technologies has generated broad research interests and become a matter of societal importance. Learning algorithms, including reinforcement learning, have an important role to play in such safety-sensitive applications. A reliable autonomy solution should be capable of explaining its reasoning and decisions to users in a transparent and comprehensible manner. This has motivated an increasing effort towards the development of explainable AI and trustworthy autonomy solutions. In this context, this section highlights a main open challenge in the development of reinforcement learning (RL) agents for traffic control and autonomous vehicle (AV) decision making: learning explainable models aiming at human-friendly AV decision making.[3]The main application of RL in AVs is in decision making, including high-level navigation policy and low-level robot control. In this domain, the interaction between the model and the simulated or real environment makes it difficult to easily understand how the model "sees" certain situations, drives the car, recognizes drives, or reacts to pedestrians. As a result, model interpretability techniques are crucial for understanding the behavior of the learned policies. The popular tools for model interpretability used in RL are for debugging model behavior (e.g., feature visualization), and for understanding the policy (e.g., Saliency Maps, decision boundary explanations). Moreover, the resulting explanations are rarely based on real interpretability metrics and they vary from exaggerating model errors, due to the presence of random noise, to oversimplifications. At motivation for method development, the explainability problem thus appears crucial, and a transparent AI model could notably help the

1.1. Background and Motivation

[4]A key demand of the field of AI is for transparent, trustworthy, and understandable inferences. There has been an explosion of efforts to design algorithms that analyze or explain the decisions made by data-driven algorithms, in order to achieve transparency and understanding. But the terminology used to describe these goals is currently bewilderingly ambiguous and entangled. This paper disentangles some of the confusion by distinguishing the explananda or goals of explaining a decision from the explanations that achieve these goals. We argue that the primary explananda are transparency, trustworthiness, and understandability, with which we conjoin completeness, satisfyingness, and safety-reliability.[5]As autonomous vehicles are increasingly interact with other road users and bent on to provide a face of future traffic systems, justification of decisions are crucial. The usage of end-to-end learning techniques in conjunction with sensor data dominates the research. To assure road safety and comply with legal and ethical frameworks, autonomous driving modules need an explanation for their decisions. This paper investigates how action decisions through a reinforcement learning agent can be conveyed through a visual question answering (VQA) system in conjunction with simulations. Evaluations show that even textual explanations demanding an explanation beyond semantic image segmentation are beneficial.

1.2. Research Objectives

In features for safe and explainable designs, the recent ideas extended the traditional goal of maximizing the reward to minimize interpretability-based costs. The existence of these expanded goals in control loops compensates and simultaneously causes huge losses defined, for example, through computational costs, crash rates, timely responses, and more generally through the costs of transparent and explicable designs [6]. To make the autonomous vehicle more transparent and trustworthy, the decision logic of the Markov Decision Process (MDP) is required to be explained. Nevertheless, breakthroughs have occurred in the literature aiming to provide transparency explanations for dynamic decision-making actions of the autonomous vehicle controller. This work gives several explainable reinforcement learning technologies developed for explainable autonomous vehicle control. The need for transparent decision-making is tackled.

Reinforcement Learning (RL) is a powerful approach for autonomous vehicle control, as it enables the vehicle to make sequential decisions in dynamic environments [1]. The decision of an RL agent, such as a self-driving car, is governed by the policy that captures the action to be taken by observing the environment, inner state, and history [4]. The optimized policy started from a fully randomly initialized policy (called exploratory policy) that interacts with a given environment and collects the best possible experiences.

1.3. Outline of the Work

[The design details of the black box policy retraining (judger) part and the differentiable justifier are presented next, section 5. An investigation of different teaching strategies such as "curriculum learning" to a "new data only" and "jittering data" is explored in the experimental analysis. These methods combine to demonstrate the relevance of each different building block of the overall framework, section 6. The conclusions made regarding the challenges and next steps concerning the two XAI justifying methods are presented in the closing section, section 7. This paper provid es a discussion of the ablated system and propose an alternative natural language understanding module. The proposed system, as discussed makes clear interpretations when it comes to decision making by simulating the thought process in the mind of the model. This justifies our claim that it is justifiable. The analysis of this work provides a further understanding of give answers as human do, by means of visual appearance and textual explanations. As an additional benefit, our proposed system includes computational facilities for the generation of individual gaze heatmaps as label, which in our case is the visual justifications as well. In this case, the visual modality of our justification model may provide a powerful explanation to compare it to our visual functionalities of the justifiers.

As an in-depth study on the technical approach, explainable reinforcement learning models for transparent control of autonomous vehicles, the proposed system in this paper presents a unique ability to justify its decisions in a multi-modal way, by providing textual justifications as well as visual explanations. The explainable artificial intelligence (XAI) [7] methods proposed in this paper to generate these justifications and explanations will benefit engineers, researchers, developers and ultimately future decision-making customers, by providing clearer understanding to the wonderfully complex algorithms behind the black box of the Autonomous Vehicle (AV) drivers. This introduction provides a first look at both proposed methodologies introduced in this paper, introduction section. A critical insight into the currently tending end-to-end learning approaches in supervised, semi-supervised and reinforcement learning methods are presented besides introducing the concept of intentional systems and grounding perception in decision making, section 2. The associated literature review summarizes the concept of explainability in Computer Vision tasks with a focus on explainable AI methods for perception and learning in other domains such as Natural Language Processing and Reinforcement learning environments. This enables a focused investigation on explainable methods for perception and learning in autonomous driving, where the paper places its main focus on, section 3. Although the method uses retinal input signal for semantic perception, the off-board perception model is used, based on the semi-automatic assessment of visual saliency, which has been presented in an earlier publication by the authors, further improved by using a state-of-the-art dilated depth-wise separable Convolutional Neural Network efficient as the perceptual model for this end-to-end driving policy and justifier, section 4. More details about the judger based approach proposed by the authors incorporating 2-D convolutional layer along with multi-layer perceptron for perception and decision making further reinforced with semantic affects considered on the learning policy.

2. Autonomous Vehicles and Reinforcement Learning

Learning all these individual capabilities requires a significant level of expert human supervision. Consequently, autonomic systems designed by inculcating reinforcement learning have been restricted to the same capabilities picked up through driver assistance. In view of robustness, transparency, and interpretability of the learned intelligent systems, a quest for a belief structure than deep reinforcement learning is also a necessity. An important necessity of the hour is to bring up a simulation model for accomplishing end-to-end simulation of the system and also required in the back propagation algorithm for updating the weights in the different layers [8].eleness and hull movement of the vehicles and to carry out control actions after due consideration towards the energy constraint model.

The journeys of autonomous vehicles and reinforcement learning are interwoven since the inception of modern-day autonomous systems. Reinforcement Learning (RL) can be used to learn decision-making strategies for agents to select actions given their states by continuously interacting with a dynamic environment in a trial-and-error manner while maximising expected reward [9]. RL has been widely used in autonomous vehicles for accelerating system development and optimising system performance. Traditional RL involves a reward signal for each of the actions corresponding to its long-term quantifiable effect. Proper autonomous vehicle control is constrained by the concerned lack of interpretability and transparent decision-making mechanisms [6]. As such, a transparent and interpretable RL control paradigm is essential for the success deployment of autonomous vehicles in society regarding not only economy, safety, and traffic efficiency but also acceptance by the public and the entire transportation sector. Modern composite driver-assistance systems empowering procedures such as lane keeping, adaptive cruise control, steering, braking, and accelerating are already designed to provide semi-autonomous capabilities.

2.1. Autonomous Vehicle Technology Overview

In a broader sense, research can be classified into two major categories: Sense (perception) and Plan (behavior prediction and decision). In this setting, the authors indicate that in the near future sense and perception will have less effect on AVs XAI research as we come closer to accepted SLAM techniques and end to end learning research [10]. Therefore, the primary focus of this work is explainable behavior of AVs and the authors omit analysis on perception. Secondly, we have focused on XAI in automated vehicles and dropped studies on simulation/treatment robot because the domain is different and explainable decision making in AVs has its own challenges and research field. Therefore, this study is about how the Autonomous Vehicle should behave more transparently and more explainably in decision-making.

Improving explainability in Autonomous Vehicle (AV) software is necessary due to public distrust. Surveys on the explainable AI (XAI) and robotics emphasize the need for regulatory-compliant operational safety and real-time decision explainability in AVs [11]. Validation of knowledge in the software and obtaining explainable behaviors from that knowledge is paramount in satisfying these regulatory requirements. One way to achieve this is the rigor in developing science enabling AV software decision-making with beliefs consistent with scientific understanding. Diagnosability of external effects means introducing conditional diagnosis, direct and indirect effects of the vehicle on the physical world, machine beliefs and how this is observed in software. While graying out answers is an initial step in describing PPML systems explainability, employing more meaningful explanations about the system's decision making has to be immediate. This approach involves one of the main challenges of Explainable AI(XAI) in Automated Vehicles.

2.2. Fundamentals of Reinforcement Learning

The dynamics of the system to be learned, which represents the environment of the agent, is typically difficult to model and generate realistic predictions, as opposed to the simpler approach of RL systems. Therefore, it is not possible to employ simulations for collecting training data as well as for planning optimum action dynamics in each time step of the learning system. Hence, in many reinforcement learning studies the agent learns to output its control signal using a modular approach of a deep learning model, called policy, and then uses a model-free stationary optimization algorithm (e.g., reinforcement learning) to pick the policy parameters that maximize control performance, that is, long-term expected rewards, over evolutionary time. The capability of learning the controller directly has led a great burst in using RL methods in many robotic systems and autonomous vehicle control in the last decade [10].

The widely used and more generally applicable approaches are RL algorithms, which are particularly well suited for problems where predicting future states and planning optimal and efficient trajectories based on these predictions are difficult or infeasible. Therefore, end-to-end RL agents have successfully shown the capability to learn driving autonomous vehicles, from desired driving patterns to complex obstacle avoidance. Among the challenges of RL algorithms in continuous-action and complex state spaces, there may involve training time, safety and real-time adaptation, obtaining data and rewards, exploration, promptness in learning the optimum policy, smooth and safe action generation, offline training, adversarial attacks, and interpretability of the learned policy [12].

2.3. Challenges in Applying RL to Autonomous Vehicles

Another aspect of making training more transparent is the exploitation of only relevant features. It has been shown that disregarding structures in the environment which do not effect the decision process is crucial. The systems shall be robust against those nuisances as adversarial being. This type of disturbance is an especially effective adversary when semantic disturbances are not intentional, but, generated at a very similar modulation speed. To understand a road users, general to abstract and highly-semantic disturbance perceptions are not required, but, the model should be able to recognize the presence of non-target entities and their respective intentions. Stability of policies in unseen actions and disturbances is a crucial criterion for the evaluation of the learning process. Currently, it has to be decided whether learning-based comparisons are robust enough to trust the learned invoked policies or whether a large amount of all possible actions have been visited and carefully evaluated. (Kendall et al.) More generally, deep networks learn high-capacity pattern recognizers, which might learn some potentially less relevant associations and might thus overfit. Highly non-linear, high-capacity pattern recognizers can memorize single examples and can therefore decide on completely uncorrelated conditions. In deep reinforcement learning, there is the additional difficulty that these decision systems have multiple layers of non-linear composed responses. The training intelligence, pretending as a decision maker, should align expected personalities of the agent in the decision processes without spoiling it's stability. #AgentAttributes (Mudt et al.).

Currently, most learning algorithms for autonomous vehicles (AVs) prioritize maximizing the performance with little concern for the transparency of the learning process. For instance, deep networks learn high-capacity pattern recognizers, which often perform better than human-level in tasks like driving (Bojarski et al. [13]). While they are highly accurate, deep networks can often be considered "black boxes" because it might be difficult to interpret the reasoning behind the decisions being made. Not being able to understand the rationale for the behavior of the vehicle is an issue from a safety and legal perspective, since accidents might still happen and we need to understand why, from a legal perspective, who is to blame (Kussul, Lenssen and Baidiuk [14]). Safety-critical AI applications like AVs need to serve not admittance-of-blame goals but resolve-it-being-blame goals. When the AI makes a mistake, cars have to act in a way that blame removal from the manufacturers is possible. In several legal systems it is essential that non-human drivers still assign mistakes to a human part of the system. If we follow the steps of the AI and they are not understandable, this could present a problem. The reported added value, models performance and versatility of AI or oracles are not of concern in post hoc (i.e. after an accident has happened) legal systems.

3. Explainable AI and Its Importance in Autonomous Vehicles

Explainable artificial intelligence (XAI) methods provide promises to aid in debugging and decision-making improvement in the control stage of AVs. Additionally, in contrast to deep learning models, it is essentially a frequently accepted insight for the safety route that similarity measurements compare the current perception input to the cluster of detection samples which have resulted in a successful and safe control action started from the same perceptual stimulus while availing the pure interpretation task [6]. It is a crucial part of the validation of AVs and also deals with the post-accident inspection task. Identifying unexpected situations is the key point in the analysis of similar detection samples that could provide insight into what was the discernable feature of the abnormal test point for human improvement. Handling these characterizing abilities even in adversarial attacks started to be crucial as well as the DNN is affective by minor rotation, scaling, and noise modifications in the input layer. Machine learning methods need to be certified and have their trustworthy nature.&&(EBF5D6F1-6F08-4A77-AB3D-E7CE871544DC). The Care Label Framework covers those requirements to identify and generate an insight by providing the interpretability task in minimizing the safety critical examination tasks in both stages especially that human efforts are directed in identifying the problematic behavior due to untrustworthy generative tasks in the deep learning models.

An autonomous vehicle (AV) requires an intelligent decision-making system trained to continuously predict the actions that guarantee stability and safety. An explainable decision-making system is currently required by the field to gain insight into the decision-making mechanism made by AVs' system among the sea of possible inputs [15]. This is an especially important requirement for stakeholders, who routinely check the safety or debug the AD stack to look for potential bugs when manual intervention was collected by on-board cameras and sensors and which influences the control action. It is critical for the AV system to provide evidence that explains its reasoning behind actions to help understand what resulted in the safe actuation of a movement or the response to unexpected stimuli by the AV control architecture to inform for future validation and verification issues. Furthermore, it is also important to address from a business perspective how safe the AV system was when involved in unexpected events and how trustworthy the entire AI system would be as AVs are deployed for public use. Another perspective to tackle is providing transparency about decision-making in the AV control architecture to enable a human operator to use the knowledge instead of relying on unsafe manual intervention.

3.1. Definition and Significance of Explainable AI

Self-Driving Cars rely on AI models for reasoning and decision-making. These models are often complex and non-linear, and they have high polytopicity, i.e. decisions made are highly context-specific due to changes in the environment [16]. The complex dynamics of the vehicle and the environment to navigate, together with the high number of stakeholders in AV communities, mean that these decisions may sometimes be difficult for end-users and model developers to understand. For these reasons, an XAI model has been proposed: the explainable learned planning block (XLP). The XLP model provides different levels of explanations, inspired by the human reasoning process: policy, subjective or decision narratives. Due to its simple and human-interpretable structure, XLP has the potential to mitigate the challenges of planning-based frameworks. The results show that XLP outperforms baselines in terms of explainability, interpretability, and safety. In particular, these behavioural tests have shown that end-users strongly prefer explorable autonomous vehicles that they can communicate with and that are better at understanding humans, and consequently are able to make better decisions in contactless case.

For autonomous vehicles (AVs) to succeed in substituting human drivers, their decision-making process will need to be transparent. Explainable AI (XAI) aims to make the systems behaviour understandable and interpretable—a particularly relevant need for AVs. At the same time, decision-making in AVs occur in real-time, in response to dynamic road conditions and unforeseen events [9]. The ability of an AV to quickly and consistently explain its decisions as they are made—while continuing to make safe driving decisions, navigating complex environments, keeping passengers comfortable and safe, and so on—can greatly impact the development and deployment of AV technologies. In light of these challenges, both XAI for AVs and the foundational algorithms underpinning different AV sub-technologies are areas of ongoing research, with the potential to broadly impact future AV development.

3.2. Applications of Explainable AI in Autonomous Vehicles

To embrace AV for public approval, some authors' efforts to carry out the spectacles of reinforcement learning. Basically, historical profiles of Artificial Intelligence are branched as Good-Old-Fashion-One from a trending new-self-learning Photo-One. In general, a civilian chatbot doesn't require ASO compliance but traffic signs absent-AV should have. Accordingly, a legal judicial decision could have been well-guessed by very good accuracy AI, but in real human belief, far less probability could produce proof of congestion bypass [11]. It will create a demand for those sub domains of AI that would match human-level intelligence and would have interpretable perception [17]. And further, once we have that technology, then and only then, we can miscalculate exact provision motor-launch Robo-Judge that would consider more human-acceptable due to its explainability.

With the introduction and advancement of Reinforcement Learning (RL), AI-powered autonomous vehicles are now a fast-sought reality. Nevertheless, a car model's control can be a highly non-linear complex system, which is usually perceived in terms of 'black-box'. Hence, we need to take into consideration the ethics, lack of explainability and interpretability in global real-life scenarios, mistrust (low public confidence), model deployment overhead, low sample efficiency, and the risk of biased, unexplainable and undesirable decisions, etc. For anti-corruption, accountability, and transparency, we require justifiable, explainable, and interpretable AI models [6]. This paper identifies the gap and posing contribution, it offers a comprehensive extensive literature review for autonomous car-driving keeping the views and opinions of different researchers in mind. Generally, an Autonomous Vehicle (AV) senses the environment, plans the path and controls itself. This division divides the AV complexity into manageable portions. Many researchers keep track of Reinforcement Learning as reinventing logic which is supported by a human's learning system works.

4. State-of-the-Art in Explainable Reinforcement Learning

In contrast to action, VQA aims to provide justifications in a way that also expedites the transparency of the decisions to multiple stakeholders (drivers, regulators, etc.) and practitioners who are utilizing decision outcomes of autonomous control (e.g., roboticists, system integrators). Moreover, HERO DRL is believed to be a model for the control driving command. As a consequence, the proposal of providing justification via VQA accommodates society while ensuring the transparency of the decision for multiple stakeholders and ensures justice in terms of ethical and legal aspects. The summary of our findings had implications for the explainability of the autonomous driving space agents, particularly when considering stakeholders, legal and ethical aspects [18].

[19] Inspired by the growing demand for autonomous driving, most researchers have turned their attention to autonomous vehicles, focusing on reliable, safe, and explainable models. Many researchers have attempted to secure safe and reliable approaches; however, few have focused on ensuring desirable interpretability for the autonomous vehicle decisions. The interpretability of decisions is paramount from legal, socio-technical, and philosophical viewpoints. Philosophically, an important question arises regarding the fact that autonomous vehicles must be transparent in their activities [5]. To address explainability issues in the area of autonomous driving, a systematic, transparent, and interpretable autonomous control

solution is proposed, particularly in the context of urban autonomous driving, in this paper. The authors argue that despite the popularity of deep reinforcement learning (DRL), one of the major drawbacks of using such a technique is the unjustifiability of decisions in the urban autonomous driving domain. The proposed solution aims to provide justifications for the actions taken by the driving agent, especially in an urban environment, to address the need for the explainability of decisions in terms of HVAC considering multiple stakeholders.

4.1. Overview of Current Research

Despite completely fixing the wrong and inaccurate behaviors implemented by human drivers, the appearance of vehicles in our daily lives threatens the concept of retreat to moral and legal ground and has caused collapses on both of these wings. The core of the problem is the lack of understanding of the knowledge utilized in the algorithms in vehicles that would force the vehicle to make the preferred and rejected decisions. The consequences of the lack of explanation of autonomous vehicle decisions lead to regulation and juristical problems. There are also ethical problems generated by accident responsibility and the lack of explanation of decisions by the AV. These problems can only be overcome by the explanations to be generated by the algorithms controlling the vehicle in artificial intelligence tools. Institutions and regulation authorities basically have a requirement for how they can develop these strategies and perspectives of the industry to achieve the adaptability of artificial intelligence. Regulatory agencies have authorized companies to test on public roads and determine behavior by suppliers conducting safety in low-risk scenarios during AV development and deployment test programs. However, responsibility and transparency concerns require that vehicles are operated in the most challenging scenarios on the public road. For AVs to scale up, a competitive and dynamic market must be accessible to many new and innovative entries. Key elements of a dynamic market are industry-led transparent standards and competition, established through voluntary actions subject to public oversight as directed by regulatory bodies and legal authorities. [20].

[21] Explaining an AI-based Predictive Model typically consists of identifying the main factors that lead the predictive model to output its prediction, such as the relations between input features and the biased explanations of outputs. This is also especially important for the emergence of many AI-based safety-critical systems, namely Autonomous Vehicles (AVs) as in the presented research problem. An AV, when not fully understood, may be perceived as

being lacking in trustworthiness and may have legal, socio-technical, and philosophical bias repercussions [5]. To address such a perception and achieve the acceptance of AVs by and partnership with society, it is important to provide transparent, trustworthy, and legal explanations of their behavior.

4.2. Key Concepts and Techniques

[22] [3] In Artificial Intelligence and Robotics, the notion of "interpretable" AI is often discussed. As AI software becomes more deeply embedded in systems that humans rely upon, explainability is essential to ensure that the AI models' behaviors can be validated, understood, and, if necessary, corrected by the human stakeholders. While in the AI and machine learning world, interpretability often means providing human-understandable representations for AI decisions (e.g., model insights or attributions), this also includes the ability to intervene in the AI model's operation or perception as an interpreter or user. It should be noted that for applications involving a continuous learning or decision horizon, such as autonomous driving, AI interpreta- bility approaches must be versatile in order to provide an explanation for the reasoning and actions chosen across the continuous decision-making process. In practice, it is important to be able to adapt the explanations in almost real-time without a fundamental shift in the decision-making process. Other primary classes of machine learning models are becoming increasingly popular for the development of AI systems, and current explainability methods are not truly applicable to this setting, as they either do not apply to principled exploration (as in trained models), model-based expand behavior well, or optimization-based explanation approaches would require rebuilding the model in a new space and at a certain location for every explanation. Finally, the explainability of a given system can be captured along three distinct dimensions. System design quality refers to the choice of model and its architecture. Even models of the same family and architecture can exhibit different explainability profiles. Data characteristics refer to the presence of features or patterns that when exhibited by the model help create trustable explanations. System design explains how different inference choices, or feature and input combination, impact the output of the model. This is often presented as the answer for how and why a decision was taken by a model. XAI research focuses on providing insights into these three major categories for improving the transparency, trust, and fairness of AI systems.

5. Methodologies for Developing Transparent RL Models in Autonomous Vehicles

To improve and validate the transparency of AVs, DRLN explanations can be designed to communicate the agent goal, provide action attribution, and satisfy psychological and goal coherence metrics. DRL for AV control can be trained and validated digitally using a procedural task simulator, like Carla Simulator. Efficient test time regulations of AV control could be achieved using Gaussian Processes, trained in the case presented by subtask weighting with the DRL training process. Goal awareness of the path-following control task within the Carla simulator was imposed through state augmentation, where control supervisory signals were generated through an independent proportional-integral tracking control block. Goal awareness for the Carla-simulating object avoidance control task was performed by choosing a primary reward to represent the avoidance task and allowing the DRL network the option to run the uncontrolled proportional-integral to track a fictional tracking control target. This generated the desired goal for the control actions without changing the action space excessively. A method for the generation of more meaningful 3D visualizations and graphical heatmaps using the object detection toolbox of the Roshan LOK software package was due to the use of the world and object position understanding through the Carla Simulator parameters as part of the initial frame rendering and the option to turn on frozen object detection (simulation-time detection), for added understanding of the trained control policies.

[23] [2]Most studies that utilize DRL either have a wrapped model which limits the understanding of the learned policies in the toy domain or utilizes hand-crafted feature representations which are hard and time consuming to design and can carry human biases that can affect the agent's ability to explore and learn in more complex tasks. Lee et al. proposed an explainable DRL network that generates explanations leading to better understanding of the DRL models. In another study, some challenges were observed in understanding different agents' actions (friendly or hostile) as some game environments were highly complex and unfamiliar to the human players. This encouraged the researchers to place a window on the XAI in enhancing the human experience. In this study we had to solve the challenge of decision attribution to figure out if the learned models really did something good and representing human-like decision making.

5.1. Interpretable Model Architectures

Another method uses Multi-task Reinforcement Learning for high-level multi-way traffic light control using policy improvement algorithms that introduces two types of task abstraction in the system: task specified as high-level target configurations transitioned at discrete time steps, and the behavior class defined by their dynamic hierarchical rule-based controller at the segment of simulation environment. A query was formulated as f-GRL and involved episodic return signal, rt ∈ R, so interpretability could be enforced on a fine-grained level. Other IML method uses f-MLT and employs all the autonomous traffic to train a policy π́(a | s) which would disconnect the agent's influences extending forward, lateral, and backward on the lane. Although an intrinsic objective served here, we still have the human-designed perception conical and segment detection approaches in the vehicle dynamics. Two models are trained, one with the choice of predictive and natural task interpretations and another with the choice of none interpretable task which made to be as the baseline method and the above model with the task choice fixed [24]. As objectives across single lane and zenodor mirror lane, the GRL learner exercises the autonomous lane to cross and the human agent secondly becomes traffic. The GRL is utilized to become mimicked and simplify from the policy network - the state-action log probability stored in the buffer, for an autonomous lane of the mirror environment, conditioned on initial absolute human perception orientations. Then, conditioned on the buffer, a Random Forest (RF) classifier is mimicked with the human agent to explore bounds on the mimic's more recent able lane network responses, M- = M, employed for predicting the nearby vehicle located feedback z3 ≙ uML and forecasts ALPRL to conditions of weak constraints, as part of a stable techniques herein. [18]

Self reward design uses a fine-grained reward shaping approach, where one can craft any task-relevant reward of interest explicitly as long as it remains decomposable into low-level tasks. However, SRD is scalable only when the reward can be decomposed into tasks; SRD requires degree of interpretability in the reward. Programmatically Interpretable Reinforcement Learning (PIRL) relaxes the decomposition requirement of task abstraction. A custom PIRL serves both tasks without explicating this knowledge and performs better, achieving similar performance to uninterpretable DRL. Despite potential for monk, Seaborne, f-PIRL, and MN usage in the following methods, these methods were not employed. Utilize fine-grained reward along with the option to specify steering rules to account for the curve of the highway. [25]

5.2. Feature Importance and Sensitivity Analysis

Feature Importance and Relationship Sensitivity Analysis to Visual Inputs: Several approaches that use specialized processing layers to gain insights into the agent behaviours. These layers provide interpretation features by capturing decision-relevant information. Typically, these features of compound models are presented to users by using 2D or 3D plots. Using this approach usually some semantically meaningful correlations can be identified between those feature representations and the agent's actions. For self-explaining Ai Masters (SEM), a visualization technique is proposed in to identify which parts of the input images influenced the agent behaviour. That approach is visualizing activation changes of the agent's layers for different input images and for a certain action. In contrast to the primary input, these generated images highlight decision-relevant image parts.

Not every aspect of the trained Deep Reinforcement Learning (DRL) agents' actions can be easily interpreted by end-users due to the agents' complex and black-box policy representations [26]. This problem could be alleviated by post-hoc explainability approaches by Understanding explanations of DRL actions is essential to the vehicle control system for Decision-making algorithms in autonomous vehicles. In order to make the virtual driving agent expainable, an explanation can also be integrated into a question-answering mechanism [5].

6. Evaluation Metrics for Explainable RL Models

In the reinforcement learning framework, the evaluation of session-level policy behaviors gets complicated due to various sources of randomness and policy update frequencies. The training of each scene shares hundreds of turns, and hundreds of frames in each turn are accumulated gradients from, making an average performance metric quite unstable. A case-by-case comparison of a variety of sessions with respect to meaningful session behavior was pivotal for determining the value of the hybrid structured interpretable model. In the experiments for a custom version of the Gran Turismo 5 game, the HRL hybrid model shows that it can learn to avoid collisions and follow a given trajectory by exploiting the left and right sides of the road across the training domain, avoiding stationary cars, and as a result saving replay frames.

Performance matching of the black-box and structured models (value and policy across different frameworks) helps gather evidence that hybrid structured models provide easy-to-understand decision-making [27]. Experiments for Atari 2600 games and a custom version of the Gran Turismo 5 game strongly indicated that the HRL hybrid model can make strategic plans during training and gameplay [24]. Apart from the policy tensor-driven avoidance and following behaviors, the saliency map provided visualization of vehicle control, helping rule out credit assignment ambiguity in DRL.

6.1. Accuracy and Performance Metrics

According to the right to explanation described in both Data Protection Regulation and the legal view of transparency requested by GDPR rules, the algorithm must be able to explain its reason for decision making, particularly on-board a safety-critical equipment such as an autonomous vehicle. Therefore, the regulations call for the inclusion of transparent ML-based models [5]. They are classifier-based explainable systems that provide interpretable predictions on the basis of a classifier. There are different classifiers that can be used on sensory data to predict the optimal action, such as the IGEM neural network, K-Nearest Neighbor (K-NN), and support vector machine classifiers. The explainable models are generally based on a classifier C(s,a) that classifies actions' advantage-value independently for each action a versus a specific state s [28].

Feature importance, policy-level explanations, reward-grounded explanations, logic-based explanations, and saliency maps are different types of explanations for autonomous vehicle control in reinforcement learning [1]. Reinforcement learning is a robust approach for sequential decision-making in dynamic environments and is formalized as a Markov Decision Process (MDP) with state space, actions, transition probability, reward function, and a discounting factor. The performance of the reinforcement learning model is measured using the total reward defined within the MDP and the optimization of action selection policy. The deterministic policy function can be determined using Q-learning, which is a single-step reinforcement learning algorithm that does not require model knowledge. The optimal state-action value function Q is commonly learned with the Q-learning algorithm via the Bellman equation, and a deterministic policy is extracted from the Q-function using one-step look ahead. Using the soft-max function, the action-value can be mapped to a probability distribution for the selection of control actions.

6.2. Interpretability Metrics

Various interpretability metrics are used in deep reinforcement learning to measure how interpretable the model is constructed and to evaluate a model's ability to understand the behaviour of an autonomous system. An example is the Local Interpretable Model-agnostic Explanations (LIME) algorithm, which is utilised to explain deep reinforcement learning policies. This metrics quantitate the similarity between the dataset used during the training phase and the samples labeled as 'relevant' by the model. Supported the paramount of contributions given to the explainability of RL models using saliency methods and visualising hidden layers in neural networks and the importance of accuracy, open issues like the lack of standards and shared benchmarks between the researchers were remained to be replied [29].

Given the increasing promises of deep reinforcement learning moving forward to real-world applications as opposed to theoretical contributions and lab-based systems given their greatest advantage to simplify intricate tasks in controlling the autonomous system in transparent vehicle control. Furthermore, numerous techniques were enforced to generate reinforcement learning algorithms interpretive effectively. The techniques could be gain a close attention to the research community in artificial intelligence and machine learning with different setup alternatives and dominences [30].

7. Case Studies and Applications

Next, we consider another case study involving the autonomous control of repositioning the satelloon in LEO, a concept being proposed by T. Sagool and C. McInnes for refueling the su avanzado MIGOL sol en casa delar futurasferoción transfer orbit cargo la terrestre oviregolith on the dark side of the moon in order to satisfy an extended shelf life requirement, titre interna cut when the departing moon is no longer available for recharging. A refueling process is generally hard to schedule and maintain effectively unless earth maneuverable satellites or teleoperators are used, but these themselves may not be an option anyway, e.g., if the conflict performed between an hypercubesat and the human of the turbulent and varied LU588 ascent part in the scenario that is considered here [31]. Therefore, the following motivational case study is relevant for various SH and science seem beyond, see also [11,9], around true anomaly 311.8127 deg t=0. In this times Mach number and flight-path angle trajectory plot dataset represented by the pure blue, dashing strips represents the expectant $k - t$ model distribution at the flying satelloon at a certain time. The pure blue segments going through each of the ellipsoids alongside of the false colored flight-path angle dataset can thus be visualized in-between the two plotted outermost expectant distributions of the predicted features for the highest SBL Flu Set, i.e., neighbors 163 and 1631 from Tables II–V. At a certain encounter time t1 in seconds, the occasion is described for this ensemble space-ship distributed Classroom of 188 × 5 = 970 expectant vehicle states falling on the 4, 5, 9, 14, 20, 29, 29, 27, 23 Being within ±3σ space, ground for the missions than J., 2026, 361, 231, T. F. T. F., 2012, I. P. III. A. I. A. I. E. 1999, 16, 13, W. B. et al. 2003, 38, 6123, A. Th. 1982, C. W. 45, 333, L. J. M., G. 2015e, 2019 sistem 2002 is denoted here for the System Earth Satelloon Nutation Relaxation, S. E. SNRO, more easily, which utilizes titled queel blanc ın our paper today_ATTRIBBLUOTSOC.

The transparent refueling process is simple for the ESA data case study, as illustrated in [32], allowing us to focus more qualitatively on what the k-t Models tell us and highlight some of the expressive and physical relationships displayed by the SBL framework. The SBL k-t Models in question are trained on input weights from the HR and MR PFs for refueling events 282, 387, and 394 with around three varied usable mass transits, arriving at fly-by speeds of 106, 107, and 108 m/s and using a settled mass flow of 0.6 kg/s into a perigee MGA trajectory [9]. The transits all end a few degrees after reaching a settled and fixed decision point of RA=105 deg. They all begin at time h=0.0655 MJD2k and have a 50 km GA after about 44 revs as MI passes. The key visualisation tool is the k-t Model's under distribution speed and flight-path angle vs time trajectories, which can be divided by the average total kinetic and potential energies of the Earth orbit at refueling in order to describe the states in units of T. We colour code the data points on the MR orbits by their k-t Model prediction errors in the speed and flight-path angle graphs, in such a way that data closest to the SBL circumscribed ellipses at the origin will be the darkest and reduced data further away the lightest. A double mass will be used for the fitted high-residual EKF orbit (light violet) in comparison to the real state (dark violet), with this being indicated by the large overlap. Of the six refueling combinations shown in the ELSA data, the dark blue and magenta plotted times correspond to the known combinations 282/395 and 394/161 obtained through direct back propagation of the orbit.

7.1. Real-world Examples of Transparent RL in Autonomous Vehicles

Two different scenarios for the justification of the vehicle's decision to overtake a slower car are displayed in the top two rows of the figure. In the leftmost situation, a bottleneck occurs in the right lane. Since the traffic in the left lane can move faster the AV overtakes this vehicle. In the right image the front vehicle changes to the left lane and forms an obstacle for the AV.

The AV thus needs to perform an overtaking manoeuvre. In the bottom of Fig. 10, the significant states are highlighted by the baseline saliency map and explanation policy which is displayed in its initial and final state. We learn different KiJ. The LRP summarization is provided in the middle row of Fig. 10 and is essentially the NiJ saliency-pruning which is learned simultaneously from the considered experience. In the two depicted situations, the explanations differ significantly. This is highlighted when comparing the top and bottom right images.

With the increasing autonomy of intelligent robotic systems, transparency to the user becomes an important requirement for the use of such systems. Lack of transparency makes it hard for a user to understand the system's decision making (Schroeer et al., 2020) and thus also impairs the capability of the user to trust the system. In particular, an autonomous vehicle may need to explain its decisions (e.g., why it will overtake another vehicle). Here, we aim at explaining the decisions of a reinforcement learning agent in terms of their expected consequences [33]. This is an important aspect because a RL agent might expose apparently strange behaviour and thus lose user acceptance and confidence [10]. The vehicle may need to take decisions that are breakable for reasons such as traffic regulations, even if the agent would have a better estimation of the expected outcome. Now, imagine the agent would ignore these traffic regulations and take another decision. In these cases, the agent should be able to explain that the alternative action has a larger expected reward and would thus increase the safety of the traffic flow. In the scenario presented in Fig. 10, we will focus on the task to find explanations for an overtaking manoeuvre, which is an important task for our agents. We will show these explanations in words and by highlighting relevant regions of the egoimages that were used as input to the RL agents. We will not present the intermediate state-values nor the activations of the policy-network's neurons presented in previous sections here. These kinds of visualizations are displayed in our previous publications [23].

8. Challenges and Future Directions

Controlling an autonomous vehicle is a challenging task that involves knowledge acquired during training. Being an internet-driven-self-acting framework, reinforcement learning (RL) provides the proper toolkit for learning and controlling as well. Recently, numerous models based on the use of RL have attempted to control an AV in real-world traffic situations. [34]. Knowledge about actions taken by the AV is crucial for humans inside and outside the vehicle.

This inside supervising process requires continuous efforts from the drivers and can cause problems. This explains why approving AVs under legally required conditions turns out to be quite difficult. Therefore, it is evident that an AV should be transparent during its decision-making to be accepted by other stakeholders involved in its control.

The alliance of reinforcement learning and deep learning is highly promising for autonomous driving tasks. When considering Explainable Artificial Intelligence-based (XAI) actuators, this partnership becomes complex due to the high-dimensional data and the non-linear function optimization employed by deep neural networks. [ref: 107dd82a-8493-4711-810e-19cab4141da5, ref: 53467337-a40d-4e02-a3a0-1b2d5997568d]. In this review, different XAI techniques for deep architectures are categorized according to the taxonomy presented by Montavon et al. In addition, a subdomain stock taking for XAI in autonomous driving applications is provided where challenges and probable solutions are discussed. Although this review targeting autonomous driving concerns mainly visual data, the presented methodologies can be ported for non-visual data as well.

8.1. Ethical Considerations

[31] The ethical and societal problems of taxi services and other intermediary platforms, robotaxis, and other on-demand transportation services continue to attract more attention. These are problems that may be related to the future of society's mobility. Ethical problems include how robots should be designed to be fair to customers, respectful, and safe when shared with others, respecting the rights of others. With the rapid development of the technology, many technical barriers are not the key to the problems of autonomous vehicles, but the moral obligation of artificial intelligence (AI) systems is the major social problem. It is a more in-depth explanation that AI explains the specific basis and reasons for its decision to carry out appropriate behavior.[35] An AI system that perfectly replaces the human intelligence and can be responsible for its actions needs to understand the inherently ambiguous human morality. The moral philosophy of the running code program to reflect its moral mission is not unattainable at the present time. An AI system that automates the learning of moral behavior based on massive amounts of human ethical knowledge can have ethical decision-making capabilities. The driver-less car project, known as the electronic breeze by the government-civil-industry to shorten the country's 60% of the loss of casualties due to human error, and to ensure significantly more security, to ensure higher air auto-

centric Turkey's needs to be developed. Preferences are likewise to be created accordingly. But the use of the road, new-fangled ideas in the circulation environment should be such a priority to work from now on in the vicinity of the great changes in morality is a necessity. An attractive research gap has begun due to the loss of the first car to get a driver-less 2454 person in our country in three separate accidents, and the giving way to the pedestrian crossing stream in the regulation of the drivers by the pedestrian crossing.

8.2. Technical Challenges and Solutions

End-user interpretability and actionable insights, and path dependence are common yet important technical challenges in the development of reinforcement learning models for autonomous vehicles. Model-based reinforcement learning (MBRL) algorithms can be used to simplify the policy search methods and to provide a simpler interface for the state–action value prediction model (see Cao et al., 2015 [36]), Environmental context-dependent path-dependent complex and non–linear, representing the end users' explanation framework and related features can guide the decision making in terms of the complex interaction of human end–users and deep learning components, is developed with a hierarchical model. Then Similarly, end–user interpretability and actionable insights development can be also a promising technical–methodological solution through the development of hierarchical models. Path-dependent model learning uses model predictive control, an end-user-friendly model generation method, which is compatible with the end users' interpretability and actionable insights demands in complex, non-linear, and real-time systems. The results of the case study show that the structural approach has a minimum of 15% test MAPE and a maximum of 3.4% test MAPE. In the automobile dataset, the interpretation framework provides approximately double the local importance local sensitivity to the policy description. Continuous Learning require the development of both automated hyper-local interface and automated hypothesis generation. Chore and attribute handling provides throttle stopping solutions with 94.42% (43.62) and an average of 92.55% (39.73 MAPE) read time [2].

Explainable reinforcement learning models face various technical challenges, many of which have been addressed by methodological strategies and technical-methodological solutions. Although the focus has been on the technical explanations of model decisions, adding end-user interpretability and actionable insights should also be emphasized. Therefore, it is necessary to deal with the technical challenges of the real-world autonomous vehicle operation environment [19].

9. Conclusion and Summary

In doing so, we make the following contributions: We adapt a prototypic convolutional neural network (CNN) and long short-term memory (LSTM) re- current neural network (RNN) - architecture to justify selfdriving car decisions via textual explanations. By guiding a reinforcement learning (RL) agent with this architecture in a simulation environment, we demonstrate that the agent makes locally interpretable, adversarially robust and situationally-aware decisions. We then visualize the internal functioning of the RL policy to validate the textual- explanation-model's ability to foresee critical states. We use experiments and an end-to-end simulator to demonstrate that the system can theoretically offer ethicality and disparity in demographic fairness in practice. Finally, when our proposed approach exceeds human-level performance, we interpret it with crowdsourcing to show that the transparent and interpretable RL agent skilfully avoids ambiguous situations [7].

In a world where self-driving cars are to autonomously make decisions in complex urban environments, it has become increasingly important that these decisions are not only safe and legally compliant, but also transparent to stakeholders. This latter property, refered to as explainability, has been widely recognized as an essential component for trustworthiness, litigation, regulatory compliance and the overall societal acceptance of self-driving vehicles [21]. While traditional planning systems consider the legality and Safeness of a candidate trajectory, many handcrafted heuristics are used to restrain them. In contrast, reinforcement learning learns human-like behavior from human demonstrations and uses these demonstrations to impute heuristics for fairness, ethicality and transparency. In this paper, we employ a reinforcement learning model that does not rely on any human demonstrations, but provides the same level of transparency as those which do [5].

9.1. Key Findings and Contributions

Admirably, ALA, by explicitly enforcing the crumple zone treatment, drives the agent to make decisions that have valuable human-like interpretations. This is indeed a difficult and ambiguous knowledge to learn from the demon: the physical reasoning for the crumple zone requires the agent be aware of pathway dependencies, inertial reactions, and have a specific and unusual intention—a seemingly tall order for a facial reading outcome recipient. Therefore, ALA provides an interesting perceptual basis for consistent and population-appropriate explanations of the policy. Nevertheless, almost as a secondary question, the human interface of this research will also be evaluated. The human advantage of the PPO model is somewhat low. There are very minor advantages in the case of ghost parameters. This must be evaluated as far as the population is representative of the confidence of the demonstration and the performance of all automatic agents. We will also show that both the sumo assistance and the long-term benefit of the population are determined by the number of GPNVs using very similar reasoning, except that the kinetic mode models that we use are trained on different human policy trajectories. [37].

The objective of the research into explainable autonomous car learning agents from human driving data is to design autonomous agents that can make understandable driving decisions. By capturing, modeling, and interpreting human driving data, we aim to develop more human-like and interpretable autonomous agents. This can be achieved using several directions of research and several goals. However, our findings show that there is no unified solution. We propose the Assess, Learn, and Act (ALA) model to address these fundamental challenges relying on the integration of a direct behavioral comparison and model-grounded intervention. With the model-grounded intervention, we can ensure human-likeness and promote vehicle interpretability without directly interacting with the autonomous agent [38]. Our findings are as follows: When learning the model and learning from the model, we explicitly ensure important human driving features in the final autonomous agent. Intuitively, the physically realistic model and direct behavioral comparison improve the ability of the ALA method to ensure that autonomous-driving decisions from the agent match the human driving recognizable criteria such as freeing the crumple zone. Human drivers (in particular, old human drivers) reduce the possibility of using important human driving criteria, but multi-model GPNV effectively prevents this behavior. Observing the physical behavior of cars, the multi-model GPNV imposes new types of contextual guidance, each agent receives a distinct spatial conditioning that highlights a particular menu of factors. This is a really important result because it is ultimately this forcing that drives the physical realism of the agent’s decisions.

9.2. Implications for Future Research

To this end, future models can be developed using ANN-based techniques. As ANN models can mimic a human driver's behavior and can also improve driving performance under different high-performance driving safety criteria, ANN models are used at ANN-RL-based autonomous vehicle control in the existing model. This RL-based vehicle control would be a suitable benchmark model for comparison with future ANN model performance and wheel slip reduction. In addition, in many real-time and high-speed maneuvering scenarios, developing race car control strategies to reduce wheel slip would be very beneficial. Using this RL technique, the model will be able to handle various driving situations, as the proposed ANN model is performance-oriented, which is mainly set up for low slip [21]. Based on the promising results, significant future work is needed to create a fundamentally improved and a more explainable RL control solutions for safety in ground robot control.

Sophisticated training techniques are being applied to artificial intelligence systems in a wide range of applications, and one of the main areas is autonomous vehicle (AV) control. In this chapter, we focused on using artificial neural network (ANN) models that are trained using reinforcement learning (RL) algorithms for AV control systems. As these models are shown to mimic the behavior of human drivers [37], in addition to their excellent driving performance in a wide range of possible driving scenarios, it is a good idea to think of different ways that safety can be improved using explanation systems. Given any potential problems associated with vehicle control decisions made without fulfilling safety requirements, it is essential to explore different strategies and solutions. These strategies can help to ensure the maximum use of efficient learning methods for ultimate control.

Reference:

1. Tatineni, S., and A. Katari. "Advanced AI-Driven Techniques for Integrating DevOps and MLOps: Enhancing Continuous Integration, Deployment, and Monitoring in Machine Learning Projects". *Journal of Science & Technology*, vol. 2, no. 2, July 2021, pp. 68-98, https://thesciencebrigade.com/jst/article/view/243.
2. Prabhod, Kummaragunta Joel. "Advanced Techniques in Reinforcement Learning and Deep Learning for Autonomous Vehicle Navigation: Integrating Large Language Models for Real-Time Decision Making." *Journal of AI-Assisted Scientific Discovery* 3.1 (2023): 1-20.

3. Tatineni, Sumanth, and Sandeep Chinamanagonda. "Leveraging Artificial Intelligence for Predictive Analytics in DevOps: Enhancing Continuous Integration and Continuous Deployment Pipelines for Optimal Performance". Journal of Artificial Intelligence Research and Applications, vol. 1, no. 1, Feb. 2021, pp. 103-38, https://aimlstudies.co.uk/index.php/jaira/article/view/104.