ScienceDirect

Expert Systems with Applications
Volume 268, 5 April 2025, 126319

Intelligent vehicle driving decision-making model based on variational AutoEncoder network and deep reinforcement learning

Shufeng Wang [a], Zhengli Wang [a], Xinkai Wang [a], Qingwei Liang [b], Lingyi Meng [a]

Show more

Share Cite

https://doi.org/10.1016/j.eswa.2024.126319
Get rights and content

Abstract

In this paper, an end-to-end driving decision-making model is proposed for intelligent vehicle, utilizing a Variational AutoEncoder (VAE) network and Deep Reinforcement Learning to address the challenges in complex and dynamical driving environments. Firstly, the traffic environment image features are extracted by VAE network, which can effectively reduce the amount of data input and improve the learning efficiency. Secondly, the Soft Actor-Critic (SAC) algorithm is improved through the application of TD error value constraints, N-step learning, etc. Then driving risk field and rule constraints are introduced into the improve SAC algorithm. Based on the real-time driving risk field, the skipping frame method can enhance learning efficiency, and the rule constraints can reduce the dangerous actions in the output of the algorithm. In order to verify the effectiveness of the model, in the CARLA simulation platform the models of scenario and algorithm are established, and the simulations are carried out. The results show that using decision-making model built by the proposed algorithm, the average driving distance by the intelligent vehicle has been improved by 91.37%, the average reward value of the task has been increased by 132.04%, the average success rate of the task has been improved by 46.56%, the training time is also significantly reduced. It demonstrated that the proposed decision-making model provides a significant improvement in driving safety and learning efficiency.

Introduction

In recent years, the development of intelligent vehicles has received extensive attention from academia and industry (Ziakopoulos et al., 2023). By means of sensors and communication technologies, intelligent vehicles are able to perceive the traffic environment in real time, make intelligent decisions, plan motion path and control vehicle safe driving, which can greatly reduce traffic accidents and improve traffic efficiency (Mesch and Dodel, 2023, Bhaggiaraj et al., 2023, Karmakar et al., 2021). Driving decision-making is the core technology of intelligent driving, accurate and efficient driving decision-making can improve traffic safety, alleviate traffic congestion, and provide more comfortable driving experience (Nees, 2019, Khare and Jain, 2023, Niu et al., 2020). However, in real traffic scenarios, driving decision-making is still a challenging task due to the complexity and uncertainty of the driving environment caused by dynamic changes in the surrounding traffic participants, traffic signals, and weather, etc. Currently, there are two main categories of methods used for autonomous driving decision-making: rule-based methods and learning-based methods (Yang et al., 2023, Zhang and Sun, 2024).

Rule-based methods rely on expert knowledge and require setting a series of working modes and transition conditions between modes for specific scenarios/tasks (Liu et al., 2019). Rule-based methods are typically represented by Finite State Machines (FSM) and Decision Tree (DT) (Cui et al., 2021).

The FSM algorithm consists of a series of discrete states, actions, and specific conditions that trigger an action or a state transition when specific conditions and states are satisfied, which are also known as rules for state transfer (Khosravi et al., 2023). In an autonomous driving system, the state describes the vehicle itself and its surroundings, and the state transfer depends on the ability of the vehicle to act accordingly to different driving scenarios. The core of these algorithms is to ensure an orderly and safe driving while complying with traffic rules (Barman et al., 2016). Wang et al. (2019) established an independent lane-changing model based on the minimum safe distance. The model triggers a lane change when the distance to the vehicle in front is less than a defined safety limit. Zhang et al. (2021) proposed a method that combines cognitive technology and decision-making to effectively recognize traffic signs, enabling intelligent vehicles to make better choices in complex situations such as intersections. Xing et al. (2021) combined System Theoretical Process Analysis (STPA) with Finite State Machines (FSM) to model vehicle states and environmental conditions, which is capable of identifying more hazardous events. The FSM model is widely used due to its simplicity and ease of use, but it ignores the uncertainty of the dynamic environment and has limited reusability, only suitable for decision-making tasks in relatively simple scenarios.

As a kind of decision-making model based on a tree structure, Decision Trees can assist vehicles in making sequential decisions and ultimately deriving appropriate driving strategies through the evaluation of multiple conditions and the top-down ""polling"" mechanism. For instance, Ashraf et al. (2021) and Ftaimi et al. (2021) applied Decision Trees models to extract pre-crash rules or evaluate the risk of collision, which provide a foundation for the decision-making of self-driving vehicles. Amin et al. (2022) developed decision-making algorithms for lane changing of vehicles on highways and suburban roads using Decision Trees. The Decision Trees model can handle relatively complex samples because it is a nonparametric model, and it offers fast computational speed and highly interpretable results. Unlike finite state machines, the control nodes of Decision Trees can be reused. However, this type of model requires the redefinition of the decision network for different scenarios. Therefore, as scenarios become more complex, the control logic may also become intricate.

Rule based methods exhibit good reliability and interpretability, but it is difficult to design comprehensive solutions that are universally applicable to real-world scenarios (Mirchevska et al., 2018). In addition, these methods often result in vehicle behavior that may lack smoothness and human like driving characteristics (Zhou et al., 2023, Negash and Yang, 2023).

Learning-based decision-making methods train the vehicle's behavior using advanced algorithms such as Deep Learning (DL), Reinforcement Learning and Deep Reinforcement Learning. The decision-making process does not require the creation of a rule base but can be achieved directly through autonomous learning from environmental samples. There are two types of solutions: decomposed and end-to-end (Chib and Singh, 2023). Decomposed decision-making architecture facilitates collaborative engineering implementation by multi-task decomposition.

However, it is challenging to be applied in complex scenarios due to complex structure and the accumulation of errors across multiple modules (Kendall et al., 2019). Whereas, end-to-end decision-making architecture provides a solution to complex scenarios, it utilizes neural networks to directly establish the mapping from on-board sensors to vehicle controllers. There are two main branches of end-to-end decision-making: supervised learning and Deep Reinforcement Learning. Supervised learning refers mainly to imitative learning for human drivers in specified driving environments (Rane et al., 2024). Supervised learning is an approach that uses labeled datasets and correct outputs to train learning algorithms how to classify data or predict an outcome. The decision-making algorithm is trained using the input information (vehicle and surroundings) and the corresponding decision outputs (human driver), which can make the vehicle decisions consistent with that of a human driver. The main challenge of this approach is that it requires a large amount of natural driving training data and ensuring the quality and diversity of the training data， otherwise it is difficulty to handle new situations that are not covered in the data(Wu et al., 2023).

To address this challenge, Deep Reinforcement Learning (Deep RL) approach is developed, which blendsreinforcement learning ↗techniques with strategies fordeep learning ↗. Reinforcement learning is dynamically learning with a trial and error method and improves with ongoing interaction with the environment, which potentially surpass human performance and ensures the system adaptable. Deep learning uses deep neural networks to learn and extract features from input data. Deep Reinforcement Learning is a data-driven self-optimizing algorithm that constructs exploration-development and trial-and-error mechanisms to autonomously search for feasible control actions and optimize the policy (human-in-the-closed-loop), master complex decision-making processes through direct interaction with the environment, which can alleviate the problems arising from the aforementioned algorithms (Silver et al., 2018). Deep RL algorithms can be categorized into three principal types: value iteration, policy iteration, and actor–critic models.

In the early stages development, value iteration-based algorithms gained popularity, which aim to establish a value function through bootstrapping and choose the action that maximizes the value function at any time, the most representative algorithm is Deep Q-learning (DQL). However, DQL algorithms suffer from the dimensional catastrophe problem as the state and action spaces grow large during interaction with the environment. As representative algorithms based on value function optimization, DQN and its improved algorithms such as dueling DQN, double DQN, and dueling double DQN (D3QN) are popular (Mnih et al., 2015). For instance, Liao et al. (2020) utilized DDQN to learn vehicle overtaking strategies on highways, demonstrating faster convergence speed and better control performance than DQN. Xu et al. (2022) employed the Rainbow DQN algorithm to build a vehicle decision-making model, achieving larger reward values and greater improvements in passing success rate and average vehicle speed compared to DDQN. However, these algorithms perform poorly in solving problems in continuous action spaces.

To address the above issue, the Policy iteration-based RL algorithms are developed, which aim to establish a differentiable policy function and update the policy for performance optimization using policy gradient algorithms. The representative algorithms: the proximal policy optimization (PPO) series gained popularity, which includes trust region policy optimization (TRPO), PPO, etc. Compared with value iteration algorithms, policy iteration algorithms can handle both discrete and continuous action problems. Siboo et al. (2023) opted for two continuous action space Deep Reinforcement Learning (DRL) algorithms: Deterministic Policy Gradient (DDPG) and PPO to address complex autonomous driving challenges. Guan et al. (2020) proposed Model Accelerated Proximal Policy Optimization (MA-PPO) to accelerate the learning process in terms of sample efficiency and improve the efficiency of vehicles passing through intersections without traffic lights. But long training times and model instability are still needed to further improve.

In recent years, there have also been many advances from Actor-Critic frameworks, including Advantage Actor-Critic (A2C), Asynchronous Advantageous Behavioral Critic (A3C), and Soft Actor-Critic (SAC) algorithm (Sivamayil et al., 2023). Actor–Critic-based RL algorithms integrate the characteristics of both value iteration and policy iteration, this coupling combines the flexibility of policy optimization with the stability of value function estimation, enabling efficient learning in more complex tasks and environments. Among them, the SAC algorithm combines the advantages of Actor-Critic algorithm and model-based strategy optimization algorithm, it efficiently solves the Reinforcement Learning problem in continuous action space, especially for high-dimensional state and action spaces. Huang et al. (2023) compared Deep Reinforcement Learning for Autonomous Driving in the urban driving scenarios, the success rate and the duration of the success episodes of SAC algorithm is better than that of PPO algorithm, Gao et al. (2021) proposed a modified SAC algorithm for autonomous driving, used ResNet-34 as actor and critic network architecture to increase the policy's robustness and verified. Cai et al. (2020) proposed a robust drift controller without explicit motion equations based on SAC algorithm, the drift control problem is formulated as a trajectory following task, where the error based state and reward are designed. The proposed controller is shown to have excellent generalization ability.

Furthermore, end-to-end algorithms learn driving decisions directly from real-time sensor data, such as semantic camera sensors and vehicle IMUs. However, the direct application of Deep Reinforcement Learning often results in longer training times and slow convergence due to the large amount of input data. To solve this problem, Li et al. (2023) developed a network that integrates separable convolution and transformation modules to extract image semantics from trajectory data time series. They combined this network with Deep Reinforcement Learning for decision-making and confirmed its superiority by evaluating it with three lane-changing scenarios. Chen et al. (2019) incorporated a variational auto-encoder (VAE) capable of generating latent codes as the state representation for the RL policy, enhancing interpretability and generalization. Chen et al. (2020) compared the effectiveness of DDQN, TD3, and SAC decision-making algorithms using the CARLA simulator output converted to a bird's-eye-view pixel-style image as model input. The results showed that SAC performed best in complex scenarios.

Despite the advantages and many achievements of Deep Reinforcement Learning algorithm in end-to-end decision-making, it still faces many challenges (Wang et al., 2024, Huang et al., 2021, Neftci and Averbeck, 2019). In most cases, the interaction between the intelligence vehicle and the environment is inefficient, and model training consumes a large amount of computational resources and time. Additionally, there are limitations in understanding complex environments, inadequate risk perception, and the potential for dangerous action output.
To address the aforementioned challenges, using complex urban road intersections as scenario, based on SAC algorithm, an end-to-end algorithm based on variant autoencoder (VAE) network and Deep Reinforcement Learning for driving decision making of intelligent vehicles is proposed in this paper, the main contributions of the work are as follows:

(1) An improved version of SAC algorithm is proposed from three key aspects: the utilization of empirical pool data, addressing gradient explosion, and improving value estimation. This is achieved through the implementation of Temporal-Difference-error value constraints, a linearly decaying learning rate, Gradient clipping and orthogonal initialization, and N-step learning.

(2) To address the challenge of large-dimensional image features, using semantic segmentation of semantic sensors to obtain environmental data, Convolutional Neural Network (CNN) and VAE networks are employed to extract image features, and the dimensionality-reduced image features are integrated into the hybrid state space of the improved SAC algorithm, which can enhance the convergence speed of the decision-making model and improve learning efficiency.

(3) The state space, action space, and reward function are designed based on the driving experience of human drivers and traffic rules, using the improved SAC algorithm, the decision-making model of intelligent vehicles is established. Based on the real-time driving risk field, the skipping frame method can enhance learning efficiency, and the rule constraints can reduce the dangerous actions in the output of the algorithm, which can improve the learning efficiency of the algorithm while ensuring security.

Section snippets

SAC algorithm

The Soft Actor-Critic (SAC) algorithm is a Reinforcement Learning algorithm specifically designed to tackle decision-making problems in continuous action spaces. It is based on the framework of Deep Reinforcement Learning and combines elements from both policy gradient methods and Q-learning. Unlike traditional Reinforcement Learning algorithms, the SAC emphasizes maintaining a 'soft' policy during training. It introduces the concept of maximum entropy within the traditional actor-critic …

Driving decision-making model of intelligent vehicle

Aiming at improving learning efficiency and safe driving in complex traffic environments, an end-to-end intelligent driving decision-making model is established. Firstly, the image features of the traffic environment are extracted through the CNN and VAE network and integrated into the hybrid state space, then the driving risk field and rule constraints are introduced into the improved SAC algorithm to learn decision-making, and the decision-making frame is shown in Fig. 4.

During the training …

Experiment design

The simulation environment is built based on Town03 map in CARLA platform. According to the needs of vehicle decision-making, a two-way 4-lane scenario is set up for the intersection and its adjacent roadway. Urban road intersections are characterized by complexity, uncertainty and high interactivity, which can simulate various challenges in real traffic.

The number of environment vehicles in the scenario is set to 30, and the vehicles are randomly generated in the scenario and managed by the …

Comparative analysis of simulation results

Due to time constraints, a uniform training of 8000 rounds is conducted in the simulation analysis. Additionally, the output data of Deep Reinforcement Learning algorithm has some fluctuations. In order to make the output results more intuitive, the data of speed, displacement, and return value are smoothed using the Savitzky. Golay filter, which is available in the built-in library of Python.

In the selection of simulation comparison indicators, the cumulative reward value of the algorithm in a …

Conclusion

In order to cope with the challenges in complex dynamic driving environments, and to address the problems of low learning efficiency, dangerous action output, and insufficient risk perception ability of the current decision-making models, this paper proposes an end-to-end driving decision-making model for intelligent vehicles by using the VAE network and Deep Reinforcement Learning algorithms.

The image features of the traffic environment are obtained by CNN and VAE networks, and then the hybrid …

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. …

Acknowledgment

This work was supported by the Postgraduate Quality Professional Degree Teaching Case Base of Shandong Province under Grantno. SDYAL2023051. …

Recommended articles

References (45)

J. Chen *et al.*
SAANet: Spatial adaptive alignment network for object detection in automatic driving
Image and Vision Computing (2020)

Q. Cui *et al.*
A hierarchical framework of emergency collision avoidance amid surrounding vehicles in highway driving
Control Engineering Practice (2021)

M.A. Nees
Safer than the average human driver (who is less safe than me)? Examining a popular safety benchmark for self-driving cars
Journal of safety research (2019)

J. Wu *et al.*
Toward Human-in-the-Loop AI: Enhancing Deep Reinforcement Learning via Real-Time Human Guidance for Autonomous Driving
Engineering (2023)

A. Ziakopoulos *et al.*
Perceptions towards autonomous vehicle acceptance: Information mining from self-organizing maps and random forests
IATSS research (2023)

M. Amin *et al.*
Detection and Risk Analysis with Lane-Changing Decision Algorithms for Autonomous Vehicles
Sensors (2022)

M.T. Ashraf *et al.*
Extracting rules from autonomous-vehicle-involved crashes by applying decision tree and association rule methods
Transportation Research Record (2021)

B. Barman *et al.*
Neuro-Fuzzy Controller Design to Navigate Unmanned Vehicle with Construction of Traffic Rules to Avoid Obstacles
International Journal of Uncertainty Fuzziness and Knowledge-Based Systems (2016)

S. Bhaggiaraj *et al.*
Deep Learning Based Self Driving Cars Using Computer Vision

P. Cai *et al.*
High-speed autonomous drifting with deep reinforcement learning
IEEE Robotics and Automation Letters (2020)

View more references

Cited by (3)

LLM-augmented hierarchical reinforcement learning for human-like decision-making of autonomous driving
2025, Expert Systems with Applications

Show abstract

Lane clearance for emergency vehicle passage in connected and automated environment: A graph convolutional soft actor-critic method
2025, Expert Systems with Applications

Show abstract

A Comparative Study of Deep Reinforcement Learning Algorithms for Urban Autonomous Driving: Addressing the Geographic and Regulatory Challenges in CARLA ↗
2025, Applied Sciences Switzerland

View full text

© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

ELSEVIER

All content on this site: Copyright © 2025 Elsevier B.V., its licensors, and contributors. All rights are reserved, including those for text and data mining, AI training, and similar technologies. For all open access content, the relevant licensing terms apply.

RELX™