Automation and Machine Learning (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/autml.2025.060210
ISSN 2516-5003 Vol. 6 Num. 1

Active Suspension Control Based on DQN-LSTM with Integrated Temporal Feature Extraction

Li Chenyang[1,a], Li Wei[1,b,*], Gao Yanfei[1,c], Zhang Hongjia[1,d]

[1]School of Automotive Engineering, Shandong Jiaotong University, Jinan, 250357, China
[a]lcy200142@163.com, [b]liwei@sdjtu.edu.cn, [c]gaoyanfei.0402@163.com,
[d]zhanghongjia@sdut.edu.cn
[*]Corresponding author

Keywords: Active Suspension, Eep Reinforcement Learning, DQN-LSTM, Ride Comfort Optimization, Intelligent Control

Abstract: To improve vehicle ride comfort and address the temporal dependency inherent in active suspension control, this study proposes a reinforcement learning–based control algorithm that integrates a Deep Q-Network (DQN) with a Long Short-Term Memory (LSTM) network, referred to as DQN-LSTM. A two-degree-of-freedom vertical dynamics model is first established as the interaction environment for the algorithm. A reward function is then designed to minimize the root-mean-square (RMS) value of the vehicle body vertical acceleration, where the DQN is responsible for policy optimization, and an LSTM layer is incorporated to extract temporal features embedded in historical state sequences, thereby enhancing the controller's capability to predict and respond to road excitations. Simulation tests on Class B and Class C random roads are conducted in MATLAB. The results indicate that, compared with the passive suspension, the DQN controller reduces the RMS of the body vertical acceleration by 12.78%, whereas the proposed DQN-LSTM controller further reduces it by 25.11%, yielding a notably smoother system response. These findings demonstrate that the proposed algorithm effectively captures temporal characteristics and exhibits strong adaptability, robustness, and application potential under stochastic road excitations.

1. Introduction

As the number of vehicles in use continues to increase, the demand for ride comfort and vehicle ride quality is growing accordingly. As a key component of the vehicle running system, the suspension system has, as one of its core functions, the effective attenuation of road-induced vibrations transmitted to the vehicle body, thereby improving both ride comfort and handling stability [1-2]. Compared with passive suspensions with fixed parameters, which can only achieve a compromise between comfort and stability, active suspensions utilise actuators to regulate the suspension force in real time to compensate for road excitations, thereby improving ride comfort and safety simultaneously [3-4]. Consequently, the development of efficient active suspension control strategies is of great significance for enhancing vehicle running performance and improving occupant comfort.

A wide variety of active suspension control methods have been proposed, each with its own limitations. Classical linear control strategies (such as PID and LQR) rely on accurate system model parameters, and their performance often degrades when suspension parameters or road conditions vary [5–7]. Intelligent control methods such as fuzzy control can handle non-linearities to some extent, but their control performance depends heavily on expert-designed rule bases and lacks systematic design guidelines [8-9].

To address the problem of active suspension control, deep reinforcement learning (DRL) methods have been introduced to achieve adaptive optimisation of control policies. Ming et al. [10] proposed a control method based on Deep Q-Networks (DQN), in which a neural network approximates the Q-value function and uses suspension states such as body acceleration and suspension deflection as inputs to autonomously generate control actions. Wang Zihao et al. [11] proposed a semi-active suspension control method based on DQN, where a reward function integrating both ride comfort and handling stability is designed to enable self-learning and reproduction of the control policy, thereby improving control performance under different road conditions. Compared with traditional methods, DRL-based approaches do not require an accurate model; instead, they obtain an optimal policy through large-scale training, enabling the suspension system to adaptively adjust itself online. As a result, they exhibit strong robustness to parameter uncertainties and varying road conditions, and can maximise ride comfort while ensuring safety [12-13]. However, the classical DQN architecture does not possess an internal memory mechanism, and its decisions are made solely based on the current observed state. Given that the dynamic response of the suspension system is closely related to the time series of road excitations, a DQN model that lacks temporal information struggles to capture the evolution patterns of the system accurately [14–16].

To overcome this limitation, this study incorporates a Long Short-Term Memory (LSTM) network into the DQN framework to enhance its capability for temporal feature extraction. Specifically, one of the fully connected layers in the conventional DQN is replaced by an LSTM layer, enabling the network to integrate long-term historical observations through recurrent connections. The improved model can thus capture the temporal correlation characteristics of road excitations, estimate the current system state more accurately, and generate appropriate control strategies, thereby significantly improving the response speed and control accuracy of the system under dynamic road inputs.

On the basis of a two-degree-of-freedom vehicle model, this paper designs a DQN control algorithm integrated with an LSTM layer. The algorithm takes the body vertical acceleration, suspension dynamic deflection and tyre dynamic displacement as state inputs, and outputs the discrete control force of the active suspension system. To improve vehicle ride quality, a reward function is constructed that minimises body acceleration while simultaneously accounting for suspension travel and tyre–road contact, thereby realising multi-objective collaborative optimisation. The main contributions of this paper are as follows:

(1) A DQN–LSTM control algorithm that fuses temporal feature extraction is proposed. By using the LSTM layer to memorise historical state information, the algorithm enhances the system's perception of and adaptability to dynamic variations in road excitation.

(2) Through systematic simulation experiments and parameter analyses, it is verified that the proposed algorithm can significantly improve vehicle ride comfort under road excitations of different grades, providing a design reference for the deployment of deep reinforcement learning in vehicle suspension control.

The remainder of this paper is organised as follows. Section 1 presents the design of the core control algorithm and details how the DQN–LSTM framework is constructed to address the aforementioned temporal dependence problem. Section 2 provides simulation validation and systematically evaluates the effectiveness of the proposed control algorithm from three perspectives: overall performance, temporal behaviour mechanisms and real-time capability. Section 3 concludes the work and outlines future research directions.

2. Design of a DQN–LSTM Control Algorithm Incorporating Temporal Feature Extraction

This chapter provides a detailed description of the DQN–LSTM control algorithm designed to address the problem of temporal dependence in suspension control. As the core part of this study, it presents a systematic introduction covering the network architecture, temporal modelling, parameter configuration and training deployment. By combining the feature extraction capability of deep learning with the dynamic characteristics of the vehicle system, an intelligent control method is constructed that can effectively suppress vehicle vertical vibrations and enhance ride comfort and stability.

2.1 State Input Modelling and Feature Pre-processing

The core state of the vehicle suspension system consists of three physical quantities: the vertical acceleration of the sprung mass, the relative displacement between the sprung and unsprung masses, and the relative displacement between the unsprung mass and the road surface. At time instant k, these three observations are combined into the vector:

$$s_k = \begin{bmatrix} \ddot{z}_2(k) \\ z_2(k) - z_u(k) \\ z_u(k) - z_q(k) \end{bmatrix} \tag{1}$$

This vector fully captures the instantaneous information on body vibration and tyre–road contact performance. Since these three physical quantities have different dimensions, directly feeding them into the neural network would lead to gradient imbalance and slow down convergence; therefore, each component is first standardised to zero mean and unit variance:

$$\tilde{s}_{k,i} = \frac{s_{k,i} - \mu_i}{\sigma_i} \tag{2}$$

where μ_i and σ_i denote the mean and standard deviation of the i-th component of the observation, obtained by computing statistics over all $s_{k,i}$ in the training data. After normalisation, a fully connected layer maps the three-dimensional input into a 128-dimensional hidden space, followed by a ReLU activation to extract non-linear features:

$$h_k^{(0)} = ReLU(W^{(0)}\tilde{s}_k + b^{(0)}) \tag{3}$$

where $W^{(0)} \in \mathbb{R}^{128\times3}$, $b^{(0)} \in \mathbb{R}^{128\times3}$. In this way, the complex coupled non-linear relationships are decoupled, and it is ensured that the subsequent network receives inputs following a stable distribution.

2.2 Temporal Dependence Modelling and LSTM Network Design

The dynamic response of the vehicle suspension system to road excitations exhibits pronounced temporal correlation: the current body state depends not only on the present road condition but is also closely related to the vibration history over a preceding time interval. To capture such time-series characteristics, a memory-capable network, namely the LSTM, is introduced in this study. The LSTM network is a special type of recurrent neural network, whose structure is shown in Fig. 1. Each LSTM unit contains a memory cell together with input, forget and output gates. These gating mechanisms effectively regulate what information is “remembered” or “forgotten”, enabling the network to retain relevant historical information over long time spans [17-18].

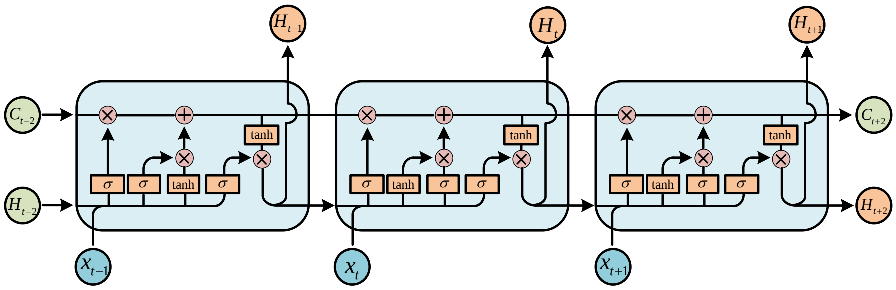

Figure 1: Architecture of the LSTM network.

In the active suspension scenario, when the vehicle is travelling on a bumpy road, neglecting historical states makes it difficult for the DQN to exploit past information to make optimal decisions for the current control. Therefore, by embedding an LSTM within the DQN network, i.e., forming a DQN–LSTM architecture, the agent is able to “remember” the effects of previous road excitations, thereby enhancing its capability to handle temporal, non-Markovian characteristics. In summary, introducing LSTM units into the DQN enables the joint consideration of past states and inputs, and offers clear advantages for tackling sequence-dependent vehicle suspension control problems.

2.3 Control Algorithm Architecture and Parameter Configuration

The LSTM output h_L is passed through a fully connected layer followed by layer normalisation (LayerNorm), and a ReLU activation is then applied to obtain the intermediate features:

$$h_k^{(1)} = ReLU(LayerNorm(W^{(1)}h_L + b^{(1)})) \tag{4}$$

Where $W^{(1)} \in \mathbb{R}^{128\times128}$. Subsequently, a 51-dimensional fully connected layer is used to output the action-value vector.

$$Q(s_k, \cdot; \theta) = W^{(2)}h_k^{(1)} + b^{(2)} \tag{5}$$

Corresponding to the action set $A = \{-50, -48, \cdots, 48, 50\}$. The reinforcement learning hyperparameters are configured as follows: the discount factor $\gamma = 0.99$ ensures a reasonable discounting of future rewards; the learning rate $\alpha = 10^{-5}$, together with the Adam optimiser, balances convergence speed and stability; the experience replay buffer capacity 10^6 and a mini-batch size of 64 are used to break data correlations; Double DQN is enabled by setting UseDoubleDQN = true to reduce overestimation of Q-values; the sequence length $L = 16$ ensures that the LSTM can capture sufficient historical information; the exploration strategy adopts an ε -greedy policy, where ε is linearly decayed from 1.0 to 0.05 to balance exploration and exploitation; and the target network is synchronised using a soft update scheme:

$$\theta^- \leftarrow (1-\tau)\theta^- + \tau\theta \tag{6}$$

Where $\tau = 10^{-3}$

To balance ride comfort, road holding and energy consumption, this study designs a penalty-type reward function as given in:

$$r_k = -[\omega_1 (z_2 - z_u)^2 + \omega_2 \ddot{z}_2^{\,2} + \omega_3 (z_u - z_q)^2 + \omega_4 u^2] \tag{7}$$

where the first term penalises fluctuations in suspension dynamic deflection, the second term penalises fluctuations in body vertical acceleration, the third term ensures tyre–road contact, and the fourth term constrains the energy consumption of the active control force.

During policy updates, the action value is iteratively updated according to the Bellman optimality equation:

$$Q(s_k, a_k) \leftarrow Q(s_k, a_k) + \alpha \left[r_k + \gamma \max_{a'} Q(s_{k+1}, a') - Q(s_k, a_k) \right] \tag{8}$$

where $r_k + \gamma \max_{a'} Q(s_{k+1}, a')$ denotes the current target value and α is the learning rate. The network parameters are updated by minimising the mean-squared TD error:

$$L(\theta) = E\left[r_k + \gamma \max_{a'} Q(s_{k+1}, a'; \theta^-) - Q(s_k, a_k; \theta)^2 \right] \tag{9}$$

In the Double DQN architecture, the current network is used to select the action, while the target network is used to evaluate its value:

$$y = r_k + \gamma Q\left(s_{k+1}, \arg\max_{a'} Q(s_{k+1}, a'; \theta); \theta^- \right) \tag{10}$$

Which further alleviates estimation bias.

2.4 Agent Training and Deployment Process

The rlSimulinkEnv interface is used to connect the Simulink active suspension model with the DQN–LSTM agent. The observation and action spaces are defined by rlNumericSpec([3,1]) and rlFiniteSetSpec(-50:2:50), respectively, and the agent is constructed using rlQValueRepresentation together with rlDQNAgentOptions. The training options are specified via rlTrainingOptions, including a maximum of 200 episodes, a maximum step number $\lceil T / T_s \rceil$, a stopping criterion based on the average reward reaching a prescribed threshold, and enabling visualisation of the training process. The command train(agent, env, trainingOpts) is then called to start training, during which the agent continuously interacts with the environment and updates its policy. After training, closed-loop simulations are performed using rlSimulationOptions and sim(env, agent, simOptions), and variables such as body acceleration, suspension and tyre displacements, and control force are recorded. A custom function plotSimulationResults is used to plot comparative time-response curves.

The training and update mechanism of the DQN–LSTM agent comprises three main components. First, an ε-greedy strategy is adopted for action selection to strike a balance between exploration and exploitation. Second, an experience replay buffer is employed to store interaction data and randomly sample mini-batches during training, thereby breaking data correlations and improving training stability. Third, a target network is used in parallel with the online network, and its parameters are synchronised at a fixed period to ensure smooth target value computation.

In terms of temporal feature extraction, the LSTM network maps the historical state sequence into a hidden vector with long-term dependencies, thus providing more accurate input features for Q-value estimation. On this basis, the agent's decision-making performance under complex road excitations is enhanced. During training, the loss function is still constructed based on the temporal-difference (TD) method, and the optimisation objective is to make the estimated Q-values approach the target Q-values, thereby achieving convergence.

3. Simulation and Analysis

To verify the performance of the DQN–LSTM control algorithm in a structured manner, the simulation analyses in this section are organised as follows. First, overall performance indices are compared under typical operating conditions to demonstrate its comprehensive superiority. Next, the temporal behavioural characteristics of the control force are examined in depth, providing direct evidence for the core innovation of “temporal feature extraction”. Finally, the engineering feasibility and real-time capability of the algorithm are evaluated.

To reflect the realism of actual vehicle driving on roads and to showcase the good ride comfort of the proposed DQN–LSTM active suspension system under complex road conditions, the simulation vehicle speed is set to 60 km/h. A combined B- and C-level road excitation signal is constructed for co-simulation, as illustrated in Fig. 2, where 0–10 s corresponds to a B-level road and 10–20 s corresponds to a C-level road.

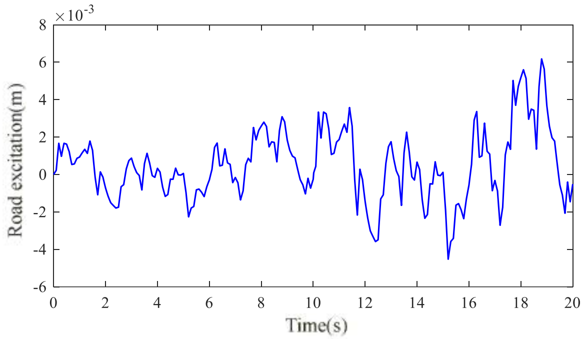

Figure 2: Road excitation signal.

To evaluate the ride comfort performance of the proposed DQN–LSTM active suspension control system, this study analyses ride comfort indices of the vehicle when the active suspension is subjected to road excitation disturbances, such as body acceleration, suspension dynamic deflection and tyre dynamic displacement, and compares the advantages and disadvantages of the DQN and DQN–LSTM control strategies. To enable a more comprehensive assessment of the active suspension control system, the vehicle simulation parameters are specified as listed in Table 1.

Table 1: Parameters of the quarter-vehicle suspension and tyre

Name	Parameter
Sprung mass/(kg)	450
Unsprung mass/(kg)	50
Damping coefficient/(N/(m/s))	2000
Tyre stiffness /(N/m)	192000
Spring stiffness/(N/m)	28000

The simulation results for body acceleration, suspension dynamic deflection, wheel dynamic displacement and power spectral density are shown in Figs. 3–6:

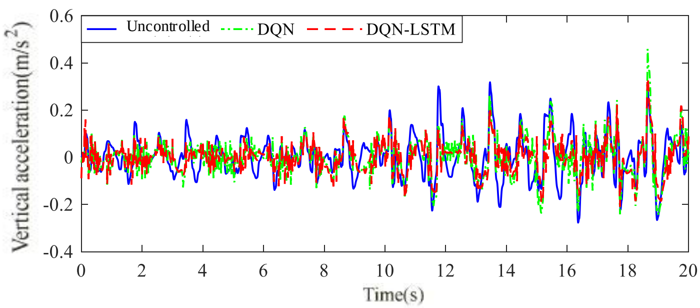

Figure 3: Comparison of vertical body acceleration

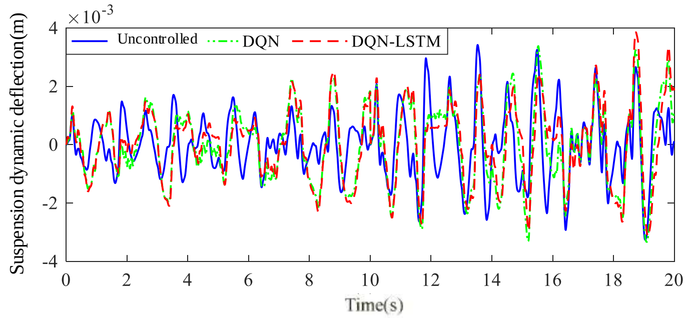

Figure 4: Comparison of suspension dynamic deflection

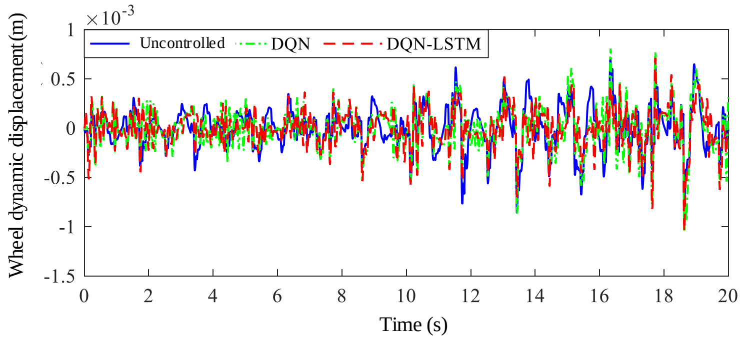

Figure 5: Comparison of wheel dynamic displacement

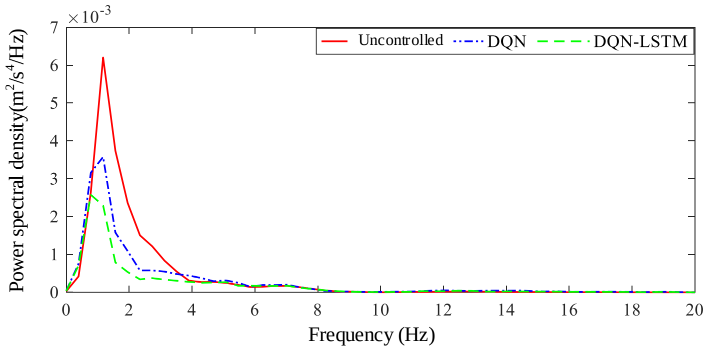

Figure 6: Comparison of power spectral density of vertical body acceleration

From Fig. 3 it can be seen that, as the road roughness level increases, both the vertical body acceleration and the vibration peaks increase noticeably. The growth in the peak vertical vibration of the body under the DQN–LSTM active suspension is smaller than that under the DQN active suspension, and the higher the road grade, the more pronounced the optimisation effect of DQN–LSTM becomes.

From Fig. 4 it can be observed that, for both DQN and DQN–LSTM active suspensions, the suspension dynamic deflection is larger than that of the uncontrolled suspension; however, the dynamic deflection of the DQN–LSTM suspension is smaller than that of the DQN suspension, and its increase remains within an acceptable range.

As shown in Fig. 5, the wheel dynamic displacement of the DQN–LSTM active suspension system is lower than that of both the DQN active suspension and the uncontrolled suspension, in terms of both overall vibration amplitude and peak vibration.

From Fig. 6 it can be seen that, in the main vehicle resonance region (1–2 Hz), both DQN and DQN–LSTM significantly reduce the PSD peak, with DQN–LSTM achieving the best effect.

Taken together, under this operating condition it can be concluded that when the road excitation becomes more severe, the DQN–LSTM active suspension system still outperforms the DQN active suspension system in terms of vertical body acceleration and wheel dynamic displacement. Owing to the introduction of the LSTM to capture the temporal characteristics of the road excitation, the optimisation effect further improves as the road grade increases.

Table 2: Comparison of ride comfort performance between the two active suspension control systems

Suspension type	Body vertical acceleration/ $m \cdot s^{-2}$	Improvement / %	Suspension dynamic deflection/ m	Improvement / %	Wheel dynamic displacement/ m	Improvement / %	Total band energy/ $m^2 / s^4 / Hz$	Attenuation/ %
Passive suspension	0.0892		0.0011		0.000218		0.0079	
DQN active suspension	0.0778	12.78	0.00136	-23.63	0.000216	0.91	0.0052	33.6
DQN–LSTM active suspension	0.0668	25.11	0.00128	-16.36	0.000192	11.92	0.0035	55.7

From Table 2, the following observations can be made:

(1) The moderate increase in suspension dynamic deflection reported in Table 2 is an inevitable consequence of the multi-objective performance trade-off inherent in active suspension systems, which must balance ride comfort, handling stability and actuator energy consumption. In order to significantly reduce body acceleration (comfort) and wheel dynamic displacement (road holding), the control algorithm needs to apply larger control forces, causing the suspension to operate more frequently within its working space. Under the same optimisation objective defined by an identical reward function, the DQN–LSTM algorithm exhibits a superior capability for intelligent trade-off: it achieves a higher performance gain (25.11% improvement in comfort vs 12.78% for DQN) at a smaller performance cost (suspension deflection deterioration of −16.36% vs −23.63% for DQN). This result demonstrates the advanced nature of the DQN–LSTM algorithm in addressing complex multi-objective optimisation problems.

(2) From the frequency-domain analysis, it can be observed that the total band energy is reduced by 33.6% under DQN control, indicating that the reinforcement learning-based method provides effective vibration suppression. In contrast, DQN–LSTM delivers even better vibration reduction across the entire frequency range, with the total energy reduced by 55.7%. This shows that the DQN–LSTM control algorithm, which is based on temporal feature modelling, can more effectively accommodate the dynamic characteristics of the suspension system and achieve significant attenuation of broadband vibrations.

In summary, under complex operating conditions, the DQN–LSTM active suspension system provides a notable improvement in overall vehicle ride comfort, further confirming the feasibility of the active suspension framework developed in this study. Therefore, under mixed road conditions, vehicles equipped with the DQN–LSTM active suspension system exhibit superior ride quality.

4. Conclusion

This study investigates an intelligent active suspension control strategy for vehicles based on deep reinforcement learning, and analyses the control performance of a DQN–LSTM algorithm on a two-degree-of-freedom suspension model. The optimisation effect of the control algorithm is evaluated using indices such as body acceleration, suspension dynamic deflection and tyre dynamic displacement. The main contributions are as follows:

(1) Deep reinforcement learning is applied to the vehicle active suspension control system, and a DQN–LSTM control algorithm incorporating a temporal memory mechanism is constructed. With the minimisation of body acceleration as its primary objective, the reinforcement learning agent realises active control of the vehicle suspension system under discretised control forces.

(2) Simulation results show that both DQN- and DQN–LSTM-based intelligent control algorithms can effectively reduce the RMS value of body acceleration under road excitations and thus improve ride comfort. Owing to the introduction of the LSTM network structure, the DQN–LSTM control algorithm possesses a stronger capability for time-series modelling and exhibits better control smoothness and convergence behaviour, particularly under long-duration excitations.

(3) As an initial feasibility study of the algorithm, the present work is conducted on a classical two-degree-of-freedom model. We fully acknowledge its limitations in capturing coupled vehicle dynamics such as pitch and roll. Therefore, the core focus of future work will immediately shift to co-simulation based on a seven-degree-of-freedom full-vehicle model using CarSim/Simulink, and a hardware-in-the-loop test procedure has already been planned to validate the effectiveness and robustness of the proposed strategy in more complex scenarios.

Acknowledgement

This work was supported by Major Projects of Shandong (2024CXGC010301)

References

[1] Yu M, Evangelou S A, Dini D. Advances in Active Suspension Systems for Road Vehicles [J]. Engineering, 2023.

[2] Llopis-Albert C, Rubio F, Zeng S. Multiobjective optimization framework for designing a vehicle suspension system. A comparison of optimization algorithms [J]. Advances in Engineering Software, 2023, 176: 103375.

[3] Zeng Q, Zhao J. Event-triggered controller design for active suspension systems: An adaptive backstepping method with error-dependent gain [J]. Control Engineering Practice, 2023, 136: 105547.

[4] Sun W, Li C Y, Wang J N, et al. Study on ride comfort of a hybrid suspension based on multi-condition modes [J]. Automotive Engineering, 2024, 46(11): 2076–2090+2099.

[5] Lin B, Su X. Fault-tolerant controller design for active suspension system with proportional differential sliding mode observer [J]. International Journal of Control, Automation and Systems, 2019, 17: 1751-1761

[6] Sun W, Li C Y, Wang J N, et al. Study on ride comfort of hybrid suspension for off-road vehicles [J]. Automotive Engineering, 2022, 44(01): 105–114+122.

[7] Cao Y L, Zhang Q. Steering control of four-wheel steering vehicles based on adaptive neuro-fuzzy inference [J]. Mechanical Design and Manufacture, 2021(03): 224–228+233.

[8] Xu M, Huang Q S. Research status of intelligent control methods for vehicle semi-active suspension [J]. Machine Tool & Hydraulics, 2021, 49(01): 169–174.

[9] Nazemian H, Masih-Tehrani M. Hybrid Fuzzy-PID control development for a truck air suspension system [J]. SAE International Journal of Commercial Vehicles, 2020, 13(1): 55-70.

[10] Ming L, Yibin L, Xuewen R, et al. Semi-active suspension control based on deep reinforcement learning [J]. IEEE Access, 2020, 8: 9978-9986.

[11] Wang Z H. Research on semi-active suspension strategy based on deep Q-network [D]. Chengdu: Southwest

Jiaotong University, 2021.
[12] Deng M, Sun D, Zhan L, et al. Advancing active suspension control with TD3-PSC: integrating physical safety constraints into deep reinforcement learning [J]. IEEE Access, 2024.
[13] Gu S Y, Jiang C H. Simulation study of model predictive control for vehicle active suspension based on RBF neural network [J]. Chinese Journal of Construction Machinery, 2025, 23(03): 410–414.
[14] Wei W Z, Xie Q Q, Sun J Z, et al. Research on a double-delay DDPG reinforcement learning control strategy for semi-active suspension [J]. Manufacturing Automation, 2025, 47(06): 85–92.
[15] Zhu X, Chen Z, Zhang S, et al. Intelligent Control Strategy of Vehicle Active Suspension Based on Deep Reinforcement Learning [C]//2022 China Automation Congress (CAC). IEEE, 2022: 4871-4876.
[16] Yu Y F. Research on control of four-corner interconnected air suspension based on multi-agent system [D]. Jinzhou: Liaoning University of Technology, 2023.
[17] Gao Z H, Bao M X, Gao F, et al. LSTM-based probabilistic multimodal anticipated trajectory prediction method [J]. Automotive Engineering, 2023, 45(07): 1145–1152+1162.
[18] Yadav H, Thakkar A. NOA-LSTM: An efficient LSTM cell architecture for time series forecasting [J]. Expert Systems with Applications, 2024, 238: 122333.