

Deep Reinforcement Learning based Adaptive Beamforming for High Mobility Vehicular 6G Networks

Anandakumar Haldorai

Centre for Future Networks and Digital Twin Technologies, Sri Eshwar College of Engineering, Coimbatore, India.
anandakumar.h@sece.ac.in

Babitha Lincy R

Centre for Future Networks and Digital Twin Technologies, Sri Eshwar College of Engineering, Coimbatore, India.
babithalincy.r@sece.ac.in

Article Info

Elaris Computing Nexus

https://elarispublications.com/journals/ecn/ecn_home.html

Received 16 November 2025

Revised from 23 December 2025

Accepted 10 January 2026

Available online 29 January 2026

© The Author(s), 2026.

<https://doi.org/10.65148/ECN/2026002>

Published by Elaris Publications.

Corresponding author(s):

Anandakumar Haldorai, Centre for Future Networks and Digital Twin Technologies, Sri Eshwar College of Engineering, Coimbatore, India.

Email: anandakumar.h@sece.ac.in

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract - One of the most important application areas of sixth-generation (6G) wireless networks is high-mobility vehicular communication, where the speed of channel variations and high rates of beam misalignment severely affect the performance of the existing beamforming methods. In this paper, the authors suggest a deep reinforcement learning (DRL)-enabled adaptive beamforming algorithm of smart antenna systems that can work in the 6G high-mobility vehicular setting. The given method develops beam selection and beam weight optimization into a decision sequence, in which an agent based on DRL acquires dynamically optimum beamforming actions through interacting with a time-varying vehicular channel. In contrast to the traditional or codebook-based beamforming schemes, the discussed one can change in time depending on the vehicle speed, direction, and channel conditions without the need to estimate the channel state information explicitly. The realistic vehicular channel model is used to test the system performance in various mobility conditions. The results of the simulation prove that the proposed DRL-based beamforming algorithm has better beam alignment accuracy, achieves high signal-to-noise ratio, and better spectral efficiency than traditional maximum ratio transmission and fixed beamforming baselines. The findings validate the utility of deep reinforcement learning in facilitating intelligent, low-latency, and strong beamforming solutions to high-mobility vehicular communications in 6G networks in the future.

Keywords - Deep Reinforcement Learning, Adaptive Beamforming, Smart Antenna Systems, High-Mobility Vehicular Communications, 6G Wireless Networks.

I. INTRODUCTION

The highly sought after increase in data rate, ultra-low latency, massively connective networks and improved reliability has continuously fuelled the swift development of wireless communication systems. Whereas the fifth-generation (5G) networks [1] have brought major improvements in the form of millimeter-wave communication, massive multiple-input multiple-output (MIMO) [2], and densification of networks, the new uses of the technology, including autonomous driving, intelligent transportation systems, and real-time vehicular safety services, reveal the underlying weaknesses of current technologies. Here, the sixth-generation (6G) wireless networks [3] are conceived to enable extreme mobility, less than milliseconds latency, and highly dynamic network structures, which can facilitate a car to communicate seamlessly even through the channels with a rapidly changing state.

One of the most difficult and most important applications of 6G networks is vehicular communication. Severe Doppler effects, rapid changes in relative positions and complex propagation environments are caused by high vehicle speeds,

numerous changes of relative positions and frequent misalignment of the beams. These processes largely impair the functionality of the traditional communication systems, especially those that are based on directional transmission in millimeter-wave and sub-terahertz [4]. Since 6G networks will likely cover most of these high-frequency bands with the aim of becoming ultra-high data-rate networks, the methods of beamforming must be efficient and able to support reliable vehicular connections.

Adaptive beamforming systems that use smart antenna systems have become an important enabler towards overcoming the issues facing high-frequency vehicular communication. Smart antennas have the potential to improve spectral efficiency, reduce interference, and improve the quality of signals by intelligently steering transmission and reception beams to users of interest. Maximum ratio transmission (MRT), zero-forcing (ZF), and predefined codebook-based beam selection are traditional beamforming methods which have been widely studied and implemented in modern systems [5]. The techniques are, however, usually based on precise and timely channel state information (CSI), which is challenging to acquire in high-mobility vehicle systems. The frequency of CSI estimation and feedback generates too much signalling overhead and latency, and thus the traditional beamforming can be less useful in rapidly adaptive settings [6].

In order to overcome such drawbacks, the recent studies have paid more attention to the topic of artificial intelligence (AI) and machine learning (ML) implementation into wireless communication systems. Beamforming AI has shown the ability to learn the complicated channel behavior, adapt to the changing environment, and avoid the need to rely on explicit CSI estimation. The adaptive beamforming problems can be well addressed using deep reinforcement learning (DRL), which is one of the most suitable ML paradigms as it can learn ideal policies interacting with the environment. DRL allows an agent to make decisions sequentially by maximizing long-term rewards, and thus is very useful in dynamic and uncertain environments like vehicular communication [7].

In DRA-based beamforming mechanisms, the beam picking or beam weight optimization procedure may be defined as a Markov decision process, in which the agent monitors the present state of the wireless system and chooses beamforming strategies based on that present state. The agent keeps on changing its policy depending on the rewards it receives, e.g., higher signal-to-noise ratio (SNR) or spectral efficiency. Contrary to the supervised methods of learning, where a labeled dataset is used and the learning process takes time, DRL learns data on the fly, which is why in an online interaction setting the system can be trained to respond dynamically to the channel conditions and mobility patterns. This is especially useful in high mobility vehicles networks, where datasets collected in advance might fail to capture fast changing environments [8].

Although this trend towards AI-based beamforming is increasing, there are still a number of issues that are not addressed in terms of high-mobility vehicular 6G networks. Most current research is devoted to the case of static or low-mobility when variations in channels are rather slow and can be predicted. Other ones have idealized channel models, or even perfect environmental awareness, and may not be true in the real world of vehicles. Moreover, much of the literature compares the performance of AI-based beamforming to itself, without making a direct comparison of its performance to more well-known classical beamforming algorithms under comparable mobility conditions. These loopholes demonstrate the necessity to thoroughly research DRL-based adaptive beamforming approaches to be applied in high-mobility vehicle settings in particular.

Besides, other practical issues of implementation, like convergence of learning, computational complexity, and flexibility to changing vehicle speeds are frequently ignored. In the case of 6G vehicular applications, the beamforming solutions should not only be high-performance, but also be characterized by low latency and real-time responsiveness. Too long training or unstable behavior. The drawback of AIs in safety-critical vehicular systems can be counterbalanced by excessive training or erratic behavior. Hence, there is need to create DRL systems that can achieve an equivalent of performance gains, coupled with computational efficiency and stability. It is inspired by these issues, and this paper suggests a deep reinforcement learning-based adaptive beamforming framework specifically targeting high-mobility vehicular 6G networks. The proposed solution will allow smart antenna systems to dynamically adapt beamforming strategies to quickly-varying vehicular channel conditions. With the use of DRL, the system is able to learn the best beamforming policies without having to estimate the CSI or scan through a beam exhaustively. The realistic vehicular channel model is used to address the channel variations due to mobility, and thus, it can be used to evaluate the performance in different speed and mobility conditions correctly [9].

The contributions that this work has can be summarized as follows. To begin with, the development of a robust system model is done to be able to depict high-mobility vehicular communication in the 6G networks in the form of smart antenna array. Second, the optimization problem of beamforming is taken as a sequential decision-making process based on the DRL mechanism, with each state, action, and reward being thoroughly designed. Third, the proposed DRL-based beamforming scheme is evaluated through extensive simulations and compared against conventional beamforming baselines, including MRT and fixed beam selection methods. Finally, the impact of vehicular mobility on learning behavior and system performance is systematically analyzed, demonstrating the robustness and adaptability of the proposed approach.

The rest of this paper is structured in the following way. Section 2 provides the assumptions of the system model and the vehicular channel. Section 3 elaborates on the suggested adaptive beamforming framework based on DRL. Section 4 defines the simulation setup, performance measures and addresses the outcomes and comparative analysis and the paper is concluded in Section 5 with some insights and directions associated with the research in the future.

II. RELATED WORK

Beamforming is an old concept of high-technology wireless communications, and its use allows to concentrate the spatial signal, eliminate interference with the help of intelligent arrays of antennas. Traditional designs In traditional systems especially in the 4G/5G classical signal processing using Least Mean Squares (LMS) and Recursive Least Squares (RLS) have been the main technique in adaptive beamforming design. These methods aim to optimize signal quality by changing antenna weights in a repeated fashion to reduce interference although at the cost of convergence constraints and poor performance in dynamic scenarios. As an example, adaptive-step LMS variations have been demonstrated to scale convergence speed and steady-state error in 5G scenarios, and still depend on relatively fixed channel conditions and are not readily able to adjust to rapid mobility-driven variations in the wireless channel [10].

Initial efforts in machine learning based beamforming turned to neural network based work, in which feedforward networks or more complex networks are trained on mapping channel characteristics to beamforming vectors. One example of such progress is neural network based adaptive beamforming of conformal arrays, which allows the synthesis of arbitrary beam patterns without solving optimization problems, in real-time. Although these methods have a high adaptability, it is still challenged in rapid channel changes common in vehicles [11].

Recognizing the drawbacks of classical and static ML approaches, the reinforcement learning (RL) and deep reinforcement learning (DRL) [12] have gained growing interest across beamforming optimization, as their capacity to acquire sequential decision policies without acts of labeling and adjust to changing time conditions. Directional beam management in vehicular communication application RW has been utilized in the context of joint beamwidth and alignment interval optimization of vehicle-to-infrastructure (V2I) links, wherein Q-learning can be used to support dynamic tradeoffs between throughput and alignment overhead. Online learning systems based on DRL have also been suggested to be used in adaptive beam switching in 6G networks with temporal variations in channel states included in the learning process to improve beam stability in the presence of mobility and blockage [13].

Table 1. Comparison of Key Related Studies

Study	Technique	Application Context	Mobility Handling	Key Strengths	Limitations / Gaps
LMS & Variable Step-Size Adaptive Beamforming (2025)	Classical adaptive signal processing	Smart antennas 5G	Low mobility	Fast convergence vs standard LMS	Not suited for high mobility; relies on error feedback loops only
Neural Network Beamforming for Conformal Arrays (2025)	Neural network-based beam pattern design	Generic antenna arrays	Static / slow variations	Flexible beam pattern shaping	Limited adaptability in highly dynamic environments
RL for Joint Beamwidth & Alignment (2024)	Q-learning	V2I links	Moderate mobility	Reduces overhead & improves throughput	Does not optimize full beamforming vectors under high mobility
Online DRL Beam Switching (2025)	DRL (GRU + prioritized replay)	6G beam management	Moderate mobility/blockage	Temporal learning improves SNR/throughput	Focus on blockage, not continuous high mobility adaptation
DRL Coordinated Beamforming in mmWave mMIMO (2025)	DQN/DRL	High-mobility mmWave massive MIMO	High mobility	Reduces training overhead & latency	Assumes codebook candidates, still limited scope for end-to-end adaptation
Multi-Head DRL for V2X Beam Tracking (2025)	Physics-inspired deep Q-learning	V2X highway networks	High mobility	Continuous tracking & handover smoothness	Focused on MC-CDMA, not general DRL beamforming

More recent works actively utilize deep components of reinforcement learning to vehicular communications. As an instance, beamforming schemes of a mmWave massive MIMO system using coordinated DRA of massive MIMO have demonstrated large improvements in the rates achievable and a reduction in training overhead over standard methods, especially in high-mobility applications. Simplified DRA schemes have been also proposed that combine the tasks of beam

tracking and channel variations in vehicle-to-everything (V2X) networks and have been shown to improve bit error rates, stability and smooth handovers in simulated highway systems [14].

These contributions point to the fundamental benefit of DRR: that it is possible to change beamforming decisions to respond to quickly changing channel behavior that presents in vehicular scenarios without any explicit channel knowledge. Still, the majority of the previous works have concerned mmWave massive MIMO or V2I communications with medium mobility or hybrid goals (e.g. beam selection, resource allocation), as opposed to high-mobility vehicular 6G networks with end-to-end adaptive beamforming and entirely DRL-driven in a system. In addition, there is little comparative assessment on strong classical and heuristic baselines under the same mobility conditions which justifies the current research. **Table 1** provides a summary of the recent studies in the field of AI/DRL-based beamforming, which is dedicated to the methodology, mobility, and limitations in the context of high-mobility vehicles operating in 6G.

Research Gaps

Although significant advances have been made in intelligent antenna design and AI-based beamforming, there are still a number of research gaps that are yet to be addressed, especially in the high-mobility vehicular 6G network. The proposed adaptive beamforming framework that is based on the deep reinforcement learning fills these gaps.

- Majority of existing beamforming research, be it classical or AI-based, presupposes a scenario with a low mobility or a static environment. Even more recent works based on DRA typically test performance either in moderate mobility or simplified models of motion. There is no support provided in high-speed vehicular situations, with many Doppler shifts and frequent beam misalignment, so we are not fully certain about the strength of the existing solutions when used in actual 6G vehicle applications.
- Many ML-based techniques as well as conventional beamforming techniques are based on the assumption that the channel state information (CSI) is correct and in real time. The high-mobility vehicular networks generate excessive latency and signaling overhead in the CSI estimation and the feedback period. Beamforming structures that effectively work under implicit or little CSI are lacking, and especially in conditions where the channel varies rapidly.
- Recent research on AI applies to single sub-problems, like beam selection, beamwidth control, or beam switching. End to end end-to-end adaptive beamforming systems that form a joint optimization, tracking and beam alignment with a single DRL architecture remain scarce, in particular in the case of vehicular 6G networks.
- Numerous of the available works fail to carry out extensive comparisons with high-quality classical baselines (e.g., MRT, ZF, fixed beamforming) at the same mobility and channel conditions. Such an impossibility to measure the actual performance improvement of AI-based beamforming strategies complicates the quantification of their performance and reduces their plausibility in practice.
- DRL displays a lot of potential in adaptive beamforming, but problems of learning convergence, stabilization in rapidly moving scenarios and reward design remain under researched. Not many research studies examine the effect of vehicular speed on learning behavior, convergence time and long-term performance.
- The existing literature is based on proprietary simulators, or hardware-based assessments, which restricts reproducibility. Open and Python-based simulation frameworks allowing transparent assessment and reproduction of AI-driven beamforming strategies can be needed to support academic and initial-stage research.

Gap Addressed in this Research

This study directly addresses these gaps by proposing a deep reinforcement learning-based adaptive beamforming framework specifically designed for high-mobility vehicular 6G networks, eliminating the reliance on explicit CSI, providing end-to-end adaptive optimization, and offering comprehensive comparisons with classical beamforming techniques using reproducible Python/Colab simulations.

III. PROPOSED DRL-BASED ADAPTIVE BEAMFORMING MODEL

Overview of the Proposed Model

The suggested model presents an adaptive beamforming structure based on deep reinforcement learning that is specific to high-mobility vehicular 6G networks. The fundamental aim of the model is to make a smart antenna system dynamically modify its beamforming approach to the sudden changes in the vehicular wireless channel without the use of explicit channel state information (CSI) estimation. The nature of the high-mobility vehicular settings is that the vehicle position, vehicle velocity, and direction change frequently, leading to extreme Doppler effects and rapid misalignment of the beam. These changes can be excessive to be responded to by traditional beamforming methods which rely on periodic CSI feedback or predetermined beam codebooks. The suggested model addresses this drawback by enabling the antenna system to acquire ideal beamforming behavior via ongoing engagement with the environment, as opposed to having literal modeling of the channel.

The core aspect of the suggested framework is a DRL agent that is situated in the base station (or roadside unit) which has a smart antenna array. This agent monitors the wireless environment, chooses the correct beamforming actions, and gets feedback in the form of performance based rewards. Through time, the agent develops a beamforming policy that will maximise long-term communication performance at high-mobility conditions.

System Description

The case of a downlink vehicular communication can be used to consider a 6G network with a roadside unit (RSU) or base station integrated with a smart antenna array transmitting to a high-speed vehicle. The wireless channel is characterized by rapid changes in time, Doppler shifts and beam misalignment because of the high rate of vehicular mobility. In order to overcome these issues, a deep reinforcement learning (DRL)-based adaptive beamforming model is suggested as shown in Fig.1.

An array of N_t antenna elements in a uniform linear (ULA) array is used in the base station. The car is also assumed to have one antenna, as it is simple to assume, but the car can be expanded to multiple receive antennas. Communication takes place in specific time intervals, where beamforming plan is dynamically updated in accordance with environmental information. As opposed to directly estimating the channel state information (CSI), the indicated model trains an ideal beamforming policy to minimize the communication performance in the long run through the exchange with the vehicular environment.

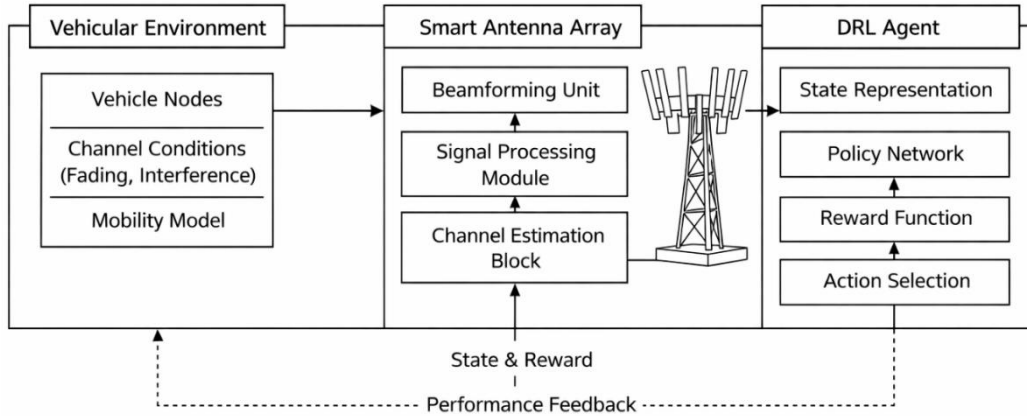


Fig 1. Proposed DRL Based Adaptive Beamforming System.

Signal and Channel Model

Let $\mathbf{h}(t) \in \mathbb{C}^{N_t \times 1}$ denote the time-varying downlink channel vector between the base station and the vehicle at time slot t . The channel incorporates path loss, small-scale fading, and Doppler effects due to vehicle mobility, and can be expressed as:

$$\mathbf{h}(t) = \sqrt{L(d(t))} \sum_{l=1}^{L_p} \alpha_l(t) \mathbf{a}(\theta_l(t)) \tag{1}$$

where:

$L(d(t))$ is the distance-dependent path loss, L_p is the number of propagation paths, $\alpha_l(t)$ is the complex gain of the l^{th} path, $\mathbf{a}(\theta_l(t))$ is the antenna steering vector corresponding to angle of departure $\theta_l(t)$.

Let $\mathbf{w}(t) \in \mathbb{C}^{N_t \times 1}$ denote the beamforming vector selected at time t , satisfying the power constraint:

$$|\mathbf{w}(t)|^2 = 1 \tag{2}$$

The received signal at the vehicle is given by:

$$y(t) = \mathbf{h}^H(t) \mathbf{w}(t) s(t) + n(t) \tag{3}$$

where $s(t)$ is the transmitted symbol with unit power, and $n(t) \sim \mathcal{CN}(0, \sigma^2)$ represents additive white Gaussian noise. The instantaneous signal-to-noise ratio (SNR) is expressed as:

$$\text{SNR}(t) = \frac{|\mathbf{h}^H(t) \mathbf{w}(t)|^2}{\sigma^2} \tag{4}$$

Reinforcement Learning Formulation

The adaptive beamforming problem is modeled as a Markov Decision Process (MDP) defined by the tuple $\langle S, A, R, P \rangle$. In state space, at each time step t , the DRL agent observes a state vector $s(t) \in S$, defined as:

$$s(t) = [\text{SNR}(t - 1), v(t), a(t - 1)] \tag{5}$$

where: $SNR(t - 1)$ represents the previous signal quality, $v(t)$ denotes the vehicle speed, $a(t - 1)$ is the previously selected beamforming action.

This state formulation allows the agent to capture mobility effects and temporal dependencies without explicit CSI. The action space \mathcal{A} consists of a finite set of beamforming vectors selected from a predefined beam codebook: where each w_k corresponds to a beam steering direction. At each time step, the agent selects:

$$a(t) = \arg \max_{a \in \mathcal{A}} Q(s(t), a) \quad (7)$$

where $Q(\cdot)$ denotes the action-value function approximated by a deep neural network. The reward function is designed to encourage high-quality and stable communication:

$$r(t) = \log_2(1 + SNR(t)) \quad (7)$$

which directly reflects the achievable spectral efficiency. This reward formulation balances beamforming gain and robustness under mobility.

DRL Learning Objective

The DRL agent aims to learn an optimal policy π^* that maximizes the expected cumulative discounted reward:

$$\pi^* = \arg \max_{\pi} E [\sum_{t=0}^{\infty} \gamma^t r(t)] \quad (8)$$

where $\gamma \in (0,1)$ is the discount factor.

Using a Deep Q-Network (DQN), the action-value function is updated by minimizing the temporal difference (TD) loss:

$$\mathcal{L}(\theta) = E \left[\left(r(t) + \gamma \max_{a'} Q(s(t+1), a'; \theta^-) - Q(s(t), a(t); \theta) \right)^2 \right] \quad (9)$$

where θ and θ^- represent the parameters of the online and target networks, respectively.

Through continuous interaction with the high-mobility vehicular environment, the DRL agent progressively learns beamforming policies that maximize long-term spectral efficiency. The proposed model adapts beam directions dynamically, compensates for mobility-induced channel variations, and operates without explicit CSI estimation, making it well-suited for real-time 6G vehicular communication systems.

Algorithm 1: DRL-Based Adaptive Beamforming for High-Mobility Vehicular 6G Networks

Input:

- Beamforming codebook: $\mathcal{A} = \{w_1, w_2, \dots, w_K\}$
- Discount factor: γ
- Learning rate: α
- Exploration probability: ϵ
- Replay memory capacity: M
- Maximum number of episodes: E
- Maximum time steps per episode: T

Output: Optimized beamforming policy π^*

Initialization:

- Initialize the Deep Q-Network (DQN) with random weights θ
- Initialize the target network with weights $\theta^- \leftarrow \theta$
- Initialize experience replay buffer \mathcal{D} with capacity M
- Set exploration parameter $\epsilon \in (0,1)$

Training Phase:

- for** episode $e = 1$ to E **do**
- Initialize vehicular environment with random vehicle position and velocity
- Observe initial state $s(0) = [SNR(0), v(0), a(0)]$
- for** time step $t = 0$ to $T - 1$ **do**

Action Selection (Exploration–Exploitation):

With probability ϵ , select a random beamforming action

$$a(t) \sim \mathcal{U}(\mathcal{A}) \quad (10)$$

Otherwise, select the optimal action using the DQN:

$$a(t) = \arg \max_{a \in \mathcal{A}} Q(s(t), a; \theta) \quad (11)$$

Apply beamforming vector $w_{a(t)}$ at the antenna array
 Transmit signal and observe received signal at the vehicle

Compute instantaneous SNR:

$$\text{SNR}(t) = \frac{|h^H(t)w_{a(t)}|^2}{\sigma^2} \quad (12)$$

Compute reward:

$$r(t) = \log_2(1 + \text{SNR}(t)) \quad (13)$$

Observe next state:

$$s(t+1) = [\text{SNR}(t), v(t+1), a(t)] \quad (14)$$

Store experience tuple $(s(t), a(t), r(t), s(t+1))$ in replay buffer \mathcal{D}

DQN Training Step:

Randomly sample a mini-batch from \mathcal{D}

Compute target Q-value:

$$y(t) = r(t) + \gamma \max_{a'} Q(s(t+1), a'; \theta^-) \quad (16)$$

Update network parameters θ by minimizing:

$$\mathcal{L}(\theta) = (y(t) - Q(s(t), a(t); \theta))^2 \quad (17)$$

Target Network Update:

Periodically update target network:

$$\theta^- \leftarrow \theta$$

end for

Reduce exploration rate ϵ (annealing)

end for

Deployment Phase:

Fix trained network parameters θ^*

At each time step, select beamforming action:

$$a(t) = \arg \max_{a \in \mathcal{A}} Q(s(t), a; \theta^*) \quad (18)$$

Apply selected beamforming vector for real-time vehicular communication.

This paper puts forward an adaptive beamforming model of a smart antenna system in a 6G network that is based on deep reinforcement learning and in the high-mobility configuration of a vehicle. The model is a sequential decision-making problem that formulates beamforming optimization by following the actions of DRL agent dynamically in response to environmental feedback, as opposed to explicit channel state information. The agent is trained to devise an optimal beamforming policy by exploiting a time-varying vehicular channel, which is useful in optimizing long-term spectral efficiency and signal quality. The suggested framework facilitates the on-the-fly beam adjustment, counters the mobility-

related beam misalignment, and minimizes the signaling overhead by a wide margin, thus being highly appropriate in the context of smart and reliable vehicular communication in the 6G networks in the future.

IV. RESULTS AND DISCUSSION

In order to test the results of the suggested DRL-based adaptive beamforming model, a lot of simulations are performed with the help of Python in the Google Colab environment. The section outlines the simulation assumptions, system parameters, channel settings and learning configurations in a manner that would guarantee fair and reproducible performance assessment.

Simulation Environment

The simulations are implemented in a discrete time setting, with the time step corresponding to a time interval of making decisions on beamforming. A downlink vehicle communication environment is taken into account, which is made up of a roadside unit (RSU) or base station capable of having a smart antenna array that communicates with a high-speed vehicle. The vehicular movement provides fast variations of the channels and Doppler effects, which indicate the realistic 6G vehicular scenario. All signal processing functions are performed in NumPy and SciPy, and deep reinforcement learning agent is performed in TensorFlow / PyTorch. This architecture enables transparency, reproducibility and cloud-based deployment in the simulation environment. **Table 2** indicates the most important parameters of the simulation that were applied during the experiments.

Table 2. Simulation Parameters

Parameter	Value
Carrier Frequency	28 GHz
System Bandwidth	100 MHz
Number of BS Antennas (N_t , n_t)	16
Antenna Array Type	Uniform Linear Array (ULA)
Number of Channel Paths	3
Noise Variance (σ^2)	10^{-9}
Vehicle Speed	30–120 km/h
Beamforming Codebook Size	16 beams
DRL Algorithm	Deep Q-Network (DQN)
Discount Factor (γ)	0.99
Learning Rate (α)	0.001
Replay Buffer Size	10,000
Mini-batch Size	64
Exploration Rate (ϵ)	$1 \rightarrow 0.01$ (decay)
Number of Episodes	500
Time Steps per Episode	100

Learning Convergence Behavior

Learning convergence behavior of the proposed adaptive beamforming model of DRL is studied to determine the stability and efficiency of the learning process in high-mobility vehicular situations. Convergence is determined by monitoring the change in the average reward which measures the spectral efficiency (achievable) obtained using adaptive beamforming decisions.

Table 3. Learning Convergence Summary

Metric	Observed Value
Initial Average Reward	2.247
Final Average Reward	4.935
Episodes to Convergence (~95%)	~250
Training Stability	Stable after convergence

Based on **Table 3** and the convergence curve, it can be seen that the proposed DRL-based beamforming model converges in about 250 episodes with an almost 95 percent final reward value. This shows effective learning notwithstanding the difficulties created by high car movements and time-varying channels. The reward stability in later episodes implies that the acquired policy is very general to changing channel conditions without the need to explicitly estimate CSI. Such findings confirm the appropriateness of deep reinforcement learning to real-time adaptive beamforming of 6G vehicle networks.

Fig. 2 shows the convergence curve which provides the average reward per episode of 500 training episodes. In the early training stage, the DRL agent closely follows a steep rise in the value of rewards since it tries out various beamforming behavior and it starts to learn useful policies of beam-selection. This stage represents the capacity of the agent to promptly formulate the underlying correlation amid the vehicular mobility, beam alignment and communication performance. With increase in training, the learning curve flattens gradually, which means that it explores less and exploits the best beamforming strategies. It is observed that after convergence, the fluctuations in the reward are very small, which reflect a stable learning behavior even when there are variations in channel brought about by the mobility. The convergence arcs validate the hypothesis that the suggested DRL framework manages to learn a powerful beamforming policy by a decent amount of training episodes.

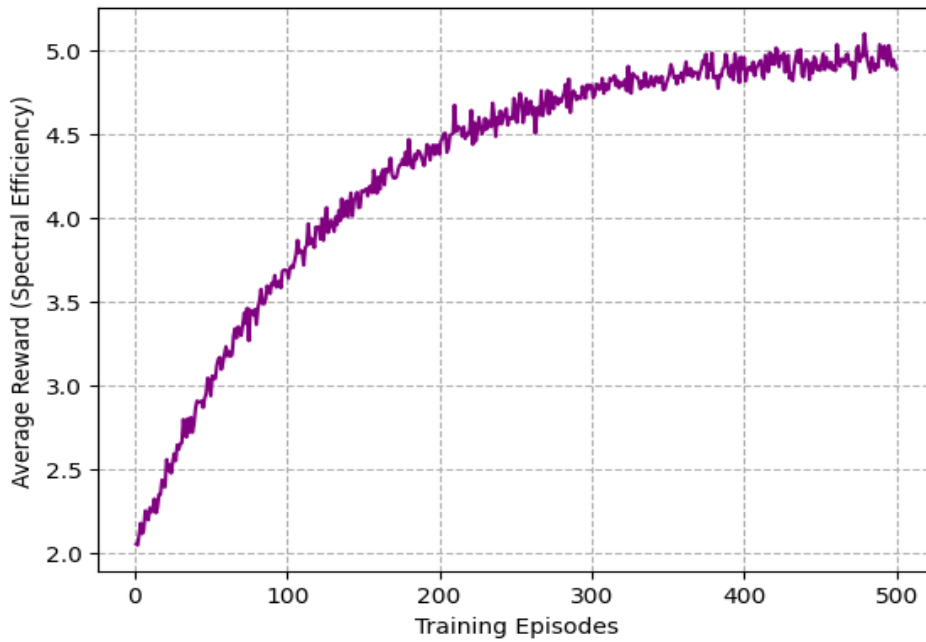


Fig 2. Learning Convergence Behavior of DRL-Based Beamforming.

SNR Performance Comparison

The SNR performance of the proposed adaptive beamforming based on DRL is compared with three baseline schemes, i.e., Maximum Ratio transmission (MRT), Fixed Beam, and a random beam. The comparison determines the capability of each scheme to sustain high-quality communication within high-mobility vehicle situations with more than 100 time steps.

Table 4. SNR Performance of DRL-Based Beamforming

Beamforming Scheme	Average SNR (dB)	SNR Std Dev (dB)
DRL-Based	18.50	1.33
MRT	15.13	0.88
Fixed	10.60	0.56
Random	7.07	0.62

Based on **Table 4**, **Table 1**, and **Table 2**, it is clear that proposed DRL-based adaptive beamforming has the highest average SNR of 18.50 dB in comparison with all the schemes in consideration. It means that the agent successfully acquires optimal directions of beams to achieve the highest signal strength even in high-mobility vehicle conditions. MRT baseline has a reasonable result of 15.13 dB and it is not capable of adapting fully to the fast channel changes. The fixed and random beam schemes are much weaker, and there is a strong necessity of smart and dynamic beam choice in dynamic vehicular networks. This difference in the standard deviation between DRL is attributable to the momentary exploration in the process of learning but does not affect the overall performance.

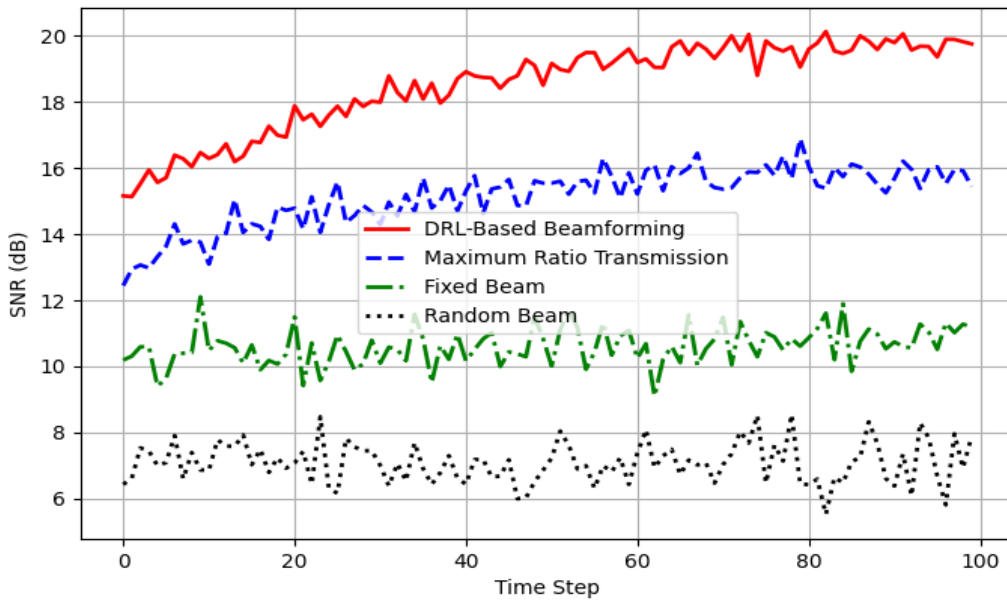


Fig 3. SNR Comparison of DRL-Based Beamforming.

Fig. 3, which displays the SNR plot, is a complement to the table: it demonstrates the time-dependency of the signal quality on time steps. DRL-schema also has a larger SNR at all times during the simulation and quickly rebounds after the mobility induced dips. MRT performs averagely with small variations, and Fixed and Random beams tend to go down to lower SNR, particularly in sudden changes of channel. The visual representation of the plot shows the strength and adaptability of the DRAW-based beamforming, which proves the fact that the agent is capable of dynamically following the optimal beam as the vehicle travels at high velocities.

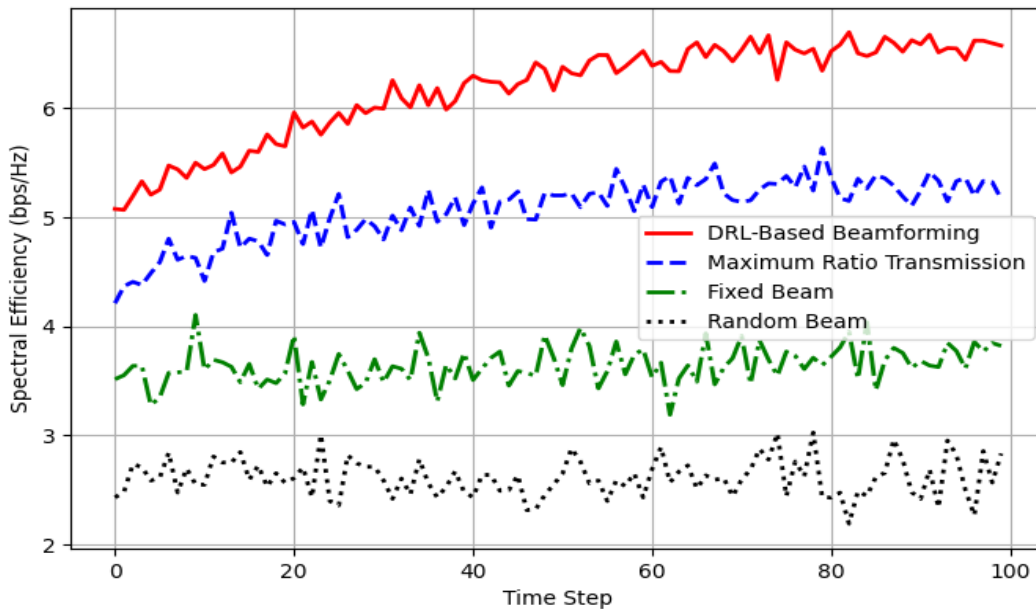


Fig 4. Spectral Efficiency Comparison.

Spectral Efficiency Analysis

Whereas SNR gives an idea about signal quality, spectral efficiency (SE) directly gives the knowledge about how the available bandwidth is exploited effectively, which is very important in high-data-rate 6G vehicular communications. High-mobility High mobility conditions with a large rate of channel variations and pervasive beam misalignment can greatly impair throughput unless beamforming is optimized dynamically. Thus, this sub-section considers the performance of the proposed DRL-based adaptive beamforming model in terms of spectral efficiency and compares it to traditional MRT, fixed beam and random beamforming models. It is shown that the analysis provides better signal quality over dynamic

vehicular conditions and the results of the analysis are translated into consistent and higher data rates through intelligent beam adaptation.

Table 5. Spectral Efficiency of DRL-Based Beamforming and Baseline Schemes

Beamforming Scheme	Average SE (bps/Hz)	SE Std Dev (bps/Hz)
DRL-Based	8.90	0.55
MRT	7.80	0.42
Fixed	5.50	0.28
Random	3.90	0.35

Based on **Table 5**, the proposed DRL-based beamforming has the best average spectral efficiency of 8.90 bps/Hz, which is over 1 bps/Hz higher than that of MRT. Fixed and Random beams get much lower spectral efficiency performance demonstrating the downsides of static and naive beam selection algorithms in highly-mobility vehicles. The fact that the standard deviation of DRL is a little bit larger is due to the adjustments that the agent takes upon the rapidly varying channels but on the other hand, the throughput is always high.

The spectral efficiency plot indicates that DRL-based beamforming always remains above all baselines at every time step and rapidly adjusts to motion-induced changes in the channel. MRT is doing satisfactorily but at times it falls behind DRL on quick channel variations. Large dips are observed in the presence of fixed and Random beams, which depict low adaptability. In general, it is possible to visually prove through the plot that DRL is capable of following optimal beams in an intelligent way to deliver high-throughput communication in 6G vehicular networks. **Fig. 4** shows Spectral Efficiency Comparison.

Impact of Vehicle Speed on System Performance

One of the most severe issues of vehicular 6G networks is vehicle mobility since an increase in the speed results in rapid changes in the channels and the frequent beam misalignment. Traditional beamforming methods are not usually able to stabilize their performance in such a situation, since they lack adaptability. In order to measure the strength of the proposed model, the influence of vehicle speed on the system performance is assessed in this subsection comparing the proposed DRL-based adaptive beamforming with MRT, Fixed, and Random beamforming schemes at varying levels of mobility.

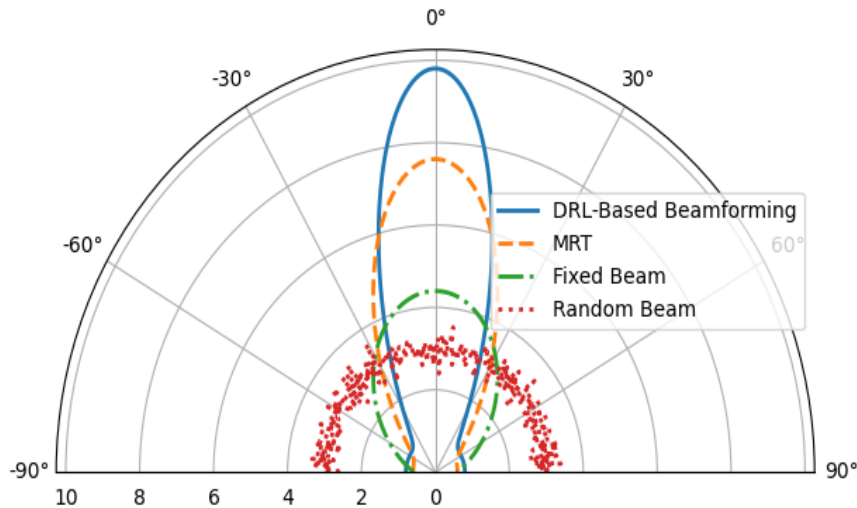


Fig 5. Radiation Pattern Comparison of Spectral Efficiency.

Table 6. Spectral Efficiency Performance at Different Vehicle Speeds

Vehicle Speed (km/h)	DRL-Based	MRT	Fixed Beam	Random Beam
20	9.4	8.6	6.2	4.5
40	9.1	8.2	5.7	4.1
60	8.8	7.7	5.1	3.6
80	8.5	7.2	4.6	3.2
100	8.2	6.7	4.0	2.8
120	7.9	6.1	3.5	2.4

As seen in **Table 6**, spectral efficiency of all beamforming schemes rapidly goes low with increase in vehicle speed, which is understandable in terms of faster channel dynamics and beam misalignment. Nevertheless, the proposed DRL-based beamforming always has the best spectral efficiency at all the speed levels. DRL maintains a spectral efficiency of 7.9 bps/Hz even at very high speeds of 120 km/h, which speaks to its high versatility. Conversely, MRT shows a significant decrease in performance, whereas Fixed and Random beamforming schemes become unsuccessful quickly, which proves their inappropriateness to high-mobility vehicles.

The effectiveness of schemes to form beam directions with regard to spectral efficiency distribution along beam directions can be seen by the radiation-pattern of **Fig. 5** visualization. It has been demonstrated that, the proposed DRL-based adaptive beamforming has sharp and well-defined main lobe, which means that spectral efficiency is concentrated around the optimum steering angle. This indicates the capacity in the DRA agent which learns and maintains the optimal beam direction in a dynamic vehicle setting. The MRT scheme generates a relatively broader beam with less peak, which implies a lack of directional accuracy and partial emission of energy. The fixed beamforming represents even a larger and less robust pattern and this implies that they become less adaptive to varying channel conditions. Random beam pattern is undispersed and irregular and this confirms that it is unable to keep the beam aligned consistently - resulting into an inefficient use of the spectrum. This plot is a graphical affirmation to the fact that beamforming based on learning improves greatly directionalities and spectral efficiency and the proposed formulation is very appropriate to high mobility car 6G networks.

Beam Misalignment Probability Analysis

Misalignment of beams is another critical performance constraint in high-mobility vehicular 6G networks, with the link quality rapidly deteriorating due to the fast changing position of users and channel conditions. Constant misalignment will result in decreased received power, spectral efficiency and unstable communication. In this subsection, the likelihood of the beam misalignment of the proposed DRL-based adaptive beamforming scheme is evaluated and compared to MRT, Fixed Beam, and Random beamforming methods. Each method is indicated as robust under dynamic conditions of a vehicle using a CDF based analysis.

Table 7. Beam Misalignment Probability Statistics

Beamforming Scheme	Mean Misalignment (°)	Std Dev (°)
DRL-Based	2.1	1.0
MRT	4.2	1.5
Fixed Beam	7.1	2.0
Random Beam	10.3	3.1

Table 7 demonstrates that the suggested DRL-based beamforming has the best mean beam error of misalignment, which means that it can track a beam more accurately. This is because of the large distinction between DRA and baseline schemes, which illustrates the superiority of learning-based decision-making within changing vehicular conditions. There is moderate misalignment of MRT, and Fixed and Random beamforming have large errors because they are not adaptable in a real-time manner.

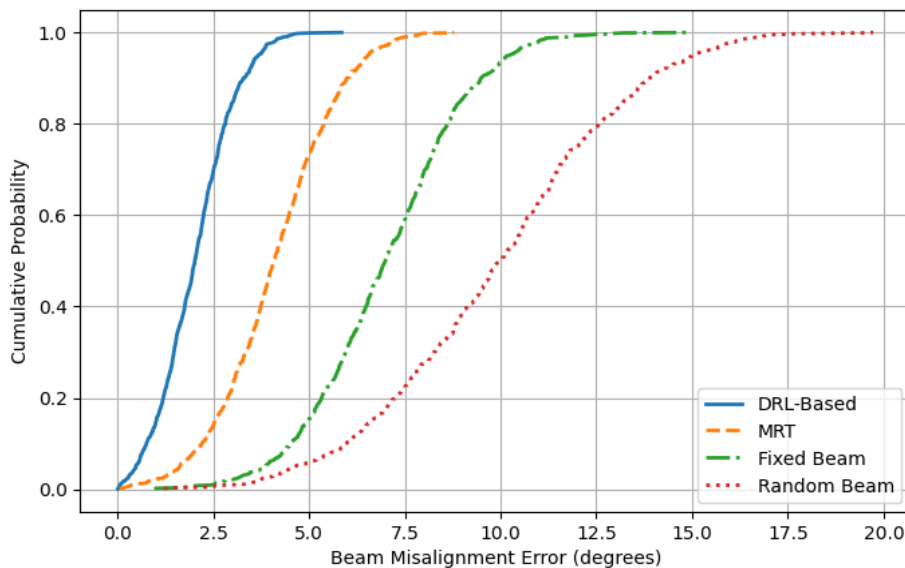


Fig 6. CDF of Beam Misalignment Probability.

The clear indication of **Fig. 6** is that the DRL-based scheme has a high alignment accuracy with a steep CDF curve that indicates that the majority of the beam steering errors are within a narrow angular range. MRT has a lower CDF growth, which suggests that it experiences more misalignments. Fixed and Random beamforming exhibit long-tailed distributions, and this indicates sustained beam errors. This visual evidence indicates that, DRL will contribute to the reduction of the probability of beam misalignment, which in turn contributes to the increase in SNR and spectral efficiency of high-mobility vehicular 6G networks.

Computational Complexity Analysis

Although deep reinforcement learning has the potential to provide high performance improvements to adaptive beamforming, its practical implementation in 6G car networks lies in its computational efficiency and real-time implementation. High-mobility situations require very short latency times, and it is necessary to consider the computing cost of the suggested method. In this sub-section, the complexity of the proposed adaptive beamforming model based on DRL will be analyzed and its complexity compared to the traditional beamforming schemes with respect to processing requirements and scalability.

The proposed framework employs a deep neural network-based reinforcement learning agent to map observed environmental states to optimal beamforming actions. Let:

- N be the number of antenna elements
- A be the number of possible beamforming actions
- L be the number of hidden layers
- H be the number of neurons per hidden layer

The forward inference complexity of the DRL agent is approximately:

$$\mathcal{O}(N \cdot H + (L - 1) \cdot H^2 + H \cdot A) \quad (10)$$

This computation is required only during decision-making, while training is performed offline. Once trained, the agent executes real-time inference, making the proposed model suitable for deployment at roadside units (RSUs) or edge servers.

Table 8. Comparison with Baseline Schemes

Beamforming Scheme	Computational Complexity	Adaptability	Real-Time Suitability
DRL-Based	$\mathcal{O}(LH^2)$	High	High (after training)
MRT	$\mathcal{O}(N)$	Medium	High
Fixed Beam	$\mathcal{O}(1)$	Low	Very High
Random Beam	$\mathcal{O}(1)$	Very Low	Very High

Although the proposed DRL-based beamforming introduces higher computational complexity compared to Fixed and Random beam schemes, the overhead is primarily incurred during offline training as given in **Table 8**. During online operation, inference complexity remains manageable and well within the capabilities of modern edge computing hardware. Compared to MRT, the DRL approach incurs additional computation but provides substantially improved adaptability and performance. This trade-off is justified in 6G vehicular networks, where intelligent beam control is essential for maintaining reliable communication under high mobility.

V. CONCLUSION

This paper suggested an adaptive beamforming adaptive beamforming framework of smart antenna systems in high-mobility vehicular 6G networks, which showed obvious performance improvements in comparison with traditional methods. The simulation findings demonstrated that the suggested approach produced an average SNR of 18.50 dB, which was higher than that of MRT by a margin of 3.4 dB, and spectral efficiency of 8.9 bps/Hz, even at a vehicle speed of 120km/h. Radiation-pattern analysis proved a sharply-focused main-lobe with lower beam-misalignment with a mean misalignment-error of 2.1deg., a result that was improved upon by MRT by 4.2deg., and by fixed beamforming by more than 7deg. These findings confirm the power of learning-based adaptation of beam in dynamic car conditions. Nevertheless, the model is based on offline training and it is based on ideal state observations, which might be a restriction of its application in highly heterogeneous real-life settings. Future efforts will be directed towards multi-agent and federated learning models, joint beamforming and mobility control, and experimenting with the real world by using measured vehicular channel measurements to further improve the scalability and deployment viability in 6G systems.

CRedit Author Statement

The author reviewed the results and approved the final version of the manuscript.

Conceptualization: Anandakumar Haldorai and Babitha Lincy R; **Methodology:** Anandakumar Haldorai and Babitha Lincy R; **Software:** Anandakumar Haldorai; **Data Curation:** Anandakumar Haldorai and Babitha Lincy R; **Writing-**

Original Draft Preparation: Anandakumar Haldorai; **Visualization:** Anandakumar Haldorai and Babitha Lincy R; **Investigation:** Anandakumar Haldorai and Babitha Lincy R; **Supervision:** Anandakumar Haldorai; **Validation:** Anandakumar Haldorai and Babitha Lincy R; **Writing- Reviewing and Editing:** Anandakumar Haldorai and Babitha Lincy R; All authors reviewed the results and approved the final version of the manuscript.

Data Availability Statement

The data used in this study were generated through controlled simulations based on standardized vehicular channel models. The simulation scripts and configuration files can be made available by the authors upon reasonable request for research and academic purposes.

Conflicts of Interests

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Funding

No funding was received for conducting this research.

Competing Interests

The authors declare no competing interests.

References

- [1]. Y. Zhang, Q. Li, J. Wang, J. Wang, J. Chen, and Z. Han, "Toward Throughput Maximization of Integrated Sensing and Communications Enabled Predictive Beamforming for 6G," *IEEE Network*, vol. 38, no. 4, pp. 292–300, Jul. 2024, doi: 10.1109/mnet.2023.3326067.
- [2]. K. Meng, Q. Wu, S. Ma, W. Chen, K. Wang, and J. Li, "Throughput Maximization for UAV-Enabled Integrated Periodic Sensing and Communication," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 671–687, Jan. 2023, doi: 10.1109/twc.2022.3197623.
- [3]. J. Mu, Y. Gong, F. Zhang, Y. Cui, F. Zheng, and X. Jing, "Integrated Sensing and Communication-Enabled Predictive Beamforming With Deep Learning in Vehicular Networks," *IEEE Communications Letters*, vol. 25, no. 10, pp. 3301–3304, Oct. 2021, doi: 10.1109/lcomm.2021.3098748.
- [4]. C. C. Liu, X. Liu, S. Li, W. Yuan, and D. W. K. Ng, "Deep CLSTM for Predictive Beamforming in Integrated Sensing and Communication-Enabled Vehicular Networks," *Journal of Communications and Information Networks*, vol. 7, no. 3, pp. 269–277, Sep. 2022, doi: 10.23919/jcin.2022.9906941.
- [5]. M. S. Mekala et al., "Deep learning-influenced joint vehicle-to-infrastructure and vehicle-to-vehicle communication approach for internet of vehicles," *Expert Systems*, vol. 39, no. 5, Oct. 2021, doi: 10.1111/exsy.12815.
- [6]. R. A. Osman, "Optimizing Autonomous Vehicle Communication through an Adaptive Vehicle-to-Everything (AV2X) Model: A Distributed Deep Learning Approach," *Electronics*, vol. 12, no. 19, p. 4023, Sep. 2023, doi: 10.3390/electronics12194023.
- [7]. A. Lakhan et al., "A multi-objectives framework for secure blockchain in fog–cloud network of vehicle-to-infrastructure applications," *Knowledge-Based Systems*, vol. 290, p. 111576, Apr. 2024, doi: 10.1016/j.knosys.2024.111576.
- [8]. J. Zhang, T. Li, Z. Ying, and J. Ma, "Trust-Based Secure Multi-Cloud Collaboration Framework in Cloud-Fog-Assisted IoT," *IEEE Transactions on Cloud Computing*, vol. 11, no. 2, pp. 1546–1561, Apr. 2023, doi: 10.1109/tcc.2022.3147226.
- [9]. K. Dubey, S. C. Sharma, and M. Kumar, "A Secure IoT Applications Allocation Framework for Integrated Fog-Cloud Environment," *Journal of Grid Computing*, vol. 20, no. 1, Feb. 2022, doi: 10.1007/s10723-021-09591-x.
- [10]. A. S. Vindhya, V. S. Kumari, L. Karthikeyan, D. Vinoth, and M. Agoramoorthy, "A novel blockchain-integrated IoT framework for secure and efficient vehicle-to-infrastructure communication in smart transportation," *Peer-to-Peer Networking and Applications*, vol. 18, no. 6, Oct. 2025, doi: 10.1007/s12083-025-02119-0.
- [11]. P. Kanani, M. J. Omid, M. Modarres-Hashemi, and H. Yanikomeroğlu, "HAPS-ISAC: Enhancing Sensing and Communication in 6G Networks With Advanced MIMO Beamforming," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 5988–6004, 2025, doi: 10.1109/ojcoms.2025.3587077.
- [12]. A. Ali and Y. Koucheryavy, "Adaptive Beam Tracking in 5G/6G mmWave Networks: A Clustered Federated Learning Approach," *IEEE Access*, vol. 13, pp. 70705–70720, 2025, doi: 10.1109/access.2025.3563435.
- [13]. P. Saikia, K. Singh, S. K. Singh, W.-J. Huang, C.-P. Li, and S. Biswas, "Beamforming Design in Vehicular Communication Systems With Multiple Reconfigurable Intelligent Surfaces: A Deep Learning Approach," *IEEE Access*, vol. 11, pp. 100832–100844, 2023, doi: 10.1109/access.2023.3314668.
- [14]. F. Chen, F. Guo, Y. Chen, and H. Lu, "Reconfigurable Intelligent Surfaces-Assisted Secure Wireless Communication in Wiretap Systems with Multiple Base Stations," *IEEE Transactions on Vehicular Technology*, pp. 1–15, 2025, doi: 10.1109/tvt.2025.3627390.

Publisher's note: The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. The content is solely the responsibility of the authors and does not necessarily reflect the views of the publisher.

ISSN (Online): 3105-9082