

Article

SDN-Assisted Deep Q-Learning Framework for Adaptive Mobility and Handover Optimization in Hybrid 5G Networks

Yahya S. Junejo ^{1,2} , Faisal K. Shaikh ² , Bhawani S. Chowdhry ²  and Waleed Ejaz ^{1,*} ¹ Department of Electrical and Computer Engineering, Lakehead University Barrie STEM Hub, Barrie, ON L4N 1R7, Canada² Department of Telecommunication Engineering, Mehran University of Engineering and Technology, Jamshoro 76062, Sindh, Pakistan; faisal.shaikh@faculty.muett.edu.pk (F.K.S.)

* Correspondence: wejaz@lakeheadu.ca

Abstract

In the evolving landscape of next-generation wireless networks, ensuring seamless mobility and high-quality service delivery for millions of devices and end users in dynamic scenarios, where the speed of a wireless device keeps changing with time, is important. The mobility, seamless and continuous connectivity, and ultra-dense deployment of wireless networks pose a significant challenge. Seamless and successful transition of a wireless device from point A to point B in variable-speed scenarios is one of the major challenges in future networks. This paper presents a novel Deep Q-Network (DQN)-based reinforcement learning (RL) framework integrated with Software-Defined Networking (SDN) for intelligent mobility management in hybrid 5G cellular networks consisting of macro and small base stations. The proposed system architecture utilizes a SDN controller to receive real-time user measurement reports, including Reference Signal Received Power (RSRP), Signal-to-Interference Noise Ratio (SINR), and user velocity, thereby classifying user mobility into distinct subclasses and dynamically determining optimal handover parameters. Leveraging the DQN's capability to learn adaptive strategies, the model enables seamless transitions between macro and small cells based on mobility profiles, thereby enhancing Quality of Service (QoS) metrics such as latency, throughput, and handover efficiency. Simulation results demonstrate consistent performance improvements over baseline and existing models in ultra-dense network environments, with handover success rates 10–15% higher across SINR and different speed scenarios, while maintaining a packet failure rate of 9% across different speed scenarios, allowing more users to transition during various environmental changes seamlessly. Our proposed model is compared with our previous work and Learning-based Intelligent Mobility Management (LIM2) models. Specifically, our previous work focused on adaptive handover management primarily for high-speed train scenarios using a learning-assisted approach tailored to fixed high-mobility scenarios, with a limitation to single mobility conditions. This work contributes to the field of merging SDN's centralized control with the predictive power of RL, paving the way for more resilient and responsive mobile networks in high-mobility scenarios. The proposed approach incorporates subclass-based mobility action abstraction, joint optimization of TTT and hysteresis margin, and dynamic target cell selection using global network information available at the SDN controller.



Academic Editor: Abderezak Touzene

Received: 7 January 2026

Revised: 2 March 2026

Accepted: 25 March 2026

Published: 2 May 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.**Keywords:** Next Generation Networks (Nex-Gen); handover; hybrid mobility; high mobility; Software-Defined Networks (SDN); machine learning (ML); seamless mobility

1. Introduction

The rapid expansion of mobile broadband services, intelligent connected devices, and the Internet of Things (IoT) ecosystem has dramatically increased global mobile data traffic. By 2029, global usage is expected to more than triple, driven by data-intensive applications such as high-definition video, cloud services, and immersive communication technologies, including Virtual and Augmented reality (AR/VR) [1]. The deployment of 5G networks has further intensified these demands by supporting applications that require stringent Quality of Service (QoS) and ultra-low latency, necessitating more efficient and scalable mobility management frameworks [2].

To meet these growing performance requirements, network densification via Ultra-Dense Small Cell (UDSC) Heterogeneous Networks (HetNets) has emerged as a key architectural strategy. By combining macro base stations operating in 6-GHz bands with dense deployments of small cells using mmWave frequencies, operators can simultaneously address coverage and capacity challenges [2]. However, this hybrid architecture introduces significant complexity in mobility management. The shorter propagation range and high sensitivity of mmWave signals lead to frequent handovers (HOs), while user mobility patterns, often spanning low-speed pedestrian movement to high-speed vehicular and train scenarios, further increase the risk of unstable connections, unnecessary handovers (UHOs), and Radio Link Failure (RLF) [3].

Effective handover performance relies heavily on optimizing Handover Control Parameters (HCPs), particularly the Time-to-Trigger (TTT) and Hysteresis Margin (HM). Static or heuristically configured HCPs are unable to adapt rapidly varying network conditions such as interference, fluctuating SINR, cell load imbalance, or changes in user velocity. As modern applications demand uninterrupted service, data-driven and adaptive HO optimization mechanisms have attracted substantial interest. Recent research [4] explores Machine Learning (ML)-based HO optimization due to its potential to reduce RLFs and improve mobility robustness. However, many ML frameworks rely on extensive offline data, and many struggle to generalize under dynamic mobility conditions.

These challenges are magnified in vehicular and transportation environments, where applications such as real-time navigation, road-safety alerts, multimedia streaming, and remote diagnostics require reliable, low-latency connectivity. As the number of vehicles continues to rise, user expectations for stable mobile service increase correspondingly [5]. Despite advancements in cellular technologies, maintaining seamless mobility for high-speed and variable-speed users remains a pressing challenge. Software-Defined Networks (SDNs) offer a promising avenue to address this limitation. By separating the control and data planes, SDN enables centralized decision-making, global visibility, and improved coordination among heterogeneous base stations, thereby enhancing mobility management, load balancing, and network automation [6].

Building on our previous work in [7], which demonstrated the benefits of adaptive mobility management for high-speed train scenarios, the previous model did not generalize across variable-speed mobility conditions. In this paper, we extend that framework by proposing a SDN controller-assisted mobility management system integrated with a Deep Q-Network (DQN) reinforcement learning algorithm. The hybrid network environment, comprising macro cells for wide-area coverage and mmWave for small cells for high-capacity services, enables the controller to select the optimal target cell based on user speed, radio conditions, and service requirements. The SDN controller maintains a global network view, whereas the DQN optimally selects the combination of TTT and HM values to minimize UHOs and RLFs. Both parameters are applied using the A5 handover triggering event, ensuring timely and robust HO decisions under diverse mobility patterns. It is also important to clarify that the proposed framework does not aim to redefine the cellular

mobility architecture; instead, it focuses on optimizing handover parameter selection using reinforcement learning under variable mobility conditions. The novelty lies in structuring the learning action space using subclasses that align with the handover control parameters guiding a UE to be handed over to the target cell. This structured design improves learning stability and policy interpretability compared to flat action representation. Traditional reinforcement learning formulations treat handover parameters as independent or flat actions, which increases the action dimensionality and slows convergence. In contrast, our subclass-based abstraction groups mobility-aware parameter combinations into discrete policy classes. This reduces exploration noise [8,9], enforcing 3GPP-consistent parameter relationships [10], and stabilizes learning under fast varying channel conditions [11,12]. Empirically, without a subclass structure, it may lead to a ping-pong effect, thereby degrading overall QoS. This model helps in improving convergence stability and sample efficiency while preserving the interpretability of learned decisions.

The following are the main contributions of this paper:

- We propose a handover mechanism for trains in high- and low-mobility scenarios, by using mmWave and sub-6 GHz bands for small and macro base stations, respectively, in 5G.
- We deploy an SDN controller-based 5G hybrid cellular network with small and macro base stations operating in mmWave frequency bands for optimal target Next Generation Node B (gNB) selection.
- We use a Deep Q-Network RL algorithm to select the best available target cell by introducing subclasses as actions in response to the current state and by choosing the optimal values of the TTT and hysteresis margin (HM).
- We develop an adaptive handover management algorithm to comprehensively evaluate the model's performance, focusing on key metrics including handover success rate, packet success rate, latency, and user throughput.
- Finally, we present an analysis and compare the simulation results with those from our previous work, and from the LIM2, DDQL, and PPO algorithms.

The rest of the paper is organized as follows: Section 2 provides a literature review of recent work on SDN- and machine-learning-based handovers. Section 3 and 4 presents a system model for the hybrid 5G cellular network for improved QoS and solution approach. Simulation results, challenges, and future directions are discussed in Sections 5 and 6.

2. Related Work

The increasing density of cellular deployments, the diversity of mobility patterns, and the stringent QoS requirements of emerging applications have motivated extensive research into intelligent mobility management for 5G and next-generation networks. Traditional handover schemes, although effective in earlier cellular generations, struggle to adapt to rapidly changing radio conditions, especially in ultra-dense heterogeneous networks and different mobility environments. Consequently, recent work has focused on SDN-assisted mobility frameworks and machine-learning-driven handover optimization, because it provides a logically centralized and programmable control plane that traditional distributed 3GPP architectures lack [13]. Secondly, this global, cross-layer visibility enables coherent, network-wide mobility decisions instead of local, signal-strength-only triggering at individual cells [14]. This section reviews key advancements that support the integration of the SDN-DQN model proposed in this research study.

Early efforts to integrate reinforcement learning into HO optimization include a deep deterministic policy gradient (DDPG)-based adaptive handover method presented in [6], where vehicles travel at 10–60 km/h based on RSRP/SINR measurements. While this model improves the handover success rate (HOSR), it optimizes only the hysteresis margin

(HM) and does not incorporate TTT, limiting scalability and robustness in more dynamic 5G environments. Our previous work in [7] addressed high-speed train scenarios with an E-SARSA-based adaptive HO mechanism. However, the design remains tailored to fixed high-mobility conditions and does not generalize to variable-speed environments.

More complex policy-learning techniques have also been explored in [4]. The PPO-based handover controller performs HO and target-cell decisions directly at the base station through an A3-triggered HO mechanism. Although the method supports mobility of up to 90 km/h, it exhibits frequent HOs and high deployment complexity due to its distributed learning architecture. Similarly, the hybrid SDN-ML architecture surveyed in [15] incorporates network slicing for high-speed trains and uses an A3-based hysteresis margin and TTT prediction. However, it relies solely on RSRP-based decision-making and incurs high signaling overhead due to SDN's centralized control.

Comprehensive surveys, such as [16], highlight the growing use of Q-learning and deep RL approaches for HO optimization, emphasizing the role of parameters such as TTT and HM under A3-triggering event conditions. These studies primarily focus on mitigating ping-pong effects and reducing RLFs; however, they also note that most existing work remains simulation-driven and has limited real-world validation. Further studies in [17,18] argue for context-aware, multi-objective HO schemes that incorporate load, QoS, and sustainability constraints beyond conventional A3-based mechanisms. Although SDN can coordinate these multi-parameter decisions, challenges remain with controller overhead, robustness, and scalability in ultra-dense deployments.

SDN-based mobility management continues to gain interest due to its global network visibility and centralized HO decision-making. The SDN-enabled HetNet architecture proposed in [13] formulates the HO decision as a linear optimization problem that integrates macro, micro, and femto cells. Despite its multi-criteria decision-making capability, the approach struggles to accurately estimate dwell time and remains unstable under rapidly changing network conditions. Additional SDN-IoT mobility frameworks, such as [19], demonstrate benefits in traffic management and flow-rule optimization but highlight persistent issues related to signaling overhead, interoperability, and lack of standardization.

Machine-learning-driven prediction models have also been explored to enhance mobility robustness. The slice-aware LSTM predictor in [20] forecasts per-slice resource demand and assists mobility along a predicted trajectory. However, inaccuracies in traffic prediction can lead to misallocation and do not directly trigger radio-layer HO events. Similarly, the Kalman filter-Q-learning combination in [21] relies on offset-based A3 HO triggering and performs well only in high-mobility environments, while suffering from inaccuracies in neighbor-cell prediction and increased ping-pong rates. Work such as [22] introduces semi-Markov and supervised ML approaches for active HO preparation using A2/A3 events. Still, these frameworks rely heavily on accurate predictions and incur additional control-plane overhead, particularly in mmWave deployments.

More recent DL-based approaches investigate fairness and proactive connectivity. The model in [23] uses deep learning to optimize both throughput and fairness by dynamically tuning TTT/HM under A3-triggered events; however, the centralized controller increases signaling overhead and suffers from high training complexity and limited generalization. A DQN-assisted pre-connect mechanism proposed in [24] used a four-layer DNN, an experience relay, and an ϵ -greedy policy to maximize proactive HO success for mobility up to 500 km/h. While effective for specific scenarios, it lacks robustness and scalability across diverse mobility conditions.

SDN-driven mobility management for high-speed trains has also been extensively studied. The approach in [25] employs Kalman-filter-based trajectory estimation to pre-install flows in the next predicted gNB. Although it reduces service disruptions, it relies heavily on

accurate telemetry and is limited by single-controller dependency and assumptions specific to certain topologies. A distributed mobility management (DMM) approach is described in [26], in which SDN dynamically selects access routers to achieve load balancing, whereas a SDN-based dynamic scheme in [25] leverages speed, location, and trajectory information for seamless HO. These solutions mitigate scalability and tunnel overhead issues but rely predominantly on RSS-based triggers, leading to frequent HOs in dense deployments.

Deep learning methods for HO preparation in [27] predict link blockage to maintain reliable connectivity, yet they do not address multi-metric HO criteria or variable-speed scenarios. Overall, the existing literature demonstrates significant progress toward intelligent mobility management but also reveals persistent challenges in robustness, scalability, multi-objective optimization, and performance under heterogeneous mobility speeds.

Table 1 summarizes the related work detailed above. Given these limitations, it is evident that a unified mobility management framework capable of handling degradations across high-, moderate-, and low mobility scenarios remains lacking. This motivates the SDN-assisted, DQN-driven hybrid mobility model proposed in the next section, which integrates macro and mmWave small cells to enhance adaptability and HO robustness.

Table 1. Related work on handover optimization techniques in Nex-Gen wireless networks.

Ref.	Year	Event A5	Use of SDN	ML Algorithms	Remarks
[3]	2025	✗	✗	MLP	This research uses the A3 handover triggering event. It has implementation complexity because it does not use real-time data, is data-dependent, and introduces high RLF. It considers different mobility scenarios, but the deployment scenario is limited to 7 base stations.
[6]	2025	✗	✗	DDPG	It does not consider the multi-frequency connectivity, only considers the A3 triggering event, is limited to only slow mobility up to 60 km/h with deployment scenarios of up to a few base stations, and does not consider the densification of network deployment. It also has a high computational cost.
[7]	2025	✓	✗	U.K.F and E-SARSA	It is only usable in high mobility scenarios with a speed above 150 km/h. It considers only a single frequency range and does not account for low- and moderate-speed vehicles. It can only be used in high-speed train scenarios.
[4]	2025	✓	✗	PPO	It uses the A3 triggering event and has an adaptability issue, is not robust to optimize handover parameters based on environmental changes, and is data-dependent, results in frequent HOs, and has high deployment complexity.
[15]	2025	✗	✓	✗	Has high complexity by using network slicing with SDN. It relies on the A3 handover event and entails a processing delay risk. It does not explicitly consider optimization; instead, it focuses on network architecture. It relies on accurate prediction and incurs a high cost in real-time deployment.
[16]	2025	✗	✗	RL	It is a survey paper that discusses the use of the A3 triggering event. It also mentions different machine learning algorithms in general. It considers only high mobility scenarios. It emphasizes hybrid scenarios, prediction scenarios, and user profiling.
[17]	2025	✓	✓	PPO	This is a survey paper that mentions the measurement-based handover optimization using RSRP/RSRQ/SINR based on an A3 triggering event and load-based handover. It also discusses multiple machine learning algorithms, such as PPO, and the general application of SDN and NFV algorithms to handover in both low- and high- mobility scenarios.
[18]	2025	✗	✓	✗	This research does not specify the triggering event. It also relies on the control layer for handover rather than optimizing the HOSR problem. It is only applicable to slow-speed UEs and entails high computational overhead. It also has signaling overhead, and is controller-dependent without evaluating the fail-over scenario.
[13]	2024	✗	✓	✗	This research considers slow mobility with load balance. It exhibits slightly elevated latency, lacks predictive learning, and limits test scenarios, primarily using indoor femto cells. It does not align with using any specific triggering event-based baseline KPI tuning.
[19]	2024	✗	✓	✗	This limits the study to the architecture of SDN deployment and its performance specifically for the use in IoT network architecture and does not discuss handover optimization.
[20]	2024	✗	✗	LSTM (RNN)	A network slice technique is used for handover between slices rather than the traditional process. The model is also insufficiently robust for high-mobility scenarios.
[21]	2023	✗	✗	K.F and SARSA	This paper uses A3 event triggering based on a prediction and estimation algorithm. It considers only a single frequency and does not account for dense multi-frequency band deployments. The algorithms used are not effective at predicting accuracy or at learning from environmental changes.

Table 1. Cont.

Ref.	Year	Event A5	Use of SDN	ML Algorithms	Remarks
[22]	2023	✓	✓	✗	This paper surveyed the different aspects of handover management and integration with the machine learning algorithms. It also discusses A3/A5 and compares them with the A2 handover triggering event.
[23]	2023	✗	✗	C-DQN and MA-DQN	Designed for low-mobility users and lower frequency bands. Additionally, there is no proper mapping of handover parameters in the research.
[24]	2023	✗	✗	DQN	Does not specify 3GPP handover baseline KPI; instead, works on RSRQ-based handover only. Dependent on the available data only.
[25]	2023	✗	✓	✗	There is no prediction or machine learning adaptiveness in the model. It is intended only for high-speed trains and is also characterized by a high-noise environment. The simulation environment is limited to 5 base stations and a 1 km route.
[26]	2022	✗	✓	✗	The paper does not report any event-based handovers, and does not analyze the results in real time with respect to network QoS. It does not explicitly mention the mobility scenarios. It has high computational complexity as it uses an anchor point as a facility location.
This paper	2025	✓	✓	✓	More accuracy in predicting the best target cell based on subclasses in variable speed scenarios.

DDPG: Deep Deterministic Policy Gradient; PPO: Proximal Policy Optimization; HO: Handovers; MLP: Multi-Layer Perceptron; RLF: Radio Link Failure; RL: Reinforcement Learning; K.F: Kalman Filter; U.K.F: Unscented Kalman Filter; C-DQN: Centralized Deep Q-Networks; MA-DQN: Multi-Agent Deep Q-Networks; E-SARSA: Expected SARSA; LSTM: Long Short-Term Memory; RNN: Recurrent Neural Network; and C-DQN: Convergent-DQN.

2.1. Handover Process

In cellular networks, handover enables connected-mode mobility by transferring an active user equipment (UE) from a serving cell to a target cell without interrupting an ongoing session. This procedure ensures continuity as the UE moves, radio conditions fluctuate, or network-level constraints such as load balancing or slice-specific QoS requirements necessitate a change in serving cell. In 5G New Radio (NR), handover procedures are standardized by 3GPP [28] and executed through Next-Generation (NG)-based signaling. The UE contributes by providing periodic or event-triggered measurements, including Reference Signal Received Power (RSRP), Signal-to-Interference and Noise Ratio (SINR), and Channel State Information (CSI). However, the decision to initiate HO and the selection of the target cell remain network-controlled functions.

Multiple handover types are supported in NR. Intra-gNB handovers occur within a single Central Unit/Distributed Unit (CU/DU) architecture, while Xn-based inter-gNB handover facilitates mobility between neighboring gNBs. NG-based handovers involve the Access and Mobility Management Function (AMF) and typically trigger a core path update [29]. Inter-Radio Access technology (inter-RAT) handovers, e.g., NR-LTE, and vertical handover across access technologies ensure service continuity across heterogeneous radio environments. These mechanisms are vital in challenging scenarios such as high-mobility vehicular routes, high-speed rail (HSR) systems, and mmWave deployments where link blockages and beam transitions occur frequently.

2.1.1. Measurement Configuration and Reporting

The source gNB configures the UE's measurements, including the events to monitor and the conditions under which reports must be generated. Commonly used NR mobility events include A3, triggered when a neighboring cell becomes better than the serving cell by a predefined margin, and A5, triggered when the serving cell falls below a threshold while a neighboring cell exceeds that threshold, and many others, depending on the operator's use of which event in their network [29]. Additionally, it uses Time-to-Trigger (TTT) and Handover Margin (HOM) or Cell Individual Offsets (CIO) to filter out short-term signal fades, which are particularly critical in the millimeter-wave (mmWave) and terahertz (THz) frequency bands [10].

2.1.2. Decision and Preparation

Upon receiving the measurement reports, the source gNB evaluates the candidate cells' performance objectives and policy constraints, including QoS requirements, network slicing rules, backhaul conditions, load status, and UE mobility state. After selecting a target cell, it initiates preparation procedures, exchanging UE context information including security credentials, data radio bearer (DRB) configurations, and QoS flow mapping with the target gNB [30]. In the case of conditional HO (CHO), the source prepares multiple HO commands, each associated with a predefined radio condition, enabling the UE to execute the most suitable command autonomously. This mechanism reduces late and failed HOs, improving mobility robustness.

2.1.3. Handover Execution

The UE executes the HO command synchronized to the target cell, often requiring beam alignment or switching in mmWave deployments, and completes the Radio Resource Control (RRC) reconfiguration. User and data-plane bearers are then resumed. While the default NR HO procedures follow a break-before-make model, resulting in a brief interruption, 5G systems also support make-before-break through dual connectivity [21]. In this configuration, the UE maintains simultaneous links to multiple cells or gNBs, enabling smoother transitions and reduced service disruptions.

2.1.4. Handover Completion

After UE reconfiguration, the target gNB acknowledges successful HO completion. Data forwarding from the source is terminated; buffered packets are delivered; and the AMF and User Plane Function (UPF) may execute a path-switching procedure during NG-based handovers [21]. During this phase, performance indicators, including interruption time, HO completion status, and signaling statistics, are updated. These metrics feed self-organizing network (SON) optimization loops, enabling long-term mobility enhancements.

2.2. Research Gap and Novelty

In view of this, the work addresses key research gaps in state-of-the-art mobility management. First, most existing DRL and ML-based handover schemes are designed for either macro or small cell environments and do not jointly optimize mobility decisions across hybrid macro–small deployments under a wide range of user speeds (50 km/h–400 km/h) [31]. Secondly, prior studies typically apply learning at the RAN or edge level but do not integrate DQN-based decisions centrally within the SDN controller, where a global network view can be exploited for mobility optimization. Thirdly, to the best of our knowledge, no work has been published yet that explicitly binds the handover parameters, such as TTT, hysteresis margin, and target cell type, into a discrete mobility subclass that can be directly used as the DRL action space, thereby limiting the interoperability and direct controllability of learned policies. Finally, existing approaches seldom employ a unified, multi-metric reward that simultaneously accounts for the user throughput, handover success, latency, and mobility preference, i.e., favoring macro cells at higher speeds and small cells at lower speeds, which is essential for delivering robust QoS in highly dynamic, ultra-dense 5G-and-beyond networks.

3. Network Model for Hybrid 5G Cellular Network

In our evaluation, handover outcomes and failure events follow the 3GPP NR definitions based on Q_{in}/Q_{out} quality thresholds and supervision timers t_{310} , N_{310} , and N_{311} , as specified in the technical reports [10,29]. A successful handover is counted when the UE completes RRC reconfiguration to the target gNB and remains connected without declaring

radio-link failure (RLF) within a short observation window after handover completion. A radio-link failure is declared when the UE observes N310 consecutive “out of sync” indications from the lower layers and the T310 timer expires before receiving N311 “in-sync” indications, or when random-access and re-establishment procedures fail in accordance with NR mobility procedures [21]. A handover failure [32] occurs when an RLF is detected during the handover procedure, or when the connection is re-established to the source cell shortly after a handover to the target cell, in accordance with mobility robustness optimization definitions. We further classify a ping-pong handover as a sequence in which the UE hands over from cell A to cell B and then back to A within a short time window, thereby increasing signaling overhead and degrading user quality of experience (QoE). This definition is consistent with recent empirical analyses of ping-pong behavior in commercial 4G/5G networks [33].

4. Solution Approach

The proposed mobility optimization framework leverages an SDNC coupled with a DQN to intelligently determine handover decisions in an ultra-dense hybrid 5G deployment comprising macro- and small-cell networks. The SDNC enables centralized visibility of network-wide measurements, while the DQN model adaptively selects the optimal handover control parameters, TTT, hysteresis margin, and appropriate target cell type to achieve seamless mobility for users traveling at speeds up to 400 km/h.

4.1. SDN Controller Layer

The top layer of the suggested network model is centered around the SDN controller located on a cloud server. This layer will collect all measurement reports from user equipment (UEs) or wireless devices via the base station. The SDN controller will play a key role in identifying the most suitable neighboring base station to ensure a smooth handover process by utilizing a deep Q-network reinforcement learning algorithm (DQN). The SDN controller is modeled as a centralized logical control entity responsible for collecting measurement reports (RSRP, SINR, speed, and user throughput) from all connected gNBs and user equipment and executing the DQN-based decision-making process for optimal handover parameter selection. The controller maintains a global network view and applies the learned policy to determine the appropriate target cell and handover control parameters for each user. The implementation is carried out within the simulation environment as a centralized software module that emulates SDN-control plane functionality, including measurement aggregation, policy inference, and decision dissemination to base stations. While full-protocol-level open flow signaling is abstracted for computational efficiency, the controller behavior reflects practical SDN-based mobility management architectures commonly adopted in 5G research, where centralized intelligence enables coordinated and adaptive handover decisions across heterogeneous cells. By analyzing the measurement reports, this Deep Q-network model will evaluate and determine the optimal neighboring cell for a seamless transition. The measurement reports include Reference Signal Received Power (RSRP), Signal-to-Interference Noise Ratio (SINR), the UE speed, and the required data throughput based on the user’s service- and speed-based requirements as listed in Table 2. The network incorporates two types of base stations: macro base stations and small base stations. The rationale for these two types of base stations is that, when a vehicle is moving at high speeds, the SDN controller prefers macro base stations to complete handovers, whereas it opts for small base stations during moderate or low-speed mobility to ensure Quality of Service (QoS).

Table 2. Speed-based mobility subclasses.

Sr.#	Subclass	Speed (km/h)	Throughput (Mbps)	Target Cell/Tech
1	A	<100	>5	Small/5G
2	B	100 ≤ 150	>10	Small/5G
3	C	150 ≤ 200	>15	Macro/4G
4	D	200 ≤ 300	>15	Macro/4G
5	E	>300	>15	Macro/4G

We propose a network architecture that integrates the SDNC, base stations, and UEs within a high-speed train, accounting for mobility. As depicted in Figure 1, we propose an ultra-dense deployment of a hybrid architecture combining both macro and small base stations. We conceptualized the network model as a hybrid network architecture where multiple macro and micro base stations are deployed. Once the UEs report the serving base station with the measurement reports of the current state, the SDN controller is updated with the current state and the handover process is triggered. We anticipate a high-speed train with UEs and wireless devices distributed randomly across an area, moving at variable speeds ranging from 50 km/h to 400 km/h. Each gNodeB (gNB) is connected to at least 10 wireless devices, and a network configuration utilizing 50 gNBs is proposed as part of the ultra-dense network setup. The UE will transmit a measurement report to the SDNC via a base station. The proposed model consists of five (5) subclasses based on mobility speed and required data throughput. Considering user needs, these subclasses will encompass the values of TTT and HM. Therefore, depending on the user's speed and throughput requirements, the Deep Q-Network algorithm will identify the most suitable mobility subclass of handover parameter values. We adopt the 3GPP A5 event as the triggering mechanism for all handover decisions. The event A3 is activated when neighboring cell becomes better than the serving cell by a predefined threshold plus an offset value, enabling handover preparation but potentially increasing ping-pong effects in dense mmWave deployments [28]. In contrast, the A5 event deploys dual thresholds, where the serving cell quality must fall below a defined level while the target cell simultaneously exceeds another threshold, thereby reducing unnecessary handovers and improving mobility robustness in high-speed or interference-prone scenarios [29]. In contrast, the A5 event employs a dual-threshold structure, thereby significantly reducing premature handovers, late HO failures, and instability across different speed scenarios. We selected the A5 event because its dual-threshold mechanism is better suited to high-speed train mobility than the A3 event. The mobility-based subclasses are categorized in Table 2. The current model assumes a very low/negligible inference delay to isolate learning behavior. In practical deployments, controller latency of up to 5 ms primarily affects transition latency but does not destabilize the learned policy. The centralized aggregation of UE measurements limits scalability. Beyond several thousands of UEs, hierarchical SDN control or distributed RIC architectures may be required. These deployment constraints are acknowledged as system-level limitations and motivate future integration with realistic RAN simulations. In addition, our simulation environment uses a latency metric that reflects the complete application-layer end-to-end (E2E) packet delivery time rather than the intrinsic 5G radio-interface latency [28]. This distinction is essential, as 5G user-plane latency may accumulate additional delay from buffering, signaling exchanges, controller processing, and retransmissions during mobility events as per the equation [34] below:

$$L_{total} = L_{radio} + L_{HO-time} + L_{controller} + L_{queue} + L_{reTX} \quad (1)$$

where 'L' denotes latency, and 'reTX' is the retransmissions.

The latency definition used in this work follows the application layer E2E delay model commonly adopted in studies of 5G mobility and vehicular communication. In this

model, the total packet delay is decomposed into several components: radio transmission delay, buffering delay, signaling delay, and retransmission delay [35]. In this work, total latency is measured as the difference between the packet generation time at the transmitter application and the packet reception time at the receiver application. This measurement reflects the user-perceived delay rather than the physical layer transmission delay. This approach aligns with the 5G E2E latency modeling used in evaluations involving mmWave and ultra-reliable communication, where scheduling, processing, and retransmissions are explicitly included in the delay budget [36]. The radio delay L_{radio} denotes the time required for packet transmission over the air interface, including scheduling and physical-layer retransmission. As a fundamental component of latency in 5G systems [37], this delay is influenced directly by factors such as retransmission time, scheduling, and link adaptation.

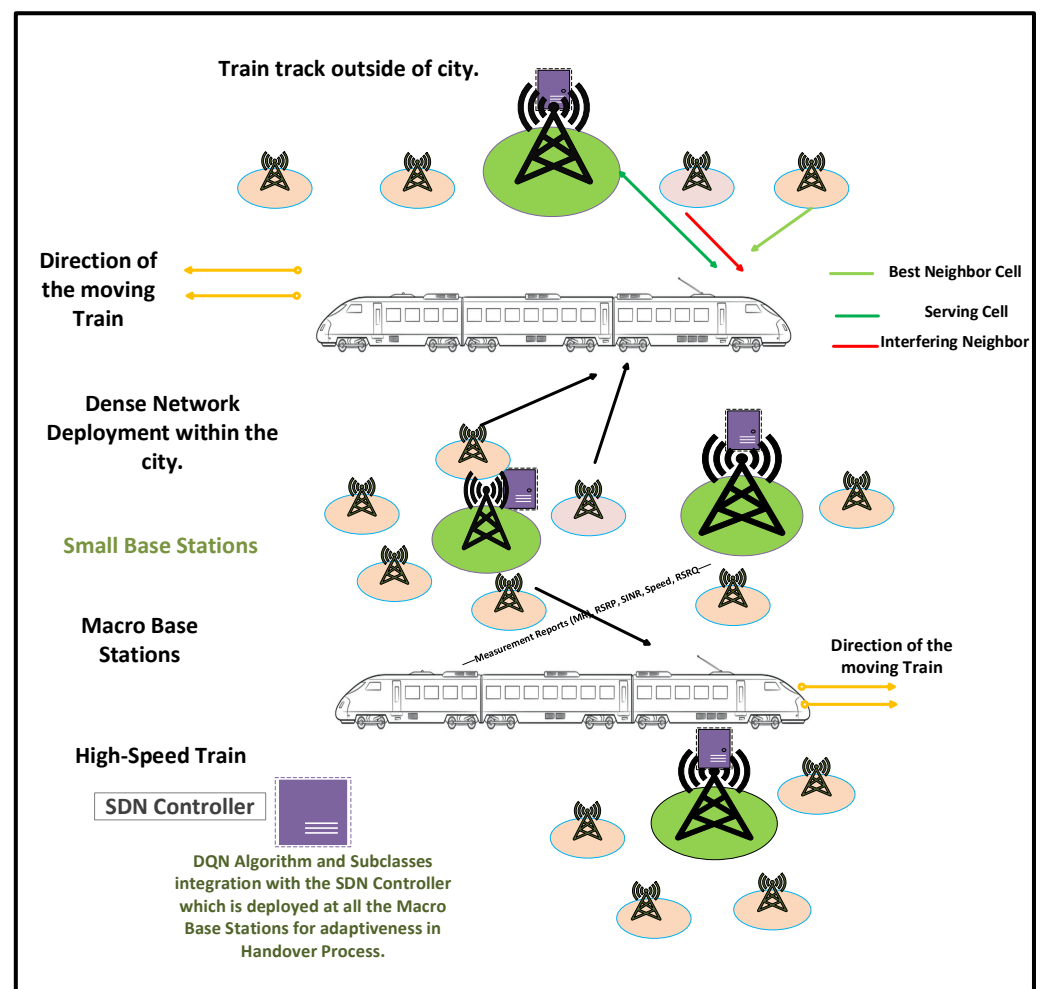
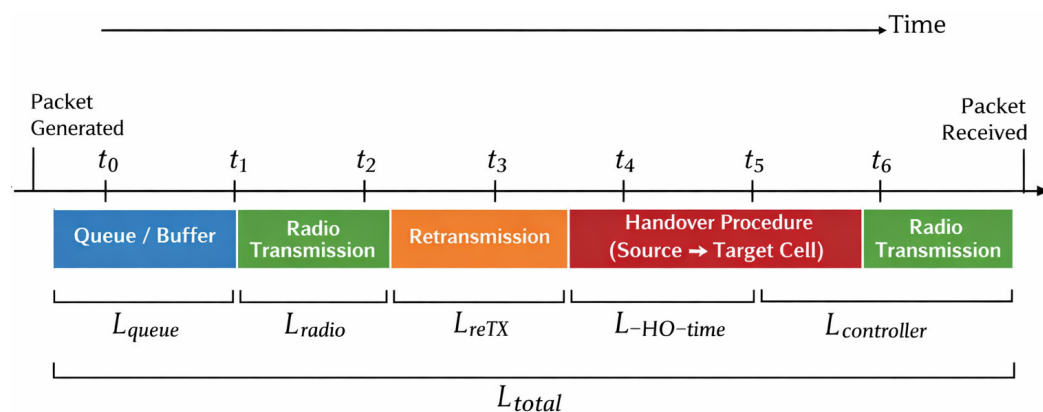


Figure 1. Network architecture (small and macro base stations).

The handover interruption delay $L_{HO-time}$ refers to the time interval between the last packet successfully received from the serving cell and the first packet received after handover is complete. This interruption time is a standard performance metric in mobility studies and is defined in the 3GPP NR RRC technical specifications [10]. Handover procedures can cause temporary data interruptions during signaling exchange and target-cell synchronization. The controller delay L_{queue} represents the time packets spend waiting in the transmission buffers at the user equipment or base station before they are scheduled for transmission. Queuing delay can become significant during mobility and periods of congestion, and it is included in E2E latency modeling for heterogeneous and vehicular networks [38]. The retransmission delay L_{reTX} represents additional delay caused by radio

link control retransmission when packet errors occur, particularly under conditions of fast mobility or poor channel quality. Retransmissions are known to increase latency in high-speed scenarios within mmWave and small-cell environments [36].

Figure 2 shows a timing diagram illustrating the decomposition of the total E2E latency used in latency equation. The total latency is measured from packet generation at the transmitter application to packet reception at the receiver application and is divided into queuing delay, radio transmission delay, retransmission delay, handover interruption time, and SDN controller processing delay. This decomposition allows accurate evaluation of mobility-induced latency during handover events in the proposed SDN-assisted handover framework.



$$L_{total} = L_{queue} + L_{radio} + L_{reTX} + L_{HO-time} + L_{controller} + L_{radio}$$

Figure 2. Latency timing diagram.

4.2. Deep Q-Network Reinforcement Algorithm

RL is a powerful machine learning approach that enables agents to develop optimal decision strategies by interacting with their environment, aiming to maximize total rewards. This method is advantageous in dynamic settings, such as cellular networks, where systems must adapt and improve performance in real-time. The capability of RL to learn through direct interactions without requiring labeled data makes it adaptable for a range of applications, including user association and handover processes in cellular networks. Studies have shown that RL, particularly when combined with deep learning techniques, can significantly enhance handover efficiency and user association, surpassing traditional methods specified by standards such as the Third Generation Partnership Project (3GPP) [23,24]. Table 3 lists the DQN hyperparameters used in our simulations to ensure reproducibility. However, the learning workflow of the proposed Deep Q-Network-based mobility optimization framework is illustrated in Figure 3. The model processes the observed network state, including RSRP, SINR, user velocity, and user throughput requirements, and maps these inputs to the optimal mobility subclass and corresponding handover parameters through Q-value estimation. The experience replay mechanism and target network update enable stable learning and convergence, allowing the SDN controller to continuously refine its handover decisions based on observed network dynamics.

The fundamental components of RL consist of the agent, actions, states, and rewards.

- State is the portrayal of the current condition of the environment.
- Action acts as the decision-maker—the environment, which reflects the context where the agent functions and is the selection made by the agent.
- Reward offers feedback regarding the agent's actions.

- Policy is the mapping from states to actions that maximizes expected cumulative reward.

The flexibility of RL algorithms enables them to operate effectively in complex environments with multiple variables, making them appropriate for situations in which traditional methods may struggle. Moreover, RL's emphasis on optimizing long-term strategies is especially beneficial for tasks that demand consistent performance over an extended period, such as energy management and resource allocation in cellular networks [4,39].

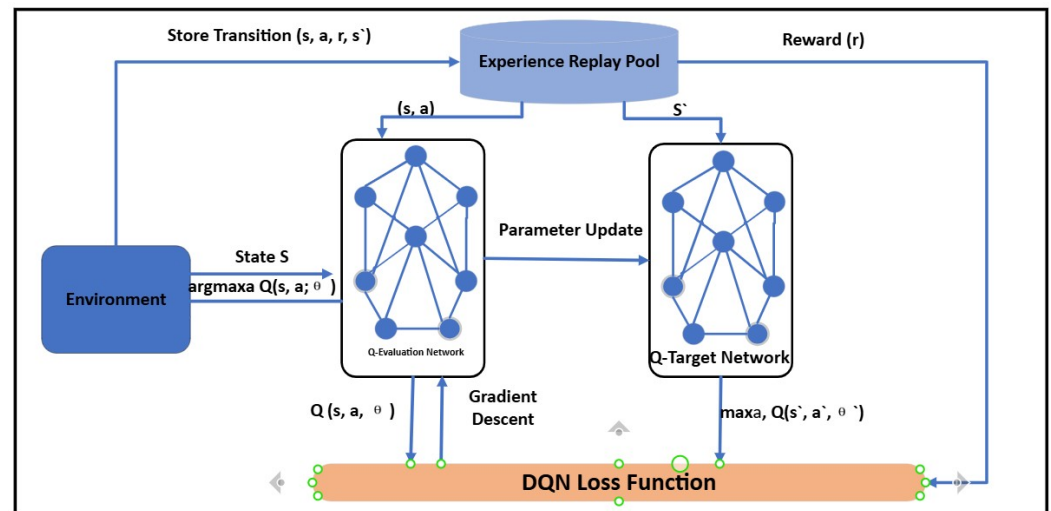


Figure 3. DQN learning model.

Table 3. DQN hyperparameters.

Neural Network Architecture	Value
Hidden layers	3
Neurons per layer	128 → 64 → 32
Activation	ReLU
Training Hyperparameter	
Replay buffer size	100,000
Mini-batch size	64
Learning rate	1×10^{-4}
Discount factor (γ)	Adam Optimizer
Target update	0.99
Training episodes	Every 10,000 steps
	5000+
Exploration policy (ϵ -greedy)	
Initial ϵ	1.0
Final ϵ	0.1
Decay	Linear steps 100,000
Reward function weights	
α (throughput reward)	0.5
β (handover penalty)	0.3
γ (latency penalty)	0.1
λ (Mobility bias penalty)	0.1

In our system model, we introduced different subclasses [40] that contain the set of handover parameter values (TTT and HM) and serve as the action stage in the Deep Q-Network model. The state space of the model contains the RSRP, SINR, UE speed, and the user throughput. Based on the state and action stages, the approach to resolve the handover challenge using the Deep Q network will be as follows. Let the state S_t of the model be:

$$S_t = [\xi_{t-t_2}, \omega_{t-t_2}, v_t, H_t, R_t^{\text{req}}] \quad (2)$$

where ξ_{t-t_2} are the current and historical coverage RSRP events, ω_{t-t_2} are the current and historical SINR values, v_t is the speed at which a UE is moving, H_t is the previous state of handover (either '0' if it did not occur or '1' if it occurred), and R_t^{req} is the average required throughput.

4.3. Deep Q-Networks

Deep Q-Networks (DQN) represent a revolutionary blend of RL and deep neural networks, designed to address complex decision-making challenges. The central concept is to train a deep neural network, referred to as the Q-network, to approximate the ideal action-value function $Q(s, a)$, which predicts the expected cumulative reward for action a in state s while adhering to the optimal policy thereafter. The algorithm progresses through several essential steps:

- **Experience Collection:** The agent interacts with an environment using an ϵ -greedy policy [41]. With probability ϵ , it selects a random action for exploration; otherwise, it chooses the action maximizing the current Q-value estimate. Each transition (s_t, a_t, r_t, s_{t+1}) comprising the current state, action, reward, and next state is stored in a replay buffer.
- **Replay Buffer Sampling:** Instead of learning from consecutive experiences, DQN randomly samples a mini batch of transitions from the replay buffer. This experience replay mechanism breaks temporal correlations, stabilizes training, and enables data reuse.
- **Q-Target Calculation:** For each sampled transition, the target Q-value is computed using a separate target network (a periodic copy of the main Q-network). The target is defined as:

$$y_j = r_j + \gamma \max_{a'} Q_{\text{target}}(s_{j+1}, a'; \theta^-) \quad (3)$$

where r_j is the immediate reward, γ is the discount factor (e.g., 0.99), and θ^- are the parameters of the target network. If s_{j+1} is a terminal state, then $y_j = r_j$.

- **Loss Minimization:** The main Q-network updates its parameters θ by minimizing the mean squared error (MSE) between its predictions $Q(s_j, a_j; \theta)$ and the targets y_j . This optimization is performed via gradient descent with the Adam optimizer [41], which iteratively updates the network parameters to minimize the prediction error. Target network synchronization helps enhance the stability of the learning process. Every 10,000 steps, the target network's weights θ^- are updated to match the main Q-network's weights θ . This "soft update" stabilizes learning by decoupling the target values from rapidly changing parameters.

The Action space at time 't' would be:

$A_i^t = \{C_1, C_2, \dots, C_n\}^t$, where each class defines the handover parameters and target base station type, and 'n' is the number of classes.

$$C_i = \{TTT_i, HM_i, TCT_i\} \quad (4)$$

where TTT = time-to-trigger, HM = δ threshold, and TCT = target cell type (macro, small).

For reward r_t , we define it as a mobility preference as:

$$r_t = \begin{cases} +\lambda, & \text{if } v_t > v_{\text{high}} \text{ and TargetCellType} = \text{macro} \\ -\lambda, & \text{if } v_t < v_{\text{high}} \text{ and TargetCellType} = \text{small} \end{cases} \quad (5)$$

The complete reward function would be:

$$r_t = \alpha \left(\frac{R_t^{ach}}{R_t^{req}} \right) - \beta HO_t - \gamma L_t + r_t^{mob} \quad (6)$$

- α : Tunable weightage factor for required and achieved user throughput.
- β : Tunable weight for the handover penalty.
- γ : Penalty weight for average measured packet latency; normally ranges between 0.01 and 0.1.
- HO_t : Handover penalty; '1' means handover occurred, and '0' means no handover.

The DQN policy and the loss function are shown as [41]:

- Policy: $Q(S_t, C_i, \theta) \rightarrow$ Expected long-term value of QoS based on the class C_i .
- Target: $U_t = r_t + \gamma * C_j^{max} Q(S_t, C_i, \theta)$.
- Loss: $L(\theta) = (U_t - Q(S_t, C_i, \theta))^2$.

4.4. Reward, Weight Selection, Sensitivity, and Mobility Preference

The reward function in this work combines throughput satisfaction, handover stability, and latency objectives into a single scalar signal for learning. Such scalarization is necessary because mobility management in dense 4G/5G deployment is intrinsically multi-objective; aggressive handover improves instantaneous radio quality but increases signaling overhead and ping-pong risk, whereas conservative handover reduces handover rate at the cost of degraded throughput and interruption events. Recent studies show that the RL-based mobility controllers explicitly highlight this tension and adopt multi-term objectives to balance handover frequency and service quality under fast varying radio conditions [23]. Similarly, Deep RL-based handover optimization for ultra-dense 5G networks shows that stable learning depends on carefully balancing throughput-driven terms against handover-related penalties to avoid degenerate policies that over-trigger handovers [8].

Weight selection methodology: The weights $(\alpha, \beta, \gamma, \lambda)$ are not selected as arbitrary constants; instead, they are treated as tuning variables in a controlled procedure aimed at robustness across mobility regimes. First, each reward component is bounded or normalized to mitigate unit mismatch, an approach commonly used in wireless DRL when combining heterogeneous metrics (e.g., handover rate, delay, and handover count) in a single objective [23]. Secondly, we select candidate weights using a two-stage process: (i) a coarse sweep within a feasible simple $\alpha + \beta + \gamma + \lambda = 1$ to identify Pareto-efficient regions, and (ii) local refinement around stable candidates using repeated training under different random seeds and speed profiles. This tuning philosophy is consistent with handover-optimization studies where handover control space is highly sensitive to mobility and fading dynamics, for example: adaptive TTT/HM optimization demonstrating that small parameter shifts can materially change handover latency, failures, and throughput trade-offs, motivating systematic calibration rather than single-point settings [21]. To ensure that the performance improvements are not dependent on a single heuristic reward configuration, a sensitivity and stability is conducted by perturbing each reward weight within a bounded range while maintaining normalized scaling among all components and evaluating the resulting impact on handover success, ping-pong effect, packet reliability, and latency. Training stability is further verified across multiple random seeds to confirm consistent convergence, as robustness evaluation is essential in DRL-based mobility optimization where exploration variance and dynamic channel conditions can otherwise bias results [23,42]. The mobility term is used as a weak regularizer rather than a hard-coded rule, consistent with reward-shaping principles that incorporate domain knowledge to improve convergence while preserving the optimal policy structure [43]. To confirm that

the agent learns autonomously rather than being guided, an ablation experiment with the mobility-preference weight set to zero is performed, if the model continues to learn effective cell-selection behavior but exhibits increased handover oscillations near mobility transitions, which demonstrates that QoS-driven learning remains the dominant factor and the preference term only enhances stability, consistent with established DRL-based handover optimization methodologies [39].

4.5. Baseline Schemes and Fair Tuning

For a fair comparison, all baseline schemes are evaluated under the same network topology, traffic model, and mobility traces as those used for the proposed SDN-DQN framework. LIM2 is implemented by having a Kalman filter predict future RSRP values and a SARSA agent adapt TTT and hysteresis margins using an ϵ -greedy policy. PPO- and DDQL-based baselines follow the same state representation (RSRP, SINR, UE speed, and a service class) and action space (TTT/HM/target-cell-type subclasses) as our method, but differ in the policy update rule and value function estimation, consistent with recent RL-based mobility designs in 5G networks [23].

4.6. Control Plane Signaling and Decision Latency

The proposed SDN-DQN controller is logically placed above the gNBs or on a train coach, within the near-real-time RIC architecture envisioned in O-RAN, and in recent SD-RAN mobility management work [14]. In such architectures, mobility-related control loops are executed on time scales between a few tens of milliseconds and roughly one second, depending on the use case and deployment, with stringent requirements for closed-loop decision latency over the E2/Xn interface and within the controller itself. In our simulation, the SDN controller ingests periodic measurement reports, executes a single forward pass of the DQN, and sends updated handover parameters and target-cell decisions back to the serving gNB. We model signaling overhead as the number of handovers and measurement reports per UE, and penalize excessive handover rates in the reward to keep control-plane load comparable to or lower than state-of-the-art RL mobility schemes. Although we assume negligible inference time for the DQN in the study, the selected network size is compatible with near-real-time (RT) deployments at the edge, where hardware accelerators routinely achieve sub-millisecond inference latency for similar models.

5. Performance Analysis

Table 4 additionally ensures fairness and reproducibility across all learning-based methods, we have explicitly documented the hyper parameters, convergence characteristics, and runtime complexity of all evaluated algorithms. Table 5 enlists the parameters and their specified values used in our simulation work derived from [7,21]. In our comparative analysis, we have considered the handover success rate, user throughput, packet transfer success rate, and end-user latency. These simulations were executed using Python 3.11. There are 35 small cells and 15 macro cells for umbrella coverage. The number of UEs is 500, and the mobility speed ranges from 50 km/h to 400 km/h. All the algorithms were implemented using the same state representation and equivalent action-space structure, ensuring a consistent comparison framework. Hyper parameters were selected through controlled tuning aimed at achieving stable convergence rather than peak isolated performance for any single method, following established best practices in reinforcement learning experimentation that emphasize robustness and reproducibility across training seeds and mobility scenarios rather than single-run optimization.

Table 4. Baseline configuration and training details.

Algorithm	Training Episodes	Learning Rate	Batch Size	Convergence Behavior	Runtime Complexity (h)
DQN	5000+	1×10^{-4}	64	Stable with low variance	2.1
DDQL	5000+	1×10^{-4}	64	Stable but slightly slower	2.4
PPO	5000+	3×10^{-4}	128	Gradual with higher variance	3
LIM2	5000+	5×10^{-4}	64	Scenario-dependent	2.7
Adaptive baseline	5000+	1×10^{-4}	64	Stable under fixed mobility	2

Table 5. Simulation parameters.

Parameters	Value
Cell Radius-Small BS	150 m
Cell Radius-Macro BS	1500 m
Center Frequency (Small BS)	26 GHz
Center Frequency (Macro BS)	950 MHz
UE Tx Power	23 dBm
gNB Tx Power	46 dBm
Velocity Model	Random Way Point Mobility
Channel Bandwidth (Small BS)	400 MHz
Channel Bandwidth (Macro BS)	20 MHz
Speed of UEs	50–400 Km/h
Number of UEs	500
Path loss Model	Log-normal path loss
Fading Model	Friis spectrum propagation loss
Propagation Model	Constant speed propagation delay
Bit error rate	0.003%
UE noise figure	9 dB
gNB noise figure	5 dB
Thermal noise density	−174 dBm/Hz
Time-to-trigger	0–7 ms
Hysteresis Margin	1–3 dB
Simulation time	20 s × 100 runs
values of α and γ	0.5 and 0.9
Number of Base Stations (small)	35
Number of Base Stations (Macro)	15

Handover Success Rate versus Speed: Figure 4 shows the handover success rate as a function of UE speed and SINR, demonstrating that the proposed SDN-based Deep Q-Network model consistently outperforms our previously proposed model, as well as the LIM2, PPO, and DDQL models, across all mobility scenarios. These results highlight the effectiveness of Reinforcement Learning (RL) in enabling context-aware, adaptive mobility management. By utilizing historical SINR, coverage patterns (RSRP), and velocity data, the SDN-based DQN model provides a strong alternative to traditional static threshold-based strategies. It significantly reduces service interruptions and dropped connections in high-mobility scenarios. Specifically, the 10–15% improvement is obtained from the comparative analysis shown in Figure 4a,b, where the proposed SDN-DQN-based model consistently outperforms the baseline models across SINR and different speed scenarios.

User Throughput versus Speed: Figure 5 presents a comparative analysis of five mobility models with respect to average user throughput for all successful UEs in the target cell after handover completion. The SDN-based DQN-based mobility model exhibits slightly reduced user throughput, but outperforms the other four reinforcement learning models due to a higher number of successful handovers and a higher HOSR. The trends in the graphs indicate that our current proposed model also achieves higher user throughput than

our previous high-mobility model. In addition to this comparative analysis, we observed significant improvements over other models including LIM2, DDQL, and PPO. The results presented indicate that the model exhibits significant adaptive target-cell selection capacity. Specifically, it reveals a tendency to favor small cells at lower speeds, particularly in scenarios with higher data throughput requirements.

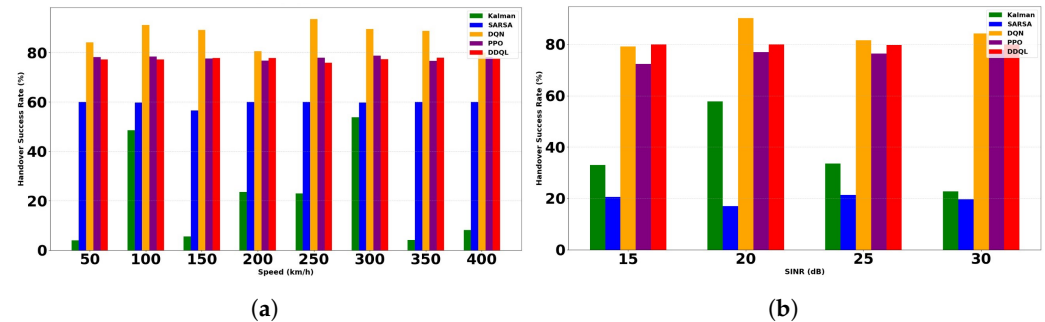


Figure 4. Handover success rate when the total number of UE is 500 versus (a) speed and (b) SINR.

Furthermore, the reinforcement learning (RL) framework employed enables intelligent, context-aware decision-making, thereby enhancing both spectral efficiency and overall throughput. This underscores the model's robust adaptability in terms of QoS. As a result, it proves especially effective in ultra-dense, heterogeneous network environments, where traditional rule-based approaches often struggle to maintain consistent user performance.

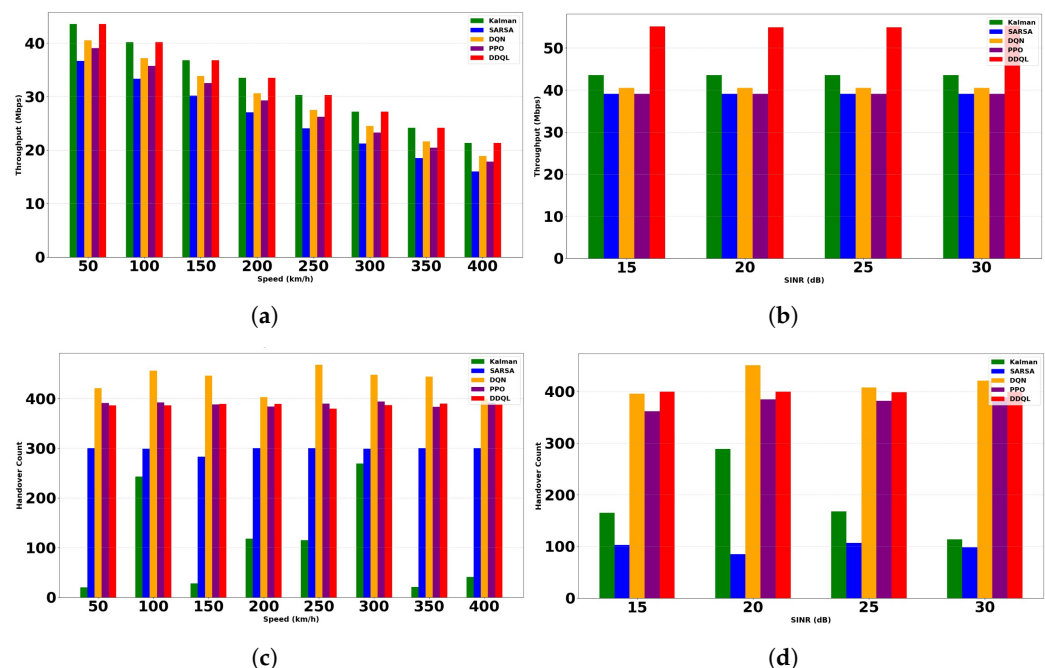


Figure 5. User throughput versus (a) speed and (b) SINR; and (c,d) handover count.

Packet Failure Ratio versus Speed: The packet failure rate graphs in Figure 6 demonstrate that the proposed SDN-based DQN model consistently achieves higher delivery reliability compared to LIM2, our previous work, DDQL, and PPO models across all mobility ranges as well as in SINR comparative analysis. As shown in the graphs, the value remains approximately below 9% across all ranges. This improvement is due to the model's ability to predict and adapt to changing channel conditions and interference by intelligently selecting mobility subclasses based on a Q-value approximation. By dynamically configuring handover parameters and determining the optimal cell types, the model ensures reliable packet transmission, even in challenging radio environments. This increased success rate is

particularly beneficial for real-time and delay-sensitive applications, mission-critical IoT, and high-mobility devices, as retransmission can negatively impact performance.

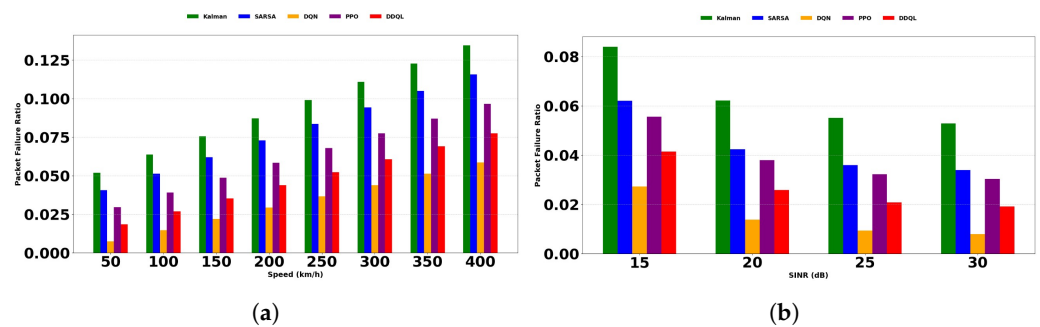


Figure 6. Packet failure ratio versus (a) speed and (b) SINR.

Latency versus Speed and SINR: The latency analysis in Figure 7 reflects the end-to-end packet delivery time, which includes handover interruption, control-plane signaling, and packet retransmissions during mobility, rather than the intrinsic 5G radio-interface latency. This explains the reason for the measured values being slightly higher than typical 5G user-plane latency targets, which are normally in the milliseconds range. The slight increase in latency observed for the SDN-DQN-based model arises from the additional SDN decision loop and more frequent proactive handovers, both of which contribute to temporary buffering and signaling overhead. Therefore, the latency metric reported in this study should be interpreted as an application-level end-to-end delay under high mobility conditions, not as a violation of standardized 5G latency specifications. Despite this, the model maintains consistent latency performance across varying UE speed variations, indicating robustness under diverse channel environments.

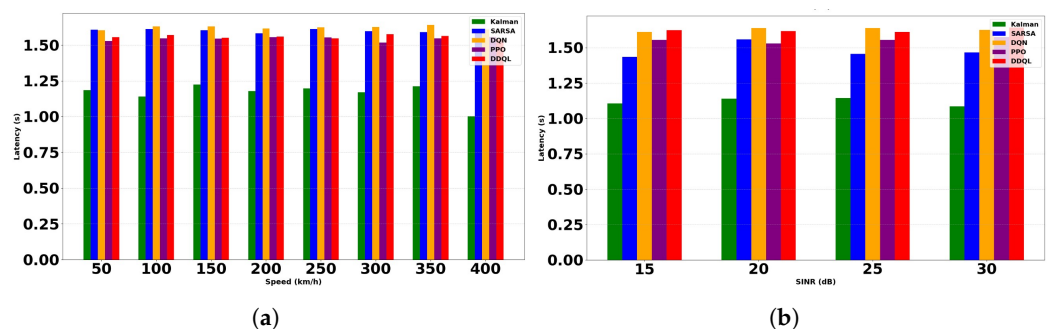


Figure 7. Latency versus (a) speed and (b) SINR.

In our simulation settings, the most likely failure regimes align with three stressors, i.e., rapid changing speed, speed coupled with SINR degradation via the exponential model, and ultra dense contention with 500 UEs served by 50 gNBs. First, mobility regime transition failures are most likely at the subclass boundaries, where the controller may switch between small- and macro-cell preferences while the SINR changes with speed via the exponential decay term. In these boundary zones, the radio environment becomes temporarily non-stationary, which can induce short-lived policy oscillations. Secondly, the other worst case scenario is packet reliability and latency spikes concentrating when high speed combines with fast SINR decay and dense neighbor competition, where the serving link quality can drop rapidly while multiple candidate cells remain close in quality, especially given short cell radii. In such conditions, handover execution may occur while simultaneously degrading the serving and target quality, producing bursts of packet failures and tail latency even if the mean latency remains acceptable. We can address these worst cases by proposing a tail-aware optimization technique that penalizes latency or packet-loss violations beyond threshold levels, ensuring the learning process accounts for rare

degradation events rather than only mean performance [23]. In high-speed scenarios where rapid SINR decay may cause burst losses during handover execution, make-before-break or dual-connectivity strategies may possibly be able to maintain link continuity and reduce interruption risk [44].

6. Conclusions

Effectively managing seamless handovers in hybrid 5G-and-beyond cellular network conditions is a complex challenge. Poor management can degrade QoS and adversely affect the end-user experience. This issue has been the focus of extensive research, and we propose an innovative Deep Q-Network-based reinforcement learning (RL) model integrated with Software-Defined Networking (SDN) to enhance handover performance in hybrid 5G environments.

Our approach introduces a mobility subclass-based framework that adaptively configures handover parameters. This framework dynamically adjusts these parameters via an online policy that accounts for environmental conditions, yielding a more flexible solution. The model is designed to address various hybrid network conditions across different speed scenarios, including low- and high-mobility scenarios, such as high-speed trains.

Simulation results show significant improvements: up to a 15% increase in handover success rates and approximately a 20% increase in average user throughput, as indicated by Signal-to-Interference-plus-Noise Ratio (SINR) values. This model is well-suited for a range of mobility scenarios, from 50 km/h to 400 km/h, within 5G networks, ensuring optimal QoS.

Moving forward, it is essential to develop a dual-technology-based model that addresses the service-based requirements of both nearby and distant mobile users, particularly in ultra-dense urban environments where diverse mobility conditions are prevalent in smart cities. Additionally, while increased flexibility in handovers is beneficial, it also exposes the system to greater risks, such as man-in-the-middle attacks and handover hijacking. Future developments should include trust models or secure reinforcement algorithms capable of detecting and avoiding malicious base stations or signal spoofing in real-time, full-stack validation using ns-3 NR or O-RAN RIC integration.

Author Contributions: Conceptualization, Y.S.J. and W.E.; methodology, Y.S.J. and W.E.; software, Y.S.J.; validation, B.S.C.; investigation, Y.S.J., F.K.S. and W.E.; writing—original draft, Y.S.J.; writing—review & editing, F.K.S., B.S.C. and W.E.; and supervision, F.K.S., B.S.C. and W.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Singh, R.; Larsen, L.M.; Ollora Zaballa, E.; Berger, M.S.; Kloch, C.; Dittmann, L. Enabling Green Cellular Networks: A Review and Proposal Leveraging Software-Defined Networking, Network Function Virtualization, and Cloud-Radio Access Network. *Future Internet* **2025**, *17*, 161. [\[CrossRef\]](#)
2. Ullah, Y.; Roslee, M.; Mitani, S.M.; Sheraz, M.; Ali, F.; Aurangzeb, K.; Osman, A.F.; Ali, F.Z. A survey on AI-enabled mobility and handover management in future wireless networks: Key technologies, use cases, and challenges. *J. King Saud Univ. Comput. Inf. Sci.* **2025**, *37*, 47. [\[CrossRef\]](#)
3. Riaz, H.; Öztürk, S.; Aldirmaz-Colak, S.; Çalhan, A. A Handover Decision Optimization Method Based on Data-Driven MLP in 5G Ultra-Dense Small Cell HetNets. *J. Netw. Syst. Manag.* **2025**, *33*, 31. [\[CrossRef\]](#)

4. Voigt, J.; Gu, P.J.; Rost, P.M. A Deep Reinforcement Learning-based Approach for Adaptive Handover Protocols. In *Proceedings of the 2025 14th International ITG Conference on Systems, Communications and Coding (SCC)*; IEEE: Piscataway, NJ, USA, 2025; pp. 1–6.
5. Dulaj, K.; Alhammadi, A.; Shayea, I.; El-Saleh, A.A.; Alnakhli, M. Harnessing Machine Learning for Intelligent Networking in 5G Technology and Beyond: Advancements, Applications and Challenges. *IEEE Open J. Intell. Transp. Syst.* **2025**, *6*, 605–633. [\[CrossRef\]](#)
6. Kumar, T.S.; Roslee, M.; Jayapradha, J.; Ullah, Y.; Sudhamani, C.; Mitani, S.M.I.; Osman, A.F.; Ali, F.Z. Advanced handover optimization (AHO) using deep reinforcement learning in 5G networks. *J. King Saud Univ.—Comput. Inf. Sci.* **2025**, *37*, 115. [\[CrossRef\]](#)
7. Junejo, Y.S.; Shaikh, F.K.; Chowdhry, B.S.; Ejaz, W. Adaptive Handover Management in High-Mobility Networks for Smart Cities. *Computers* **2025**, *14*, 23. [\[CrossRef\]](#)
8. Kwong, C.F.; Shi, C.; Liu, Q.; Yang, S.; Chieng, D.; Kar, P. Autonomous handover parameter optimisation for 5G cellular networks using deep deterministic policy gradient. *Expert Syst. Appl.* **2024**, *246*, 122871. [\[CrossRef\]](#)
9. Aouedi, O.; Ortiz, F.; Lagunas, E.; Vu, T.X.; Chatzinotas, S. Hybrid Model-Aided Learning for 5G-NTN Handover in High-Mobility Platforms. In *Proceedings of the IEEE INFOCOM 2025-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*; IEEE: Piscataway, NJ, USA, 2025; pp. 1–6.
10. 3GPP. NR; Radio Resource Control (RRC) Protocol Specification; Technical Report TS 38.331; 3GPP: Sophia Antipolis, France, 2024.
11. Sohaib, R.M.; Onireti, O.; Tan, K.; Sambo, Y.; Swash, R.; Imran, M.A. Meta-transfer learning-based handover optimization for V2N communication. *IEEE Trans. Veh. Technol.* **2024**, *73*, 17331–17346. [\[CrossRef\]](#)
12. Kabeer, M.; Nordin, R.; Behjati, M.; Lun, L.S. Sequence-Based Deep Learning for Handover Optimization in Dense Urban Cellular Network. *arXiv* **2025**, arXiv:2510.02958. [\[CrossRef\]](#)
13. Khan, A.; Ahmad, S.; Ali, I.; Hayat, B.; Tian, Y.; Liu, W. Dynamic mobility and handover management in software-defined networking-based fifth-generation heterogeneous networks. *Int. J. Netw. Manag.* **2025**, *35*, e2268. [\[CrossRef\]](#)
14. Prado, A.; Ciki, M.; Mehmeti, F.; Kellerer, W. Enhanced mobility management with SD-RAN in 5G networks. In *Proceedings of the 2024 IFIP Networking Conference (IFIP Networking)*; IEEE: Piscataway, NJ, USA, 2024; pp. 159–167.
15. Khurana, J.; Tripathy, A.; Mishra, A.; KN, R.P.; Louis, M.T.; Yadav, R.K. Handover Management for High Speed Mobile Train Communications in 5G Networks. In *Proceedings of the 2025 International Conference on Networks and Cryptology (NETCRYPT)*; IEEE: Piscataway, NJ, USA, 2025; pp. 365–370.
16. Viviana, P.V.; Danny, S.G.; Felipe, G.; Pablo, L.M. A Systematic Literature Mapping of Machine Learning for Handover Optimization in Cellular Networks. *IEEE Access* **2025**, *13*, 144699–144732. [\[CrossRef\]](#)
17. Saoud, B.; Shayea, I.; Alnakhli, M.A.; Mohamad, H. Mobility and Handover Management in 5G/6G Networks: Challenges, Innovations, and Sustainable Solutions. *Technologies* **2025**, *13*, 352. [\[CrossRef\]](#)
18. Dao, D.T.; Vo, D.B. A software-defined networking-based load-balanced handover control scheme for distributed mobility management in vehicular communication. *Univ. Danang—J. Sci. Technol.* **2025**, *23*, 55–61. [\[CrossRef\]](#)
19. Al-Shareeda, M.A.; Alsadhan, A.A.; Qasim, H.H.; Manickam, S. Software defined networking for Internet of things: Review, techniques, challenges, and future directions. *Bull. Electr. Eng. Inform.* **2024**, *13*, 638–647. [\[CrossRef\]](#)
20. Tariq, M.A.; Saad, M.M.; Ajmal, M.; Siddiqua, A.; Seo, J.; Haishan, Y.; Kim, D. Network slice traffic demand prediction for slice mobility management. In *Proceedings of the 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*; IEEE: Piscataway, NJ, USA, 2024; pp. 281–285.
21. Karmakar, R.; Kaddoum, G.; Chattopadhyay, S. Mobility management in 5G and beyond: A novel smart handover with adaptive time-to-trigger and hysteresis margin. *IEEE Trans. Mob. Comput.* **2022**, *22*, 5995–6010. [\[CrossRef\]](#)
22. Zaidi, S.M.A.; Farooq, H.; Rizwan, A.; Abu-Dayya, A.; Imran, A. A framework to address mobility management challenges in emerging networks. *IEEE Wirel. Commun.* **2023**, *30*, 90–97. [\[CrossRef\]](#)
23. Prado, A.; Stöckeler, F.; Mehmeti, F.; Krämer, P.; Kellerer, W. Enabling proportionally-fair mobility management with reinforcement learning in 5G networks. *IEEE J. Sel. Areas Commun.* **2023**, *41*, 1845–1858. [\[CrossRef\]](#)
24. Wei, Y.; Lung, C.H.; Ajila, S.; Cabrera, R.P. Deep Q-networks assisted pre-connect handover management for 5G networks. In *Proceedings of the 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
25. Sinha, A.; Uduthalapally, V.; Das, D.; Mahapatra, R. SDN-based seamless mobility management for B5G services in high-speed railways. In *Proceedings of the 2023 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*; IEEE: Piscataway, NJ, USA, 2023; pp. 294–299.
26. Amiri, E.; Wang, N.; Vural, S.; Tafazolli, R. Dynamic anchor point selection in software defined distributed mobility management. In *Proceedings of the 2022 IEEE Symposium on Computers and Communications (ISCC)*; IEEE: Piscataway, NJ, USA, 2022; pp. 1–7.
27. Vankayala, S.K.; Gollapudi, S.K.S.; Singh, S.; Jain, B.; Yoon, S.; Bashir, A.K. Deep-learning based proactive handover for 5G/6G mobile networks using wireless information. In *Proceedings of the 2022 IEEE Globecom Workshops (GC Wkshps)*; IEEE: Piscataway, NJ, USA, 2022; pp. 461–466.
28. 3GPP TS 22.261; Service Requirements for the 5G System; Stage 1. 3GPP: Sophia Antipolis, France, 2024.

29. 3GPP. NR; NR and NG-RAN Overall Description; Stage 2; Technical Report TS 38.300; 3GPP: Sophia Antipolis, France, 2024.
30. Mahamod, U.; Mohamad, H.; Shaye, I.; Alhamadi, A.; Asuhaimi, F.A.; Othman, M. Handover Performance of Mobility Robustness Optimization for Beyond 5G Networks. In *Proceedings of the 2024 IEEE 14th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*; IEEE: Piscataway, NJ, USA, 2024; pp. 245–249.
31. Baz, A.; Logeshwaran, J.; Natarajan, Y.; Patel, S.K. Enhancing mobility management in 5G networks using deep residual LSTM model. *Appl. Soft Comput.* **2024**, *165*, 112103. [[CrossRef](#)]
32. Alraih, S.; Nordin, R.; Abu-Samah, A.; Shaye, I.; Abdullah, N. F. A survey on handover optimization in beyond 5G mobile networks: Challenges and solutions. *IEEE Access* **2023**, *11*, 59317–59345. [[CrossRef](#)]
33. Zidic, D.; Mastelic, T.; Kosovic, I.N.; Cagalj, M.; Lorincz, J. Analyses of ping-pong handovers in real 4G telecommunication networks. *Comput. Netw.* **2023**, *227*, 109699. [[CrossRef](#)]
34. Larrabeiti, D.; Contreras, L.M.; Otero, G.; Hernández, J.A.; Fernandez-Palacios, J.P. Toward end-to-end latency management of 5G network slicing and fronthaul traffic. *Opt. Fiber Technol.* **2023**, *76*, 103220. [[CrossRef](#)]
35. Zheng, K.; Zheng, Q.; Chatzimisios, P.; Xiang, W.; Zhou, Y. Heterogeneous vehicular networking: A survey on architecture, challenges, and solutions. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 2377–2396. [[CrossRef](#)]
36. Mezzavilla, M.; Zhang, M.; Polese, M.; Ford, R.; Dutta, S.; Rangan, S.; Zorzi, M. End-to-end simulation of 5G mmWave networks. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 2237–2263. [[CrossRef](#)]
37. Sun, S.; Rappaport, T.S.; Shafi, M.; Tang, P.; Zhang, J.; Smith, P.J. Propagation models and performance evaluation for 5G millimeter-wave bands. *IEEE Trans. Veh. Technol.* **2018**, *67*, 8422–8439. [[CrossRef](#)]
38. Akyildiz, I.F.; Kak, A. The internet of space things/cubesats. *IEEE Netw.* **2019**, *33*, 212–218. [[CrossRef](#)]
39. Wang, H.; Li, B. Double-deep Q-learning-based handover management in mmWave heterogeneous networks with dual connectivity. *Trans. Emerg. Telecommun. Technol.* **2024**, *35*, e4907. [[CrossRef](#)]
40. Ejaz, W.; Ibnkahla, M. Multiband spectrum sensing and resource allocation for IoT in cognitive 5G networks. *IEEE Internet Things J.* **2018**, *5*, 150–163. [[CrossRef](#)]
41. Wu, Y.; Zhao, G.; Ni, D.; Du, J. Dynamic handoff policy for RAN slicing by exploiting deep reinforcement learning. *EURASIP J. Wirel. Commun. Netw.* **2021**, *2021*, 61. [[CrossRef](#)]
42. Mollel, M.S.; Abubakar, A.I.; Ozturk, M.; Kaijage, S.F.; Kisangiri, M.; Hussain, S.; Imran, M.A.; Abbasi, Q.H. A survey of machine learning applications to handover management in 5G and beyond. *IEEE Access* **2021**, *9*, 45770–45802. [[CrossRef](#)]
43. Ng, A.Y.; Harada, D.; Russell, S. Policy invariance under reward transformations: Theory and application to reward shaping. *ICML* **1999**, *99*, 278–287.
44. Hassan, A.; Narayanan, A.; Zhang, A.; Ye, W.; Zhu, R.; Jin, S.; Carpenter, J.; Mao, Z.M.; Qian, F.; Zhang, Z.L. Vivisecting mobility management in 5G cellular networks. In *Proceedings of the ACM SIGCOMM 2022 Conference, Amsterdam, The Netherlands, 22–26 August 2022*; pp. 86–100.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.