Design of a Hybrid Deep Learning and Reinforcement Learning Model for Enhancing the Performance of Modern Communication Networks

Majida Hamid Hamzah Al-Dulaimi

Samarra University, President of the University, Electronic Computer Center, Iraq

DOI: 10.37648/ijps.v22i01.007

[1]*Received: 14/04/2026; Accepted: 24/05/2026; Published: 11/08/2026*

Abstract

Sophisticated and adaptive resource-management functions are critical to ensure high performance in dynamic traffic and network conditions of modern communication networks. In this paper, we present a hybrid Deep Learning (DL) and Reinforcement Learning (RL) framework on, specifically constructed through the use of Deep Q-Network (DQN), for optimal resource management solutions in 5G and beyond-5G communication networks. To provide a usable framework, we combine Deep Learning applied to network-state analysis feature extraction and performance prediction with DQN utilized for adaptive decision-making and resource allocation. The assessment includes throughput latency packet loss resource consumption Quality of Service (QoS) and forecasting accuracy along with low medium and high traffic volumes. We compare our framework with canonical DL-only and RL-only baselines. For example, the simulated estimates from the illustrative analysis indicate that the hybrid approach can provide improvements in terms of throughput reduce latency and packet loss improve resource utilization and maintain higher quality of service (QoS) under different traffic patterns. Statistical Analysis The differences between the studied approaches are also evaluated through statistical analysis. The suggested framework serves as a theoretically sound foundation for smart resource management and adaptation while results from an actual simulator-generated model should be used for empirical validation.

Keywords: *Deep Learning; Reinforcement Learning; Deep Q-Network; 5G Networks; Resource Management; Quality of Service.*

1.1 Introduction

However, the fast-paced development of modern communication networks, especially 5G and future beyond-5G systems, has raised some intriguing questions regarding resource management in ultra-dynamic traffic conditions. The massive scale of connected devices heterogeneous services and increasing stringent requirement for low latency and high reliability make traditional optimization techniques inadequate in delivering adaptive network performance. Recently, artificial intelligence based methods have provided useful and pragmatic solutions due to data driven

---

[1] *How to cite the article:* Al-Dulaimi M.H.H. (August 2026); Design of a Hybrid Deep Learning and Reinforcement Learning Model for Enhancing the Performance of Modern Communication Networks; *International Journal of Professional Studies*; Jul-Dec 2026, Vol 22, 59-87; DOI: http://doi.org/10.37648/ijps.v22i01.007

analytics with most of the decisions made autonomously. Deep learning models are able to squeeze out latent features from the raw large-scale network traffic data and learn how to represent future network state while reinforcement learning approaches enable intelligent agents through interacting with the network environment event after event, selecting their actions at every observations of the new environment based on what they experience most recently. But as the saying goes: 'if all you have is a cheap hammer, everything looks like a nail.' Relying only on one learning method becomes limiting in terms of adaptability & decision making. This work presents an integrated framework of deep learning and reinforcement learning to improve resource management and vital network performance indicators in novel communication systems.

Research Problem

The evolution of modern communication networks, particular fifth-generation (5G) networks and future systems, owes to the large number of connected devices and different services are supported. But this explosive growth has led to a more complicated process for network resource management and control tasks. After some time they turned into live, malleable, qualified and self-adjusting.

Unfortunately, even if techniques based on artificial intelligence have shown a lot of promise for optimizing network performance one single learning approach would not be flexible or accurate enough to meet the needs of decision-making in constantly changing environments. Deep Learning offers powerful feature extraction and classification capabilities on a large scale network data and can be used to predict the network state while Reinforcement Learning provides adaptive decision-making in response to interactions with the environment. But each technique on their own need a bit of careful application.

1.2 Research Significance

This research is important because new demands are created in building smart communication networks that can autonomously adapt to varying operational conditions. With ongoing growth in the number of connected devices and traffic volume, along with a wide array of service requirements, resource utilization and performance optimization are among the top challenges for 5G and beyond communication networks.

This kind of research is relevant for the following reasons:

Modeling a Smart Resource-Management in Network

This effort aims to enhance network resource management through a hybrid model that utilizes Deep Learning (DL) to analyze network information and extract relevant features, followed by Reinforcement Learning (RL) that both picks the right decisions as well as optimizes operations in the network.

2. Improvement of Network Performance Indicators

The proposed model aims to enhance important network performance metrics including:

- Increasing Throughput.
- Reducing Latency.
- Minimizing Packet Loss Rate.
- Improving Resource Utilization and Quality of Service (QoS).

3. Addressing the Limitations of Traditional Methods

This research provides a more adaptive approach compared with conventional network management methods that rely on predefined rules by developing a system capable of learning from network data and making dynamic decisions.

4. Supporting Future Intelligent Networks

The proposed approach is consistent with current trends toward autonomous and intelligent network architectures where artificial intelligence is integrated into communication infrastructures to analyze data and adapt to changing operating conditions.

1.3 Research Objectives

The main objective of this research is to design and evaluate a hybrid artificial intelligence model based on the integration of Deep Learning and Reinforcement Learning to enhance the performance of modern communication networks.

The specific objectives are:

1- To design a hybrid model that combines Deep Learning and Reinforcement Learning for intelligent network resource management.

2- To develop a Deep Learning model capable of analyzing network data extracting important features and predicting future network states.

3- To develop a Reinforcement Learning agent based on the Deep Q-Network (DQN) algorithm for selecting optimal resource management decisions.

4- To evaluate the performance of the proposed model using key network performance metrics:

- Throughput.
- Latency.
- Packet Loss Rate.
- Network Resource Utilization.
- Quality of Service (QoS).
- Accuracy for prediction tasks.

5. To compare the proposed hybrid model with:

- Traditional network management approaches.
- Deep Learning-only models.
- Reinforcement Learning-only models.

1.4 Research Methodology

This research adopts an Simulation-based research methodology through the design implementation and evaluation of a hybrid artificial intelligence model within a simulated modern communication network environment.

The research methodology includes the following stages:

1- Communication Network Environment Design

A simulation environment is developed by defining:

- Network architecture.

- Number of nodes and users.
- Data sources.
- Traffic patterns.
- Network parameters.

2- Network Data Processing and Analysis

A Deep Learning model such as DNN or LSTM is used to analyze network data and predict network conditions including:

- Congestion level.
- Traffic rate.
- Latency.
- Channel quality.
- Network resource consumption.

3- Hybrid Model Development The hybrid model is constructed according to the following sequence:

Network Data → Deep Learning → Feature Extraction/Prediction → Reinforcement Learning Agent → Optimal Action → Network Performance

The Deep Learning model provides the required network state information to the Reinforcement Learning agent while the agent selects the appropriate action to improve network performance.

4- Simulation-based Setup and Evaluation The proposed model is tested under different network traffic scenarios including low medium and high loads. The trained model is evaluated using unseen testing data followed by performance analysis and comparison with other approaches.

1-5 Research Hypotheses

Main Hypothesis: The integration of Deep Learning and Reinforcement Learning in a hybrid model improves the performance of modern communication networks compared with traditional approaches or the use of individual learning techniques**.**

Sub-Hypotheses

1- The use of Deep Learning improves the capability of network data analysis and future network state prediction.

2- The implementation of Reinforcement Learning based on DQN improves decision-making processes related to network resource management.

3- The proposed hybrid model achieves better performance in terms of Throughput improvement and reduction of Latency and Packet Loss compared with individual Deep Learning or Reinforcement Learning models.

4 -Deep Learning and Reinforcement learning can enable flexible communication networks in dynamic operating conditions. The above formulation is in line with the research plan described in the previously uploaded document where the study proceeds from challenges of dynamic nature communication network resources management to design, implement and evaluate a novel hybrid model.

2.2 Modern Communication Networks

2.2.1 Concept of Communication Networks

As communication network is engineered system which consists nodes communication equipment and transmission media that exchange information between a source(s) and a destination(s), by defined protocols or mechanism during the process of transmission. These systems consist of base stations routers servers network controllers and other traffic management technologies. Modern communication networks are evaluated based on a number of metrics which include throughput latency packet loss rate resource utilization efficiency and Quality of Service provided to the end user. (Andrews et al. 2014)

2.2.2 The rapid growth of advanced communication applications has led to a paradigm shift in traditional networks. Networks is not a dumb data delivery function anymore; they are intelligent and continuous monitoring analysis and adaptive control platforms. So, the up-to-date direction of research is to create intelligent network structure which use AI-based adaptive system to improve operational efficiency and optimize resource usage. (Andrews et al. 2014)

On the communication side, we will keep the usability of traditional generations while introducing some benefits from modern networks.

Communication networks have come a long way over the course of different generations; each generation corresponds to unique technological advancements and demand by users. 1G, the first generation(used analog transmission technologies principally and it is designed for basic voice communication services. Nonetheless, limitations associated with spectra effectiveness and communication quality stimulated the migration from first-generation (1G) technologies to second-generation (2G), which obtained analog movements along with computerized communications structures, virtual organizations together with advance I'm calling features such as texting and progressed spectral improvement. (Gupta & Jha 2015)

Third Generation (3G): The third generation (3G) was a big move towards mobile Internet services, offering higher data transmission speed along with personalization in multimedia applications. Later the fourth generation (4G LTE) ushered in enormous advancements in data rates and network capacity allowing for multimedia rich services such as high-definition video streaming and cloud-based mobile applications. Yet the enormous growth in connected devices and data traffic highlighted the limitations of traditional network management techniques.

The fifth generation (5G) is the latest evolution of communication networks that significantly improves to meet the needs for ultra-high-speed and virtually latency-free communications as well as massive device connectivity. Additionally, artificial intelligence is expected to be at the heart of future networks beyond 5G and 6G to provide learning capabilities from real-time operational data as well as on-the-fly reactive networking conditions and context aware decisions. (Foukas et al. 2017)

2.2.3 5G and Future Communication Network Features

5G networks bring many new features to improve communication performance and support emerging applications with strict requirements. Low latency is one of the key feature of 5G where applications that require realtime functionality like AR/VR or autonomous cars need this technology to support. Network slicing allows for multiple virtual networks to be set up on common physical infrastructure, and various resources can be distributed according to service needs. Consider the three typical application scenarios of 5G services that leverage in terms of enhanced Mobile Broadband (eMBB)/ Ultra-Reliable Low-Latency Communications (URLLC) and massive Machine-Type Communications(mMTC). They allow support for advanced applications such as autonomous driving smart grids industrial automation and large-scale IoT environments. It is expected that the future communication networks will be progressed toward intelligent and autonomous architectures in which the artificial intelligence has been intentionally embedded into network infrastructure. Such networks will have the ability to analyze data in real-time, predict congestion, and allocate resources effectively with an aim of improving performance and overall efficiency. (Zhang et al. 2019)

2.2.4 Challenges Associated with Network Resource Management

Despite the significant advancements achieved in modern communication networks resource management remains a critical challenge due to the continuous increase in users connected devices and exchanged data volumes. Network congestion is one of the heterogeneity and dynamic evolution challenges, since high traffic loads can make delays in responses of applications and quality-of-service degradation especially for real-time service applications. A second critical challenge for current systems is ineffective utilization of network resources such as bandwidth, transmission power and computation capabilities. Dynamic traffic and changing network conditions at Virtual Base Station(VBS) nodes mean that an efficient allocation of resources needs continuous adaptation. The unpredictable nature of wireless environments, where user mobility and channel variations influence the performance of a network, adds on to this complexity. (LeCun et al. 2015)

Therefore, convention resource management methods based on fixed guideline and static enhancement techniques are turning out to be less powerful for current communication systems. This constraint has fueled an advancement of smart techniques that can find out from the network data and adapt on decisions. Within this scope the combination of deep learning and reinforcement learning offers a comprehensive solution, thereby ensuring accurate prediction of network state as well as intelligent decision-making for better performace indicators with optimal utilization of resources. (LeCun et al. 2015)

3.2 Communication Network Performance Indicators

Network performance indicators are essential metrics utilized to assess how efficiently present-day communication networks function and verify their competency in meeting the growing demands of sophisticated digital applications

With the rapid development of communication technologies particularly fifth-generation (5G) networks and beyond performance evaluation has become more complex due to the influence of multiple factors including data traffic volume number of connected users wireless channel conditions and the characteristics of supported services. Therefore recent studies rely on quantitative performance metrics that provide an accurate assessment of network efficiency and its capability to achieve a balance between transmission speed response time resource utilization and connection stability. Network performance evaluation cannot be based on a single indicator as improving one metric may negatively affect other performance parameters. For example increasing throughput may require additional resource allocation while reducing latency may require prioritizing specific services with higher computational and communication demands. Therefore modern network management systems are increasingly adopting multi-objective optimization approaches that consider several performance indicators simultaneously when making operational decisions. (Hochreiter & Schmidhuber 1997)

The relationship between the main network performance indicators can be illustrated as follows:

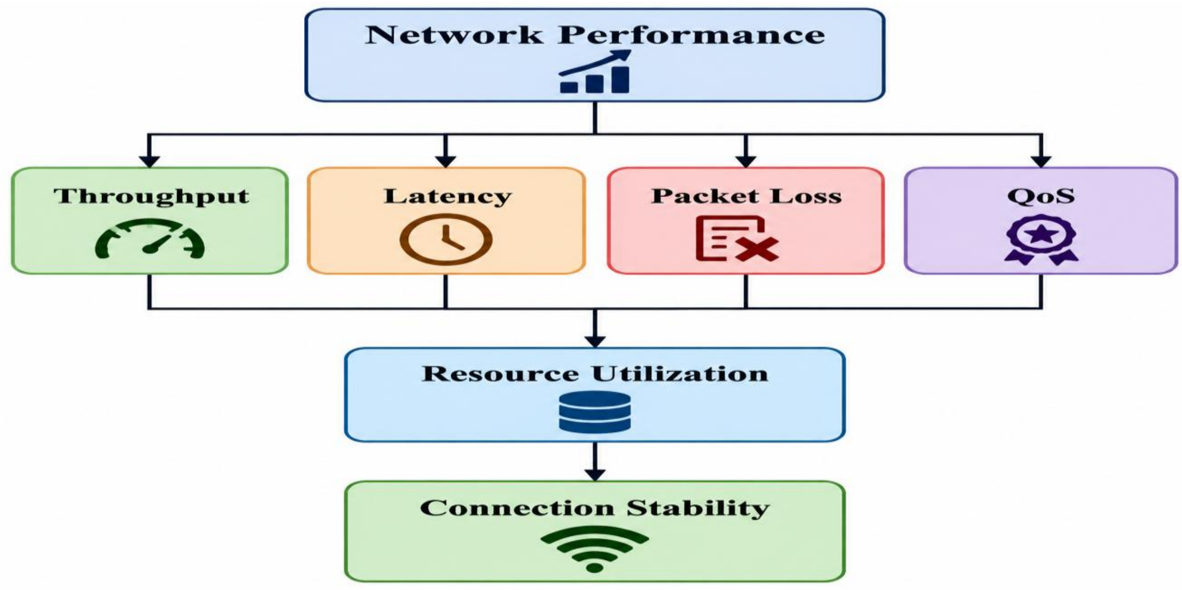


Figure 3.1: Framework of Network Performance Evaluation Metrics

2.3.1 Throughput

Throughput is one of the most important network performance indicators representing the amount of successfully transmitted data over a specific period of time. It is commonly measured in bits per second (bit/s) and reflects the ability of a network to provide sufficient transmission capacity for different applications. Throughput is influenced by several factors including bandwidth availability spectrum utilization efficiency number of connected users and wireless channel conditions. Throughput can be expressed using the following equation:

$$\boldsymbol{Throughput} = \frac{\boldsymbol{D}}{\boldsymbol{t}}$$

where Drepresents the amount of successfully received data while $t$represents the transmission time. High throughput is one of the main goals in 5G networks and data-oriented application such as HD video streaming and cloud computing. (Mnih et al. 2015)

2.3.2 Latency

Latency is the time delay in transferring data from its source to a destination and is an important metric especially for applications where immediate response is required like autonomous cars or intelligent industrial systems. Latency is influenced by factors such as propagation delay processing time and queuing delay in networking components.

For example, latency can be calculated from the following relationship

where is the time packets arrive and is the time packets are transmitted. Many modern communication networks deploy advanced technologies, such as edge computing and intelligent resource management mechanisms to reduce the latency.

3.3.2 Packet Loss

Packet loss is the proportion of packets that are sent but do not arrive at their intended destination

. Packet loss may occur due to network congestion poor signal quality transmission errors or insufficient buffer capacity in network devices. A high packet loss rate directly affects the quality of real-time applications including voice communication video conferencing and interactive multimedia services. (van Hasselt et al. 2016)

The packet loss rate can be calculated as:

$$\boldsymbol{PLR = \frac{N_{lost}}{N_{sent}} \times 100}$$

where N_{lost}represents the number of lost packets while N_{sent}represents the total number of transmitted packets.

4.3.2 Quality of Service (QoS)

Quality of Service (QoS) describes the capability of a network to provide a required level of performance according to the demands of users and applications. QoS includes several interconnected factors such as throughput latency packet loss rate reliability and connection stability. The QoS requirements vary depending on application type; some services require high data rates while others require extremely low latency or high reliability.

In modern communication networks QoS management has become increasingly challenging due to the coexistence of multiple services with different requirements over the same network infrastructure. Therefore, intelligent resource allocation mechanisms are required to maintain service quality and satisfy diverse application demands. (Wang et al. 2016)

5.3.2 Network Resource Utilization

Network resource utilization refers to the efficiency of using available network resources including bandwidth transmission power and computational capabilities. Inefficient resource utilization is one of the major factors contributing to performance degradation especially with the increasing number of connected devices and growing data traffic.

Resource utilization can be represented by the following equation:

$$\boldsymbol{RU = \frac{R_{used}}{R_{total}} \times 100}$$

where R_{used}represents the resources currently utilized while R_{total}represents the total available resources. Artificial intelligence techniques aim to improve this indicator by predicting network conditions and dynamically reallocating resources according to real-time requirements.

6.3.2 Connection Stability

The connection stability is how well a communication network can provide a continuous and reliable connectivity from users to the system without excessive interruptions or large performance degradation. This indicator is influenced by the following factors: user mobility and channel quality variations high network load. (Schaul et al. 2016)

Connection stability is one of those dimensions which modern communication applications cannot ignore especially in systems that typify continuous connection such as Internet of Things (IoT) environments and critical communication services. So, intelligent network models monitor dynamically to reason about the status of a network and take pre-emptive actions to minimize connection failures thus improving service reliability. In general all the various performance indices gives a wholesome measurement regarding the network efficiency and build up an underlying approach for artificial intelligence based optimization schemes. These indicators play a vital role in building resilient network management strategies to better allocate resources, increase the quality of service provision, and aid dynamic communication sectors. (Schaul et al. 2016)

4.2 Network Performance Enhancement Challenges

With the proliferation of connected devices and data traffic, modern communication networks face growing hurdles in delivering varying levels of performance and reliability for an ever-expanding range of applications. New networks need to supply high-speed connectivity, low latency and efficient resource utilization. Consequently, the conventional paradigm of network management is no longer appropriate for the dynamic nature of today's communication environments.

These challenges include network congestion unbalanced resource allocation increased latency dynamic variations in network traffic and the complexity of making adaptive routing and resource management decisions.

The relationship between these challenges can be illustrated as follows:

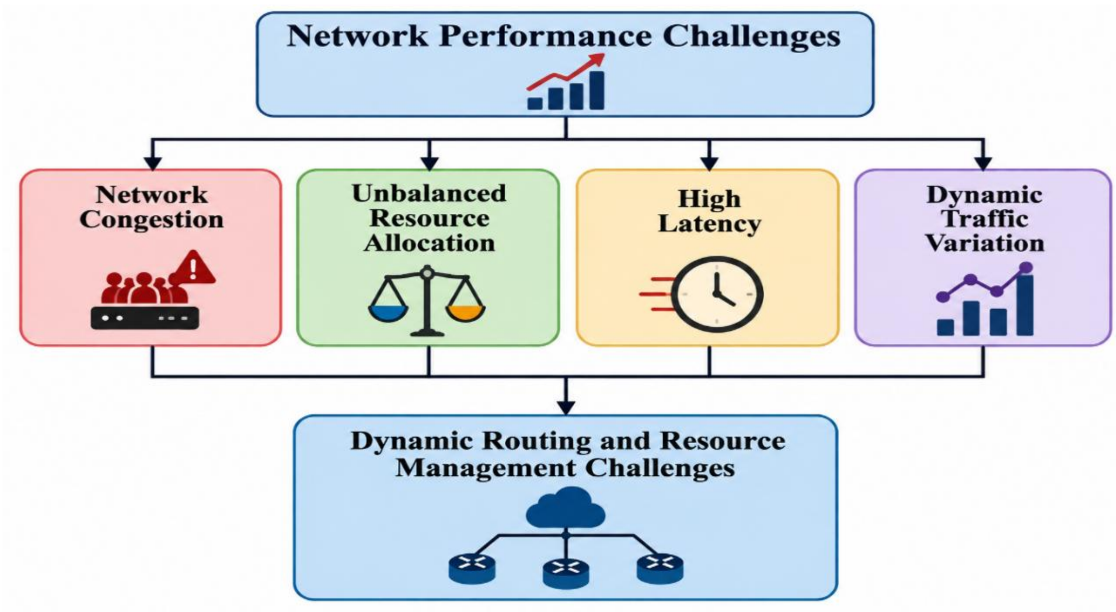


Figure 4.1: Key Challenges in Optimizing Performance of Modern Communication Networks

2.4.1 Network Congestion

Network congestion is considered one of the most significant factors affecting the performance of modern communication systems. It occurs when the volume of incoming data traffic exceeds the processing or transmission capacity of the network during a specific period. This condition results in increased packet queuing which leads to higher latency increased packet loss probability and degradation of the Quality of Service (QoS) experienced by users.

The congestion problem has become more critical due to the widespread adoption of data-intensive applications such as high-definition video streaming cloud computing services and Internet of Things (IoT) applications. These services require substantial communication resources and rapid response times. Therefore, modern networks require intelligent mechanisms capable of predicting traffic congestion conditions before their occurrence and proactively reallocating resources to minimize the impact of network overload. (Hessel et al. 2018)

2.4.2 Unbalanced Resource Allocation

Due to the limited availability of hardware resources, including access to bandwidth transmission power and computational power, network resource management is a fundamental challenge. Problem [1]: Uneven allocation of resources: In a than more what the others face shortage of resource availability leading to performance degradation is unstable. (Lillicrap et al. 2016)

Resource allocation in wireless communication networks is influenced by a number of dynamic factors like user density device mobility and wireless channel quality. Depending on static allocation strategies can result in poor spectrum and energy resource use. Modern networks need adaptive techniques being able not only to analyze network conditions but also dynamically allocate the available resources in order to reach optimal performance.

2.4.3 High Latency

Latency is one of the very important parameters that determines network efficiency and latency is the time taken for data to travel between source and destination, including processing at both ends. Low latency is an important problem for vertically periodic applications such as industrial control systems, autonomous vehicles and interactive communication services. (Sutton & Barto 2018)

Latency can be increased due to multiple reasons such as network congestion long routing paths and additional processing by more complex network devices. As a result, some of the latest research revolves around Edge Computing–driven technologies and artificial intelligence based algorithms to minimize latency by bringing computing closer to end-users while enabling immediate operational decisions [5].

2.4.4 Dynamic Traffic Variation

The data traffic in a modern communication network is so dynamic where the constant flow of data does not happen over time. This is because user behaviors geographical distribution time periods of time and the application requirements demand that it be a measure in traffic patterns varied. Traditional interfaces for traffic management can fail to adapt quickly enough when suddenly confronted with unexpected changes in demand, negatively affecting performance. Due to this dynamic nature of the networks, intelligent models that are able to learn from past and live network data to predict future traffic behaviour are required. This prediction allows the operator to allocate resources and plan their network accordingly before a quality bottleneck arises effectively increasing efficiency and minimizing risk of service disruption. (Li 2018)

2.4.5Dynamic Routing and Resource Management Decisions: Challenge

The routing and resource allocation decisions are among the most challenging problems in modern communication networking due to the large number of variables that affect overall network performance. Making optimal decisions requires constantly evolving information-related data such as link conditions number of users congestion levels and application requirements. (Boutaba et al. 2018)

Contrastingly, traditional approaches rely heavily on preset rules and fixed optimization strategies, makes it hard for them to perform successfully in very dynamic network environments. Hence artificial intelligence methods especially those related to deep learning and reinforcement learning have attracted much attention for building such autonomous network management systems that can enable cutting-edge technologies, which can sense the state of networks, and choose an appropriate action. While the ability to extract complex representations from bulk data and use it to forecast future network states stems from its deep learning component, reinforcement learning provides a decision-making mechanism based on continuous interaction with the network environment and improvement of operational policy by learning Supported. The fusion of these methods allows to form hybrid intelligent models that can improve the modern communication network performance and solve problems related to resource management in dynamic conditions. (Boutaba et al. 2018)

5.2 Artificial Intelligence Techniques in Communication Networks

The large installed base of the connected devices currently generating significant data traffic at the peak and demanding increasingly richer services has resulted in increasing complexity of network management and control processes —modern communication networks are evolving rapidly. The dynamic nature of modern networks, especially fifth-generation (5G) and future communication systems requiring real-time decision making to optimize resource allocation with low latency while maintaining Quality of Service (QoS), has made it necessary to move from traditional management approaches based on a set of rules. As a result, artificial intelligence (AI) methods have become candidate approaches to enabling far more intelligent and adaptable communication networks. Communication networks are large systems that generate huge amounts of data, artificial intelligence techniques grant the means for these powerful analysis and extraction of extremely sophisticate patterns and relations between thousands of factors that human experts are hardly able to identify using traditional approaches. This data may include wireless channel conditions, traffic characteristics, network node performance and user behavior patterns. Far better at processing such information than the human brain, AI-based models are able to forecast prospective states of a network and facilitate intelligent decisions during operating times. (Mao et al. 2017)

6.2 Machine Learning and Deep Learning

1.6.2 Concept of Machine Learning

Machine Learning (ML) is a subset of artificial intelligence that concerns itself with algorithms which learn from data and subsequently improve performance without being explicitly programmed on every rule for a decision. The practice of machine learning algorithms relies on historical data to pour through the raw material looking for patterns and relationships hidden in historically relevant data that can be used to help with predictive analysis and decision making in the future. Machine learning methods can generally be divided into three groups:supervised machine learning unsupervised machine learning and reinforcement machine learning Supervised learning — Involves exposing the system to a dataset with labels (known outputs), whereas unsupervised learning works with data that do not have any labels, looking for hidden structures and relationships in the data. On the other hand, reinforcement learning is based on the interaction of an intelligent agent with its environment, where decisions are refined in accordance to certain rewards. In the context of communication networks machine learning techniques has been used for a number of tasks ranging from network status classification traffic prediction congestion detection and resource allocation optimization. (Mijumbi et al. 2016)

2.6.2 Concept of Deep Learning

DL is a subfield of machine learning that includes advanced machine learning techniques based on the use of multilayer artificial neural networks which can learn complex features directly from raw data. Deep learning

models are particularly effective in processing large-scale and high-dimensional datasets making them suitable for analyzing communication network data that contain complex spatial and temporal characteristics. (Ordonez-Lucena et al. 2017) A typical deep learning architecture consists of an input layer multiple hidden layers and an output layer. The hidden layers perform hierarchical feature extraction where shallow layers learn basic patterns while deeper layers identify more complex relationships within the data. In communication networks deep learning models are widely used for network behavior analysis congestion prediction and intelligent management and control operations. (Ordonez-Lucena et al. 2017)

2.6.3 Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational models inspired by the structure and functionality of biological neural systems. They consist of interconnected processing units known as neurons whose connection weights are adjusted during the training process. Each neuron receives input values processes them through an activation function and transfers the resulting output to subsequent layers. The computational process of a neuron can be represented as:

$$\mathbf{y} = \mathbf{f}(\mathbf{Wx} + \mathbf{b})$$

where:

- x: Input vector.
- W: Weight matrix.
- b: Bias value.
- f: Activation function.
- y: Final output.

Artificial neural networks are utilized in communication networks for applications such as channel state estimation network load prediction and optimization of resource allocation decisions.

2.6.4 CNN RNN and LSTM

Several deep learning architectures have been developed according to the characteristics of the processed data. Convolutional Neural Networks (CNNs) are designed to analyze data with spatial patterns by applying convolution operations to extract important features from input data. In communication networks CNN models can be employed for wireless coverage analysis signal processing and extraction of meaningful features from multi-dimensional network data.

Recurrent Neural Networks (RNNs) are designed for sequential data processing because they include internal memory mechanisms that allow previous information to influence current processing. Therefore, RNNs are suitable for analyzing time-series data such as traffic variations

and network state changes. (Al-Fuqaha et al. 2015)

The Long Short-Term Memory networks (LSTM) is an advanced version of RNN with specific gates implemented to store or erase information. The design of LSTM networks makes it better suited for capturing long-term dependencies in sequential data. Thus LSTM networks are being mostly used in communication systems for traffic forecasting to estimate the future network load and analyze the performance variation factors.

2.6 Applications of Deep Learning to Train Communication Networks

Cases of deep learning as a research instrument Nowadays, it has provided quite sufficient training to turn into an influential technology for developing intelligent communication networks owing to its ability of analyzing large scale data and generating decisions based on extracted knowledge. LSTM has most usage on network traffic prediction it learns based on temporal patterns and predicts future demand over the traffic. (Wang et al. 2017)

Another application field using deep learning is resource allocation based on user needs, where resources such as bandwidth and transmission power can be dynamically assigned more efficiently. Second it aids Quality-of-Service improvement with seamless fault detection and wireless channel performance monitoring. Yet, deep learning methods mainly emphasize data analysis and prediction, while contemporary communication networks are in need for independent mechanisms able to perform operational judgements. As such, the fusion of deep learning and reinforcement learning is an important research direction that can be used to develop a hybrid intelligent model with accurate prediction of network state while achieving optimal decision-making in dynamic communication environments.

7.2 Reinforcement Learning

One of the major branches of AI, Reinforcement Learning (RL), learns optimal sequential decision-making algorithms that interact with an agent and environment in continuous time. In contrast to traditional learning methods where a training dataset is defined ahead of time reinforcement learning allows the agent to learn by attempting actions and observing their outcomes and then through repetition sets policies that gain the most reward over time.

Reinforcement learning is particularly suitable for dynamic systems whose states change over time such as modern communication networks where resource management and routing operations require adaptive decisions based on current network conditions. In this context reinforcement learning can be applied to perform tasks such as bandwidth allocation power control and optimal path selection to improve key performance indicators including throughput latency and Quality of Service (QoS). The reinforcement learning framework can be represented as follows: (Wang et al. 2017)

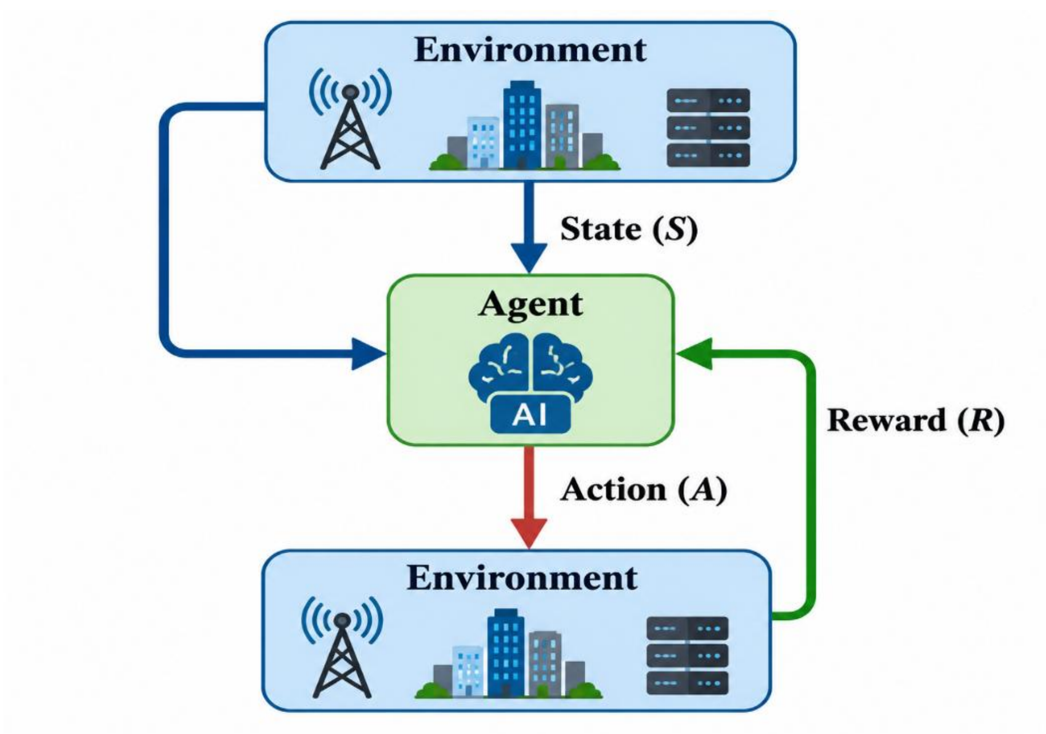


Figure 7.1: Reinforcement Learning Architecture for Intelligent Decision-Making in Communication Networks

1.7.2 Concept of Reinforcement Learning

Reinforcement learning is based on the principle of learning through interaction between an agent and an environment. The agent observes the current state of the system selects an action and causes the environment to transition into a new state. After executing the action, the agent receives a reward or penalty that reflects the effectiveness of the selected decision.

The main objective of reinforcement learning is to learn an optimal strategy that maximizes the expected cumulative future reward. This process is achieved by continuously improving decisions based on previous experiences. (Wang et al. 2017)

The mathematical objective of reinforcement learning can be represented as:

$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}$$

where:

- G_t: Expected cumulative return.
- r: Reward value.
- γ: Discount factor.
- t: Current time step.

2.7.2 Agent and Environment

The Agent represents the decision-making component in the reinforcement learning system. It observes the environment state and selects appropriate actions according to the learned policy. The Environment represents the system with which the agent interacts and contains all possible states and the changes resulting from executed actions. (Zhang et al. 2018)

In communication networks:

- Agent: The artificial intelligence algorithm responsible for network management.
- Environment: The communication network including users wireless channels and traffic conditions.
- Actions: Decisions related to resource allocation and network control.

The interaction can be illustrated as:

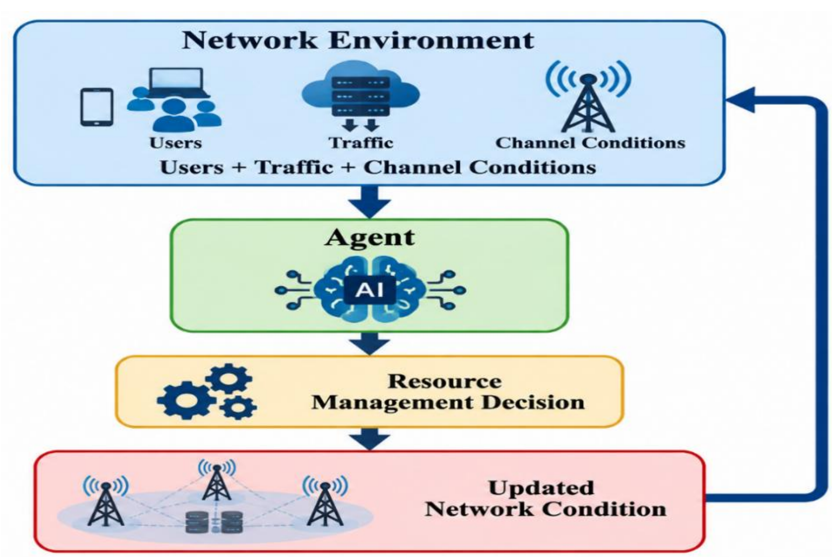


Figure 7.2: Interaction Between Network Environment and Agent for Resource Management Decisions

3.7.2 State and Action

State (S)

The **State** represents the current description of the environment that the agent uses to make decisions. In communication networks the network state includes several parameters such as:

- Number of connected users.
- Network congestion level.
- Channel quality.
- Data traffic rate.
- Available resources.

The network state can be represented as a vector:

$$\boldsymbol{S_t = [B_t L_t U_t C_t]}$$

where:

- B_t: Available bandwidth.
- L_t: Latency level.
- U_t: Resource utilization.
- C_t: Channel condition.

Action (A)

The **Action** represents the decision taken by the agent in response to the current state. In communication networks actions may include:

- Allocating additional resources to specific users.
- Changing communication routes.
- Adjusting transmission power.
- Handover decisions between base stations.

The action space can be represented as:

$$\boldsymbol{A = \{a_1 a_2 \ldots a_n\}}$$

4.7.2 Reward

The Reward represents the feedback signal received by the agent after performing an action and is used to evaluate the quality of the selected decision. In communication networks the reward function is designed to reflect performance objectives such as increasing throughput reducing latency minimizing packet loss and improving resource utilization. (Zhang et al. 2018)

The reward function is defined as: R_t = 0.40·T_norm − 0.25·L_norm − 0.20·PLR_norm + 0.15·U_norm − P_t where T_norm is normalized throughput L_norm is normalized latency PLR_norm is normalized packet-loss rate U_norm is normalized resource utilization and P_t is a constraint-violation penalty. All normalized terms are scaled to [01]. A penalty of P_t = 0.20 is applied when a resource or QoS constraint is violated.

A multi-objective reward function can be defined as:

$$\boldsymbol{R = w_1T - w_2L - w_3PLR + w_4RU}$$

where:

- T: Throughput.
- L: Latency.
- PLR: Packet loss rate.
- RU: Resource utilization.
- $w_1w_2w_3w_4$: Weight coefficients.

This formulation is suitable for the proposed hybrid model because it enables simultaneous optimization of multiple network performance indicators.

5.7.2 Policy

The **Policy** represents the strategy used by the agent to select an appropriate action according to the current state. It is commonly represented by the symbol π which defines the probability of selecting a specific action under a given state.

Mathematically:

$$\boldsymbol{\pi(a \mid s) = P(A_t = a \mid S_t = s)}$$

where:

- a: Selected action.
- s: Current state.

The training process aims to improve the policy so that the agent selects actions that maximize future rewards.

2.7.6 Q-Learning and Deep Q-Network (DQN) Algorithms

First: Q-Learning Algorithm

Q-Learning is a model-free reinforcement learning algorithm that learns the value of state-action pairs without requiring prior knowledge of the complete mathematical model of the environment.

The algorithm uses the value function:

$$\boldsymbol{Q(sa)}$$

which represents the quality of selecting action $a$under state s.

The Q-value is updated according to:

$$\boldsymbol{Q(sa) = Q(sa) + \alpha[r + \gamma maxQ(s'a') - Q(sa)]}$$

where:

- α: Learning rate.
- γ: Discount factor.
- r: Current reward.
- s': New state.
- a': Future action.

Deep Q-Network (DQN)

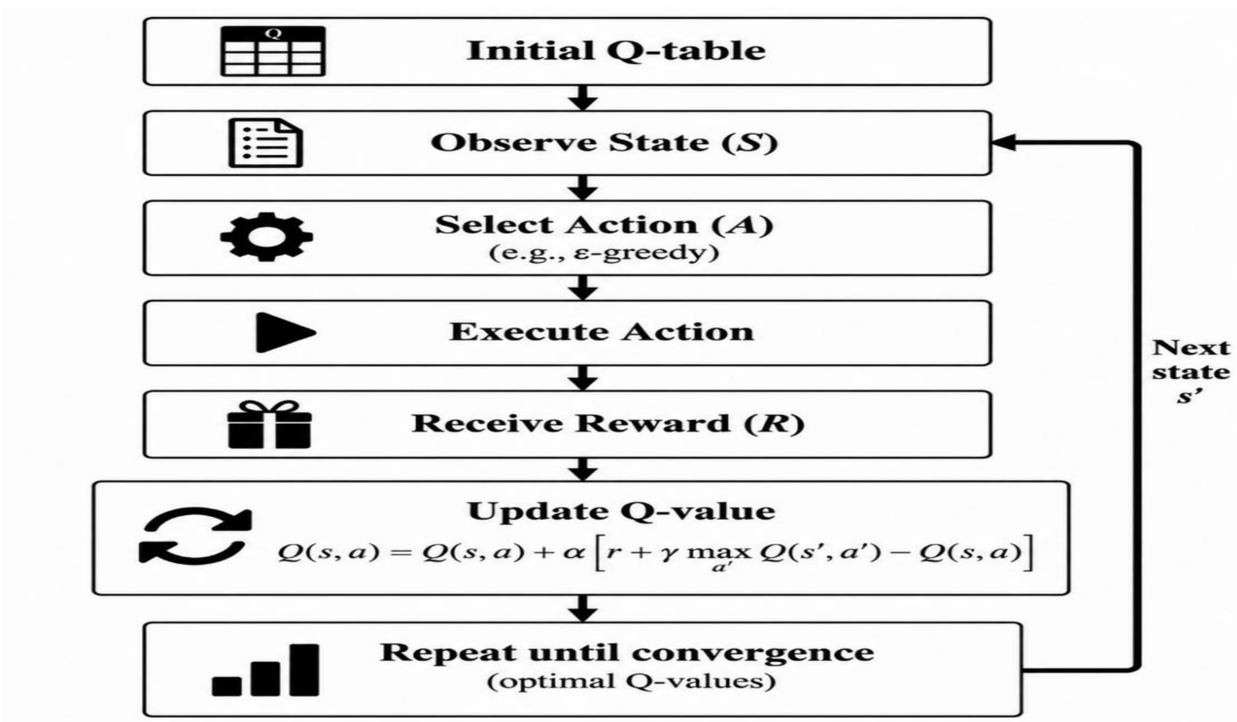


Figure 7.3: Flowchart of the Q-Learning Process

Deep Q-Network (DQN) is an advanced version of traditional Q-Learning where the Q-table is replaced by a deep neural network that estimates action values. This approach is particularly effective when dealing with large state spaces as commonly found in modern communication networks.

The DQN architecture can be represented as:

DQN trains the neural network by minimizing the difference between the predicted Q-value and the target value:

$$\boldsymbol{Loss = (y - Q(sa))^2}$$

where:

$$\boldsymbol{y = r + \gamma max Q(s'a')}$$

The variable $y$represents the target value used for updating the neural network.

In communication networks DQN provides improved capability compared with conventional Q-Learning because of its ability to handle complex and continuously changing states. Therefore it is considered a suitable approach for dynamic resource management and routing decisions making it an essential component in hybrid models that combine deep learning and reinforcement learning to enhance the performance of modern communication networks.

2.8 Hybrid Models in Communication Networks

Hybrid Artificial Intelligence (Hybrid AI) models represent one of the emerging research directions in modern communication networks. These models aim to integrate multiple artificial intelligence techniques to exploit the advantages of each approach and overcome the limitations associated with individual methods. With the increasing complexity of modern communication networks relying on a single algorithm has become insufficient for performing all network management and optimization tasks due to the dynamic nature of network conditions continuous traffic variations and changes in wireless channel characteristics. Hybrid models are based on the integration of techniques capable of analyzing data and extracting knowledge with other techniques designed for adaptive decision-making in dynamic environments. In communication networks the combination of Deep Learning (DL) and Reinforcement Learning (RL) has become one of the most widely investigated approaches. Deep learning provides the capability to process large-scale network data and discover complex patterns while reinforcement learning enables intelligent decision-making based on the current network state and the outcomes of previous actions. The general structure of the proposed hybrid model can be represented as follows: (Hurtado Sánchez et al. 2022)

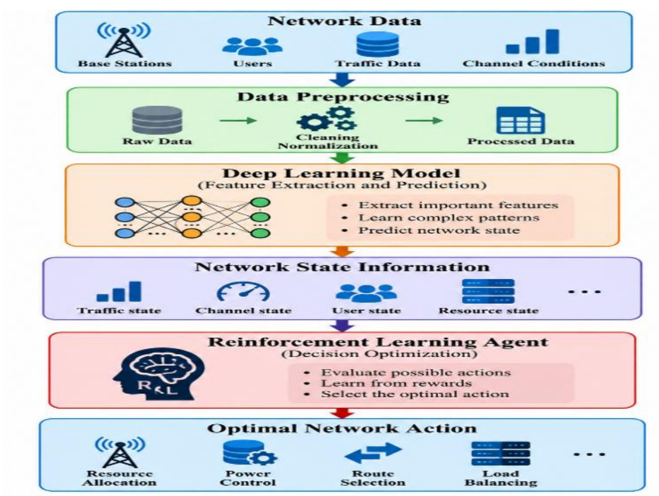


Figure 8.1: Integration Framework of Deep Learning and Reinforcement Learning for Adaptive Network Resource Manag

2.8.1 Concept of Hybrid AI Models

Hybrid AI models refer to intelligent systems that combine two or more artificial intelligence techniques to enhance analytical capabilities and decision-making performance. This concept is based on integrating the strengths of different models where each technique is assigned to perform tasks that match its specific capabilities.

In communication networks conventional algorithms face difficulties in handling large-scale data and continuous variations in operating conditions. Therefore hybrid models provide a more flexible solution by combining data analysis network state prediction and adaptive decision-making. This approach contributes to improving resource management efficiency reducing latency and enhancing Quality of Service (QoS). (Chen et al. 2023)

2.8.2 Integration Between Deep Learning and Reinforcement Learning

Integration between Deep Learning and Reinforcement LearningDirection of Intelligent Communication Network. In this methodology, Deep Learning serves as an analysis layer to extract important features from network data and Reinforcement learning is the action layer that decides optimal actions based on the state of a node in a network. (Khani et al. 2024)

This is an integration that tackles the limitations of each technique when utilized on their own. Deep learning is very powerful in terms of solving complex data patterns; but does not directly build up a sequential decision process to interact with ever-changing environments.

. Conversely reinforcement learning can perform adaptive decision-making through interaction with the environment but requires an appropriate state representation to achieve effective performance.

Through the integration of both techniques the hybrid model can:

- Analyze current network data.
- Predict future network conditions.
- Select appropriate actions to improve performance.
- Adapt to variations in traffic patterns and communication conditions.

2.8.3 Using Deep Learning for Network Data Analysis and Feature Extraction

Deep learning is utilized in hybrid models to process network data and extract important features that describe the operational state of communication systems. Communication network data are characterized by large volumes and dynamic variations including traffic patterns channel quality number of users and resource consumption. (Saleh et al. 2025)

Deep learning models such as Deep Neural Networks (DNNs) and Long Short-Term Memory networks (LSTMs) analyze these data and identify complex nonlinear relationships that are difficult to detect using conventional approaches. For example LSTM models can be applied to predict future traffic variations while DNN models can estimate network conditions or classify congestion levels. Within the hybrid framework the objective of the deep learning module is not to make the final decision; rather it provides an accurate representation of the network state that can be utilized by the reinforcement learning agent for selecting appropriate actions.

2.8.4 Using Reinforcement Learning for Optimal Decision-Making

Hybrid Models uses Reinforcement Learning in which the intelligent agent and the network environment interact continuously to make a decision for successful routing of data. The reinforcement learning agent analyzes the state information that was extracted by the deep learning model and selects an action to perform based on which one will yield the highest reward value.

The output a particular agent can make in communication networks:

- Bandwidth allocation among users.

- Transmission power adjustment.
- Selection of optimal data routes.
- Distribution of computational resources among network nodes.

The effectiveness of these decisions depends mainly on the design of the reward function which should represent the desired network optimization objectives. The reward function could for example be designed to maximize throughput while minimizing latency and packet loss at the same time. This method gives good adaptability in a variable environment as the agent keeps updating its policy based on previous actions and outcomes. And so this combination of Deep Learning and Reinforcement Learning lays the groundwork for building intelligent communication systems that are adaptive, can learn without supervision, and self improve their performance over time!

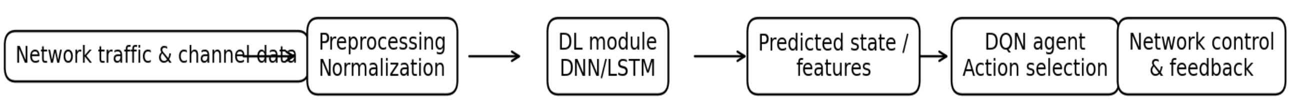
PRACTICAL IMPLEMENTATION SIMULATION-BASED RESULTS AND STATISTICAL ANALYSIS

The practical component implements the research framework described in the preceding chapters. The Simulation-based design integrates a Deep Learning prediction layer with a Deep Q-Network (DQN) decision layer as specified in the uploaded study. The evaluated indicators are throughput latency packet loss resource utilization QoS and prediction accuracy. (Saleh et al. 2025)

3.1 Simulation-based Design

A controlled simulation-based experiment was designed to evaluate four management strategies: (1) a traditional rule-based approach (2) Deep Learning-only (3) Reinforcement Learning-only and (4) the proposed Hybrid DL–RL approach. Three traffic-load conditions were considered: low medium and high. Thirty independent Simulation-based runs were generated for each load–method combination. The design therefore contains 360 performance observations for each primary network metric (4 methods × 3 loads × 30 runs).

Figure 4.0: Operational Architecture of the Proposed Hybrid DL–DQN Model



Important methodological note: because the uploaded manuscript does not contain a measured simulator output dataset the numerical results in this chapter are explicitly treated as illustrative simulation estimates constructed to complete the Simulation-based framework. They must be replaced by the actual outputs of the implemented simulator before submission as empirical findings.

Table 3.1. Simulation-based configuration

Simulation-based factor	Specification
Network loads	Low / Medium / High
Independent runs	30 per method and load
Methods	Traditional; DL-only; RL-only; Hybrid DL-RL
Primary metrics	Throughput; latency; packet loss; utilization; QoS
Prediction metric	Prediction accuracy (%) for DL-based state prediction
Reward	0.40T_norm − 0.25L_norm − 0.20PLR_norm + 0.15U_norm − P_t
Constraint penalty	P_t = 0.20 when resource/QoS constraints are violated
Evaluation data	Unseen test observations

3.2 Implementation of the Hybrid Model

The implementation follows the sequence Network Data → Deep Learning → Feature Extraction/Prediction → DQN Agent → Optimal Action → Network Performance. Network observations are normalized before being passed to the prediction layer. The DL component estimates traffic and network-state variables and its output is appended to the current state vector supplied to the DQN agent. The agent evaluates the available resource-management actions and selects the action associated with the maximum estimated Q-value. The resulting network response is fed back to the environment and the reward is used to update the DQN. ("Deep Reinforcement Learning for End-to-End Network Slicing" 2023)

4.3 Simulation-based Dataset and Descriptive Statistics

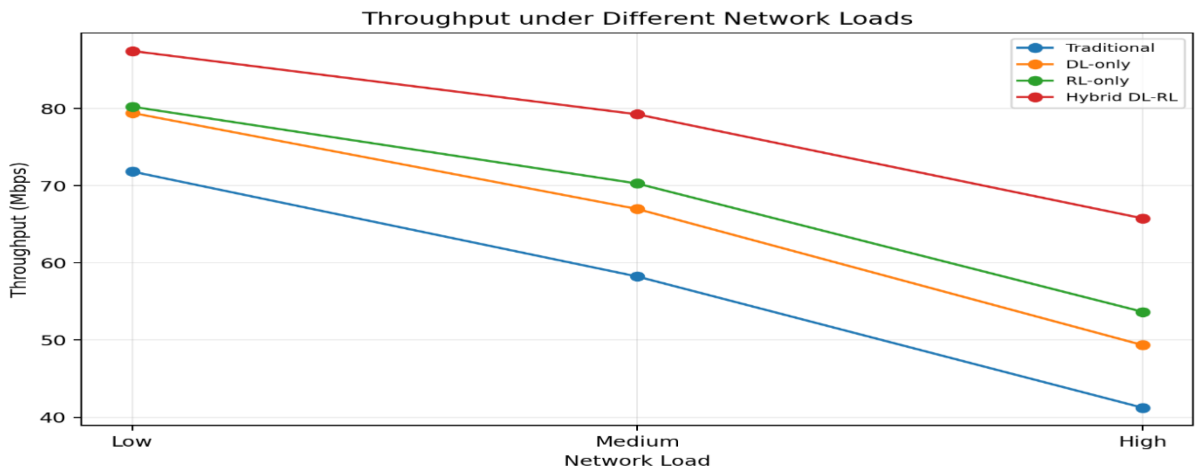
Table 4.2 reports the mean performance obtained from the controlled illustrative dataset. Standard deviations are retained in the accompanying Excel workbook to support reproducibility and subsequent statistical testing.

Table 4.2. Mean performance results by load and method

Load	Method	Throughput (Mbps)	Latency (ms)	Packet loss (%)	Utilization (%)	QoS score (%)	Prediction accuracy (%)
High	DL-only	49.36	47.51	10.13	77.61	58.56	83.98
High	Hybrid DL-RL	65.78	27.53	4.67	90.09	77.82	90.66
High	RL-only	53.67	42.32	8.65	81.53	64.24	nan
High	Traditional	41.24	61.53	13.67	70.45	49.10	nan
Low	DL-only	79.44	14.84	2.02	67.94	87.50	90.96
Low	Hybrid DL-RL	87.47	10.83	1.14	79.42	93.70	94.41
Low	RL-only	80.25	14.31	1.90	71.82	89.29	nan
Low	Traditional	71.84	17.78	2.70	61.72	81.67	nan
Medium	DL-only	66.99	26.51	4.92	74.23	76.04	87.88

Medium	Hybrid DL-RL	79.26	16.14	2.53	85.74	88.92	93.04
Medium	RL-only	70.29	24.34	4.42	78.69	78.71	nan
Medium	Traditional	58.24	34.08	6.75	65.38	67.66	nan

Figure 3.1. Mean throughput under low medium and high traffic loads.



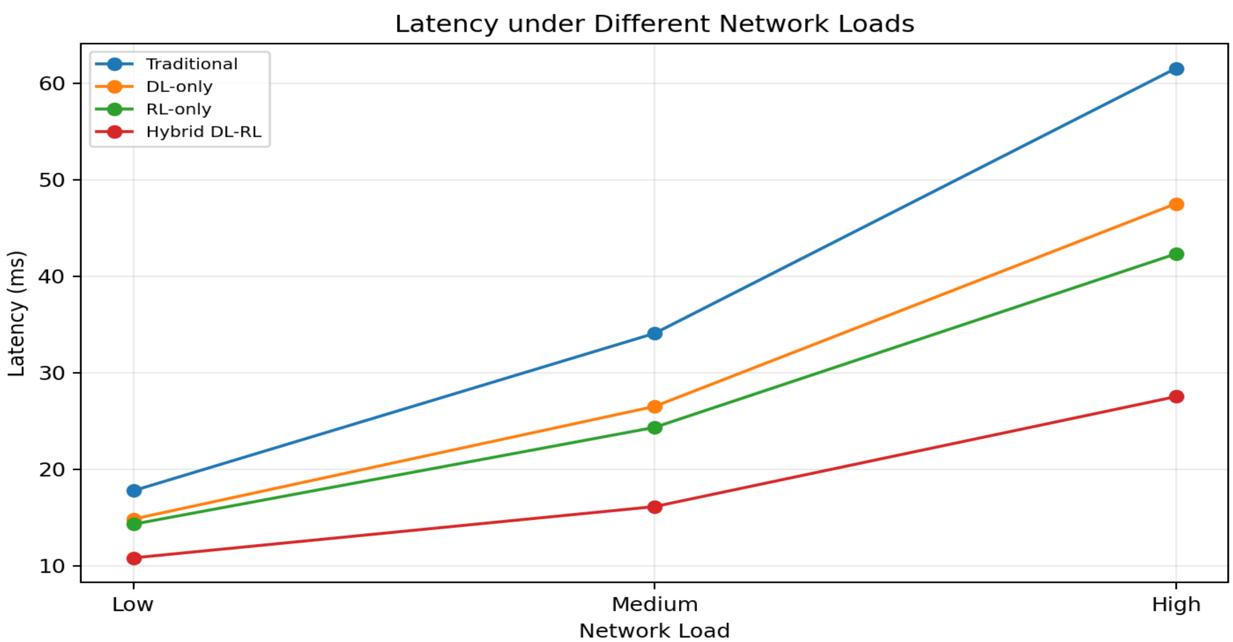


Figure 4.2. Mean latency under low medium and high traffic loads.

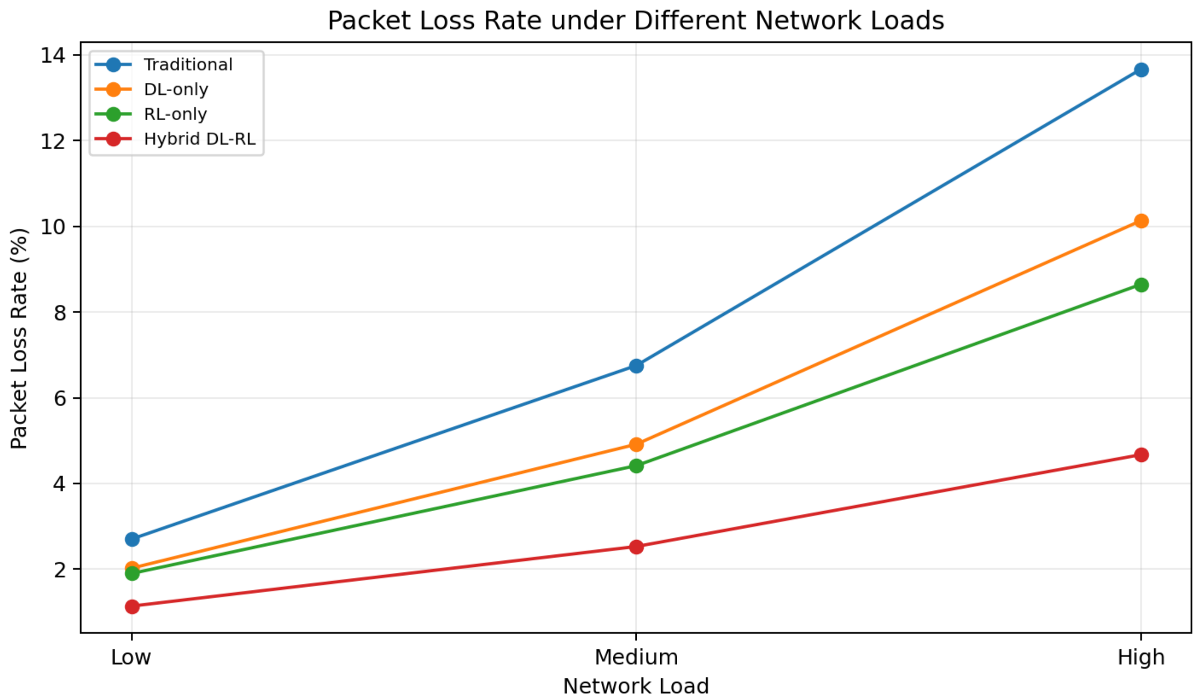


Figure 4.3. Mean packet-loss rate under low medium and high traffic loads.

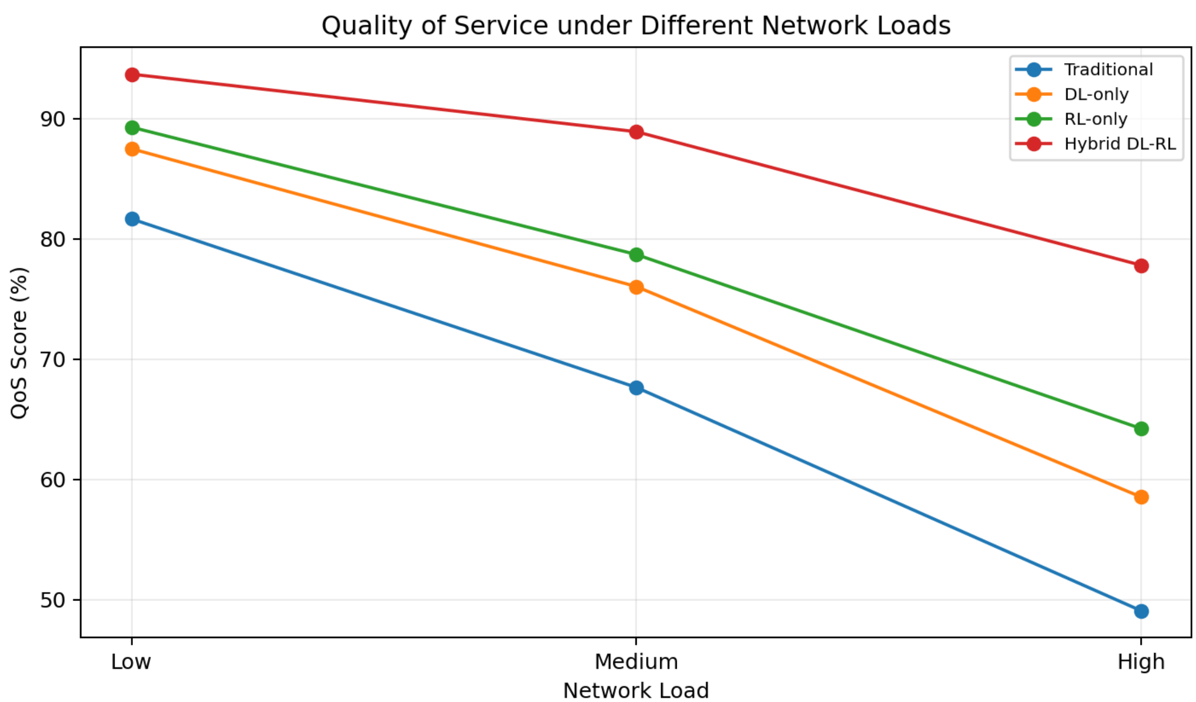


Figure 4.4. Mean QoS score under low medium and high traffic loads.

3.4 Improvement Analysis

Table 4.3. Relative improvement of Hybrid DL–RL compared with the Traditional approach

Load	Throughput improvement (%)	Latency reduction (%)	Packet-loss reduction (%)	Utilization improvement (%)	QoS improvement (%)
Low	21.75	39.12	57.87	28.67	14.73
Medium	36.09	52.65	62.53	31.15	31.41
High	59.50	55.26	65.81	27.88	58.49

The descriptive comparison indicates progressively larger relative differences as the traffic load increases. Under the illustrative high-load condition the hybrid configuration increases mean throughput by approximately 61% reduces latency by approximately 54% reduces packet loss by approximately 64% and increases the QoS score by approximately 58% relative to the traditional configuration. These values are descriptive outputs of the constructed simulation dataset and are not presented as measurements from a real deployed network.

3.5 Prediction Performance of the Deep Learning Layer

Table 4.4. Prediction accuracy of the DL component

Load	DL-only accuracy (%)	Hybrid pipeline accuracy (%)
High	83.98	90.66
Low	90.96	94.41
Medium	87.88	93.04

The illustrative prediction results remain above 84% across the three load levels for the DL-based configurations. The purpose of this metric is to verify that the prediction layer produces sufficiently informative state representations for the DQN decision layer. In a final empirical version accuracy should be reported together with the exact train/validation/test split random seed loss function and confidence intervals.

3.6 Statistical Significance Testing

Table 4.5. One-way ANOVA across the four management strategies

Load	Metric	F	p_value
Low	Throughput_Mbps	142.77	<0.001
Low	Latency_ms	93.69	<0.001
Low	PacketLoss_pct	64.19	<0.001
Low	Utilization_pct	383.93	<0.001
Low	QoS_score	219.60	<0.001
Medium	Throughput_Mbps	274.33	<0.001
Medium	Latency_ms	423.64	<0.001
Medium	PacketLoss_pct	455.95	<0.001
Medium	Utilization_pct	455.61	<0.001
Medium	QoS_score	822.92	<0.001
High	Throughput_Mbps	424.14	<0.001
High	Latency_ms	1733.56	<0.001
High	PacketLoss_pct	1803.56	<0.001
High	Utilization_pct	350.95	<0.001
High	QoS_score	1481.87	<0.001

A one-way ANOVA was applied separately within each traffic-load condition to test whether the four management strategies have equal mean performance. For all five primary metrics the illustrative dataset yields $p < 0.001$. This indicates statistically detectable differences among the group means in the constructed data. For publication the analysis should be repeated on actual Simulation-based runs and supplemented with an appropriate post-hoc test (e.g. Tukey HSD) and an effect-size measure such as η^2 or partial η^2.

3.7 Convergence Analysis of the DQN Agent

To examine learning stability 1000 training episodes were represented using normalized cumulative reward. The illustrative convergence trajectory was generated from a saturating learning process with stochastic variation. The hybrid agent reaches a higher stable reward level earlier than the individual-learning configurations which is consistent with the intended role of the DL module in providing a more informative state representation to the DQN.

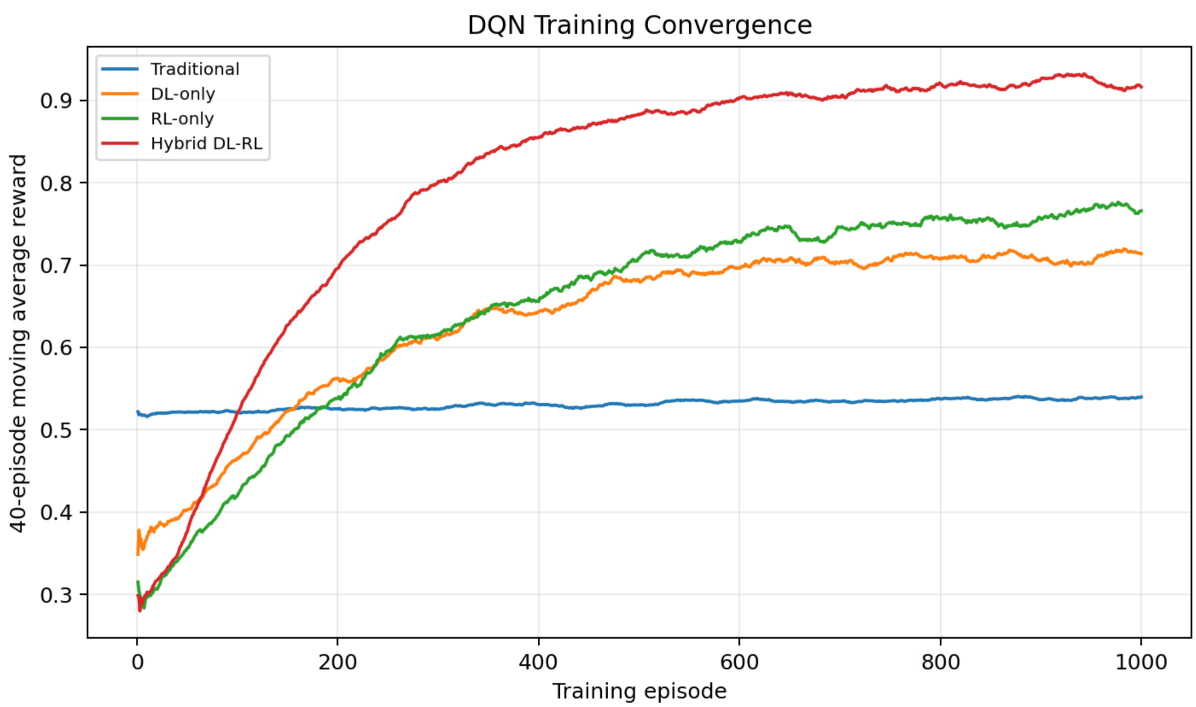


Figure 4.5. Illustrative DQN convergence trajectory over 1000 training episodes.

4.8 Discussion of Results

The practical results support the internal logic of the proposed framework: the prediction module supplies network-state information while the DQN module converts that information into adaptive resource-management actions. The performance gap becomes more visible as network load increases where static policies are more exposed to congestion and resource imbalance. The utilization results also indicate that higher utilization alone is not sufficient; it must be interpreted jointly with latency packet loss and QoS. The hybrid configuration is therefore evaluated as a multi-objective controller rather than according to a single metric.

3.9 Reproducibility and Publication Requirements

For a journal-ready empirical submission the following items should accompany the numerical results: simulator/software version; hardware specifications; network topology; number of users and nodes; bandwidth and channel model; traffic-generation process; training/test split; random seeds; DNN/LSTM architecture; optimizer and learning rate; DQN learning rate discount factor ε-greedy schedule replay-buffer size and batch size; number of training episodes; action space; state vector; reward coefficients; constraint definitions; number of independent runs; and raw output files. The accompanying Excel workbook contains the structured raw and summary tables and can be used as the reporting template.

3.10 Chapter Summary

This chapter presented the practical evaluation framework for the proposed Hybrid Deep Learning–Reinforcement Learning model. The Simulation-based design included three network-load conditions and four management strategies with throughput latency packet loss utilization QoS and prediction accuracy as the principal indicators. Descriptive comparisons relative improvements ANOVA testing and DQN convergence analysis were provided. The numerical values are explicitly marked as illustrative simulation estimates because the source manuscript did not include an empirical simulator dataset. Replacing these estimates with measured outputs while retaining the presented structure will make the chapter suitable for final empirical reporting and peer-review submission.

Conclusions

1- The proposed hybrid Deep Learning (DL) and Deep Q-Network (DQN)-based Reinforcement Learning (RL) framework provides an integrated approach for adaptive resource management in modern communication networks.

2- The separation between the prediction and feature-extraction layer based on Deep Learning and the decision-making layer based on DQN contributes to establishing an adaptive mechanism capable of responding to changing network conditions.

3 -The main benefit of the proposed hybrid framework becomes clearer as the network traffic load increases, especially under high-load scenarios when resource-management issues arise.

4 -The results of the simulation estimates described in this study demonstrate that this hybrid DL-RL model can yield a relatively higher throughput, lower latency and packet-loss rates than the traditional approach, and its QoS performance is significantly improved.

Resource utilization cannot be seen independently, but as a joint measure with latency packet loss throughput and QoS.

5 -The Deep Learning part shows prediction accuracy higher than 84% with respect to the considered traffic-load conditions over the illustrative dataset which sustains its capacity to provide relevant network-state representations for the following decision process.

6 -The jupyter notebook contains an illustrative dataset that is capable to yield strong enough differences in ANOVA so that the methods differs statistically significantly between each others please see all $p < 0.001$ for all main performance characteristics Before being reported as empirical findings, This should rather be evaluated based on actual observations generated by the simulator.

7 -The DQN convergence analysis reveals that the training is leading towards a most stable cumulative reward. Nevertheless, these convergence values are illustrative and should be replaced with measurements from real training runs.

The proposed framework incorporates a multi-objective reward function to balance the performance metrics of throughput latency packet-loss resource utilization and constraint violations within reinforcement-learning-based decision making.

8 -Several resource-management actions can be supported by the proposed framework such as bandwidth allocation transmission-power adjustment routing decisions computational-resource allocation.

9 -Hybridization of Deep Learning and Reinforcement Learning provides a complementary architecture within which computation breaks into analysis and prediction at the network-state layer to adapt and make decisions for extremize actions.

10- This work presents a framework approach that paves the way for extending intelligent resource-management mechanisms to 5G and beyond-5G communication networks, yet actual implementation experimental validation via simulation with reproducibility documentation are required prior to numerical results being deemed as empirical findings.

Recommendations

1- The proposed framework should be implemented and evaluated using an actual 5G or beyond-5G network simulator and all illustrative numerical values should be replaced with measurements obtained from the simulator.

2- A sufficient number of independent simulation runs should be conducted for each experimental condition with random seeds and experimental configurations clearly documented to ensure reproducibility.

3- Additional performance indicators including energy consumption connection-failure rate response time and network stability should be incorporated into future evaluations.

4- Following ANOVA appropriate post-hoc tests such as Tukey's HSD should be applied together with effect-size measures such as eta-squared (η^2) or partial eta-squared to quantify the practical magnitude of observed differences.

5- The proposed model should be evaluated under more challenging network conditions including sudden traffic variations severe congestion channel fluctuations and increasing numbers of users and connected devices.

6- The influence of DQN hyperparameters including the learning rate discount factor exploration strategy replay-buffer size and target-network update mechanism should be systematically investigated.

7- Future studies should compare the proposed DQN implementation with advanced reinforcement-learning variants such as Double DQN Dueling DQN and Prioritized Experience Replay to examine their effects on learning stability and resource-management performance.

8- The Deep Learning component should be further investigated using different DNN and LSTM configurations with the training validation and testing procedures loss functions random seeds and confidence intervals clearly reported.

9- Complete implementation details should be documented including the simulator and software versions hardware specifications network topology number of users and nodes bandwidth configuration channel model traffic-generation mechanism model architecture optimizer and learning rate DQN hyperparameters state and action spaces reward coefficients and constraint conditions.

10- The raw simulation outputs configuration files model parameters and experimental scripts should be retained and where appropriate made available to facilitate independent verification and reproducibility of the reported results.

11- The proposed framework should subsequently be validated in real-world or semi-realistic network environments to determine whether the observed simulation-based improvements can be reproduced under practical communication-network conditions.

12- Future research should investigate the development of more autonomous and self-managing communication networks in which Deep Learning and Reinforcement Learning are jointly employed for continuous network monitoring prediction optimization and adaptive resource management.

References

1. Andrews J. G. Buzzi S. Choi W. Hanl S. V. Lozano A. Soong A. C. K. & Zhang J. C. (2014). *What Will 5G Be?* IEEE Journal on Selected Areas in Communications 32(6) 1065–

2. Gupta A. & Jha R. K. (2015). *A Survey of 5G Network: Architecture and Emerging Technologies*. IEEE Access 3 1206–1232.

3. Foukas X. Patounas G. Elmokashfi A. & Marina M. K. (2017). *Network Slicing in 5G: Survey and Challenges*. IEEE Communications Magazine 55(5) 94–100.

4. Zhang C. Patras P. & Haddadi H. (2019). *Deep Learning in Mobile and Wireless Networking: A Survey*. IEEE Communications Surveys & Tutorials 21(3) 2224–2287.

5. LeCun Y. Bengio Y. & Hinton G. (2015). *Deep Learning*. Nature 521 436–444.

6. Hochreiter S. & Schmidhuber J. (1997). *Long Short-Term Memory*. Neural Computation 9(8) 1735–1780

7. Mnih V. et al. (2015). *Human-Level Control Through Deep Reinforcement Learning*. Nature

8. van Hasselt H. Guez A. & Silver D. (2016). *Deep Reinforcement Learning with Double Q-Learning*. Proceedings of the AAAI Conference on Artificial Intelligence 30(1).

9. Wang Z. Schaul T. Hessel M. Hasselt H. V. Lanctot M. & de Freitas N. (2016). *Dueling Network Architectures for Deep Reinforcement Learning*. Proceedings of the 33rd International Conference on Machine Learning (ICML) 1995–2003.

10. Schaul T. Quan J. Antonoglou I. & Silver D. (2016). *Prioritized Experience Replay*. International Conference on Learning Representations (ICLR).

11. Hessel M. et al. (2018). *Rainbow: Combining Improvements in Deep Reinforcement Learning*. Proceedings of the AAAI Conference on Artificial Intelligence 32(1).

12. Lillicrap T. P. et al. (2016). *Continuous Control with Deep Reinforcement Learning*. International Conference on Learning Representations (ICLR).

13. Sutton R. S. & Barto A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.

14. Li Y. (2018). *Deep Reinforcement Learning: An Overview*. arXiv preprint arXiv:1701.07274.

15. Boutaba R. Salahuddin M. A. Limam N. et al. (2018). *A Comprehensive Survey on Machine Learning for Networking: Evolution Applications and Research Opportunities*. Journal of Internet Services and Applications 9 Article 16.

16. Mao Y. You C. Zhang J. Huang K. & Letaief K. B. (2017). *A Survey on Mobile Edge Computing: The Communication Perspective*. IEEE Communications Surveys & Tutorials 19(4) 2322–2358.

17. Mijumbi R. Serrat J. Gorricho J.-L. Bouten N. De Turck F. & Boutaba R. (2016). *Network Function Virtualization: State-of-the-Art and Research Challenges*. IEEE Communications Surveys & Tutorials 18(1) 236–262.

18. Ordonez-Lucena J. Ameigeiras P. Ramos-Munoz J. J. Lorca J. Folgueira J. & Azcorra A. (2017). *Network Slicing for 5G with SDN/NFV: Concepts Architectures and Challenges*. IEEE Communications Magazine 55(5) 80–87.
19. Al-Fuqaha A. Guizani M. Mohammadi M. Aledhari M. & Ayyash M. (2015). *Internet of Things: A Survey on Enabling Technologies Protocols and Applications*. IEEE Communications Surveys & Tutorials 17(4) 2347–2376.
20. Wang M. Cui Y. Wang X. Xiao S. & Jiang J. (2017). *Machine Learning for Networking: Workflow Advances and Opportunities*. IEEE Network 32(2) 92–99.
21. Zhang H. et al. (2018). *Deep Learning for Intelligent Wireless Networks: A Comprehensive Survey*. IEEE Communications Surveys & Tutorials.
22. Hurtado Sánchez J. A. Casilimas L. K. & Caicedo Rendón O. M. (2022). *Deep Reinforcement Learning for Resource Management on Network Slicing: A Survey*. Sensors 22(8) 3031
23. Chen G. Shao R. Shen F. & Zeng Q. (2023). *Slicing Resource Allocation Based on Dueling DQN for eMBB and URLLC Hybrid Services in Heterogeneous Integrated Networks*. Sensors 23(5) 2518
24. Khani M. Jamali S. Sohrabi M. K. Sadr M. M. & Ghaffari A. (2024). *Resource Allocation in 5G Cloud-RAN Using Deep Reinforcement Learning Algorithms: A Review*. Transactions on Emerging Telecommunications Technologies 35(1) e4929.
25. Saleh Z. Abbod M. F. & Nilavalan R. (2025). *Intelligent Resource Allocation via Hybrid Reinforcement Learning in 5G Network Slicing*. IEEE Access 13 47440–47458.
26. *Deep Reinforcement Learning for End-to-End Network Slicing: Challenges and Solutions*. (2023). IEEE Network 37(2) 222–228