An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol. 16, No. 3, pp.794-815 (2026)
https://doi.org/10.36922/IJOCTA025500230

REVIEW ARTICLE

A comprehensive review of intelligent end-to-end networking solutions through the integration of graph neural networks and deep reinforcement learning

Muhammad Kamran[1,2,3†], Salwa Muhammad Akhtar[4†], Muhammad Zain ul Abideen[5], Junaid Asghar[6], Muhammad Farman[1,7*], Aseel Smerat[8,9], and Mohamad Hafez[2,10]

[1] Mathematics Research Center, Department of Mathematics Near East University, Mersin, Turkey
[2] Department of Mathematics, Faculty of Engineering and Quantity Surviving, INTI International University Colleges, Nilai, Negeri Sembilan, Malaysia
[3] International Center for Interdisciplinary Research in Sciences, The University of Lahore, Lahore, Pakistan
[4] Department of Information Systems, University of Management and Technology, Lahore, Punjab, Pakistan
[5] Department of Mechanical Engineering, Faculty of Engineering, University of Central Punjab, Lahore, Punjab, Pakistan
[6] Department of Computer Science Information Technology, Faculty of Information Technology, University of Lahore, Lahore, Punjab, Pakistan
[7] Research Center of Applied Mathematics, Khazar University, Baku, Azerbaijan
[8] Department of Mathematics, Faculty of Educational Sciences, Al-Ahliyya Amman University, Amman, Jordan
[9] Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India
[10] Department of Management, Faculty of Management, Shinawatra University, Sam Khok, Pathum Thani, Thailand
kamrankfueit@gmail.com; salwa.akhtar@umt.edu.pk; mzainua10@hotmail.com; mjasghar52@gmail.com; farmanlink@gmail.com; smerat.2020@gmail.com; mohdahmed.hafez@newinti.edu.my

ARTICLE INFO

Article history:
Received: December 14, 2025
Revised: January 19, 2026
Accepted: January 26, 2026
Published Online: April 30, 2026

Keywords:
Graph neural networks
Deep reinforcement learning
Intelligent networking
End-to-end network optimization
Autonomous network management
Resilient infrastructure

ABSTRACT

Topology awareness and scalable, adaptive network control have become critical with the development of 5G/6G, the Internet-of-Things, vehicular networks, and edge computing. Traditional rule-based and centralized networking models are unable to support dynamic topologies, heterogeneous traffic models, and demands with strict quality-of-service requirements. Structural and topological dependencies are encoded using graph neural networks (GNNs) and combined with deep reinforcement learning (DRL) to make decisions sequentially. Exploiting rewards is an avenue toward intelligent end-to-end network optimization. This review is a systematic examination of modern GNN–DRL models implemented in routing, congestion control, chaining of service functions, vehicular communication, and the optimization of optical networks. It also highlights their performance strengths, including topology awareness, cross-topology generalization, high sample efficiency, and high scalability, as well as their weaknesses, such as inference overhead, inconsistent benchmarking practices, low real-time deployability, and sensitivity to noisy or partial state observations. The main findings of this review are: (i) a coherent taxonomy of GNN-based, DRL-based, and hybrid GNN–DRL effective designs; (ii) comparative analysis of algorithms, architecture components, and learning pipelines; (iii) generalized performance trends in major areas of

* Corresponding Author
† *These authors contributed equally to this work.*

intelligent networking; and (iv) a collection of grounded research directions to be followed in the future, lightweight architecture, transfer learning pipeline, fault-tolerant learning, and unified evaluation frameworks. Finally, this review focuses on enabling resilient infrastructure through intelligent, scalable, and autonomous end-to-end networking solutions.

1. Introduction

Highly adaptable and dynamic communication infrastructures require smart, flexible, and scalable control infrastructures that can meet dynamic topologies, heterogeneous traffic characteristics, and strict quality-of-service (QoS) requirements. Conventional rule-based networking designs are typically unable to satisfy these needs because they lack scalability (they are inflexible) and cannot make inferences in highly dynamic contexts.[1,2] Although several studies have examined graph neural networks (GNNs) and deep reinforcement learning (DRL), there is no integrated review of how these two concepts can be combined to achieve complete end-to-end networking. Current reviews either (i) are limited to a study of GNN modeling, (ii) look at DRL-based control, or (iii) do not address scalability, generalization, and real-time deployment issues. New network models, including 5G/6G, Internet-of-Things (IoT) networks, vehicle-to-vehicle communications, and multi-access edge computing, provide further incentive to the use of data-driven, learning-enabled, and autonomous optimization models.[3,4]

Graph neural networks and DRL are two technologies that have become highly promising in the next-generation intelligent networking world. GNNs offer a principled system for representing and modeling communication networks, their topologies and interactions, and the structural correlations between nodes.[5,6] Complementary to this, DRL provides strong sequential decision-making, and networks can optimize routing, resource distribution, congestion, and spectrum management via reward-based learning on their own.[7]

Both GNNs and DRL are complementary, and neither can attain the capabilities of the other. GNNs provide a topology-aware representation of the network, capturing node relationships, node dependencies, link congestion patterns, and global network behavior.[8] DRL, on the other hand, is a sequential decision-maker that enables autonomous routing, scheduling, and resource optimization through reward-based policies. GNN embeddings serve as effective state representations for DRL agents when combined, enabling generalization to unseen topologies, reducing state ambiguity, and improving policy stability. This synergy creates a coherent, end-to-end intelligent system that can optimize large, dynamic, and heterogeneous networks much more efficiently than GNNs or DRL.

Recent studies have demonstrated that adopting GNN-based topology representations in DRL controllers can improve decision-making in most networking tasks. Graph embeddings have been applied to enhance the performance of DRL models for routing optimization, virtual network function (VNF) placement, wireless resource allocation, and multi-domain network control, where structural dependencies are usually ignored by traditional DRL agents.[9] As a result of such embeddings, DRL agents can obtain more information about congestion, bottlenecks, and optimized actions by making informed topology choices. Besides, GNN–DRL was found to be more scalable and resilient in large and distributed network environments, enabling collaborative decision-making and improving its generalization to novel network environments.[10]

Although there has been major progress, the current studies are still rambling. The majority of studies either focus solely on GNN-based topology modeling or solely on DRL-based control, without providing an integrated view of their capabilities for working together to create intelligent end-to-end networking.[10,11] Scalability, generalization capability, constraints on real-time inference, and evaluation practices are also diverse across current approaches, making it difficult to compare results across works. This fragmentation and lack of methodological coherence underscore the need for a general review that combines both GNN and DRL paradigms.[12]

Figure 1 shows a general interaction between GNNs and DRL in optimizing autonomous networks. The capacity of GNNs to model node dependencies, linkages, and the topological structure of communication networks has made them effective models of communication networks. GNNs can produce expressive embeddings for a wide variety of tasks, such as traffic prediction, routing estimation, anomaly detection, and link-state inference, through message-passing mechanisms.

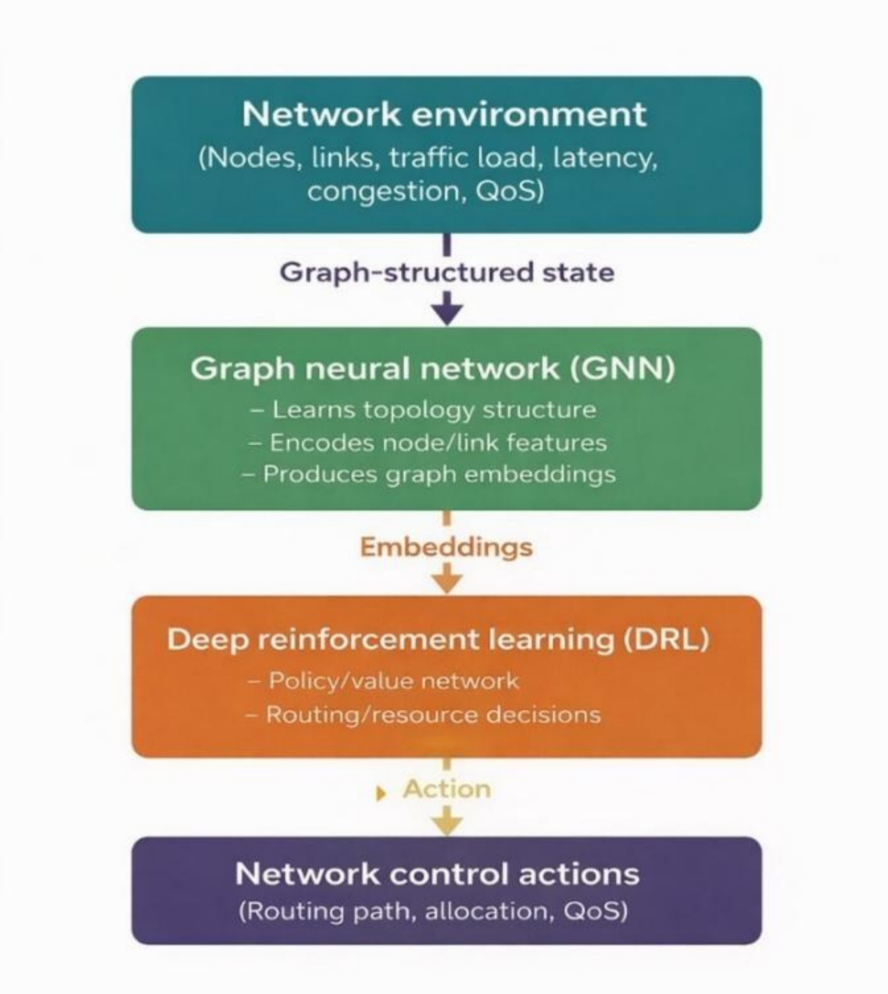

Figure 1. High-level framework of GNN–DRL integration for intelligent end-to-end networking
Abbreviation: QoS: Quality of service.

Recent intelligent networking studies investigate GNNs and DRL independently but do not comprehensively compare their combined capabilities to enable end-to-end autonomous network operation.[13] Although 5G/6G, IoT, and vehicular systems are adopting it rapidly, the current literature lacks consistent benchmarking, scalability testing, and real-time testing of this approach, leaving the GNN and DRL integration opaque across studies.[14]

The main contributions of this review are:

(i) An extensive overview of the GNN-based, DRL-based, and hybrid GNN–DRL models for intelligent end-to-end networking in key application fields.
(ii) An integrated GNN–DRL taxonomy that hierarchically categorizes the existing approaches according to learning paradigm and networking task, which has not been previously reviewed.
(iii) An end-to-end framework based on topology-aware representation learning, adaptive DRL-based control, and closed-loop feedback of autonomous networking.
(iv) A cross-domain comparative study that covers routing and congestion control, service function chaining (SFC), vehicular networks, optical networks, and IoT networks, and identifies gaps in scalability and benchmarking.
(v) Critical architectural and evaluation requirements discovery in existing research, resulting in grounded research directions to scalable and deployable intelligent networking systems.

The paper is structured as follows: Section 2 gives background on the foundations of GNNs and DRL. Section 3 provides the methodology framework, combining GNN-based representation learning and DRL-based adaptive control. Section 4 discusses areas of application and comparative knowledge. Section 5 describes available challenges and future research directions. Section 6 concludes the paper.

2. Background and fundamental concepts

Intelligent learning-based networking systems are becoming popular for managing complex, dynamic, and large-scale communication scenarios.[15] Modern networks require models that can learn structural dependencies, adapt to dynamic traffic patterns, and make autonomous decisions in cases of uncertainty. This section provides an overview of GNNs and DRL, and explains how they can be used to enable end-to-end intelligent networking.

2.1. Graph neural networks in networking

Graph neural networks can be seen as a structured framework for modeling communication networks, where nodes, links, and topologies naturally form graphs. As emphasized previously,[16] GNNs, including graph convolutional networks (GCNs), graph attention networks (GATs), adaptive graph encoders, and spatial-temporal GNNs, are effective at capturing node-to-node relationships, link attributes, and topology-conscious dependencies. The features of GNN render them applicable to the following tasks: routing estimation, traffic forecasting, anomaly detection, and link-state inference. By means of message passing and aggregation, GNNs produce expressive embeddings that can be generalized to different (or unseen) network structures.[17] Research has shown their great abilities in topology-sensitive representation learning and in modeling graph-based functionality for networking problems.[18] However, GNNs are also good at learning structural patterns but do not include sequential decision-making or adaptive real-time control per se, which implies they must be complemented with alternative techniques.

2.2. Network control deep reinforcement learning

Deep reinforcement learning has also emerged as an enabling factor in autonomous network control, enabling sequential, reward-driven decision-making in dynamic and uncertain environments. Deep Q-networks (DQNs), proximal policy optimization (PPO), deep deterministic policy gradient (DDPG), and asynchronous advantage actor-critic (A3C) algorithms have demonstrated significant routing, congestion control, resource allocation, and spectrum management efficiencies by learning optimal policies through interaction with the environment.[19,20] In contrast to static or rule-based controllers, DRL agents can adapt to traffic variability, topology changes, and heterogeneous QoS requests without explicitly modeling the network.[21,22] DRL has been demonstrated to be useful for software-defined networking (SDN) routing, IoT resource allocation, bandwidth control, and even for activities such as intrusion and distributed denial-of-service (DDoS) detection, with significant improvements in throughput, latency, and stability.[23] Furthermore, it is common for DRL agents to fail to generalize when network structures greatly change or when raw state representations do not reflect graph-based relationships.[21] The topology-sensitive encoders, like GNNs, which yield structure-based representations of states that enhance the robustness and learning capabilities of DRL, are highlighted by this limitation.[24]

2.3. Current graph neural network–deep reinforcement learning models and applications

Recent work has shown that combining topology representations via GNNs with a DRL policy improves decision-making across many networking areas.[25,26] GNN embeddings also enable DRL agents to reason about structural dependencies that classical DRL approaches do not consider, enhancing routing optimization, VNF placement, wireless resource allocation, and multi-domain network control. These are enhanced representations of states that help DRL agents identify congestion points, bottlenecks, and the best courses of action. Moreover, hierarchical and multi-agent GNN–DRL models have demonstrated higher scalability and robustness in large, distributed, and dynamic networked settings.[22] These hybrid models are a synthesis of two paradigms: topology-aware representation learning in GNNs and adaptive control in DRL, thereby enabling more efficient and generalizable network intelligence.

2.4. Limitations of the existing literature

Despite the promise of hybrid GNN–DRL models, a number of drawbacks exist in the contemporary literature. Most research assesses its methods only on small or artificial topologies with few traffic variations, which limits their applicability to real-world, large-scale networks.[27] The high cost of training and inference, which is high in both GNNs (due to message passing in an iterative manner) and DRL (because of policy updates), complicates deployment in high-latency settings, such as vehicle-to-everything (V2X) and industrial IoT.[28] Furthermore, current techniques are susceptible to noisy measurements, biased observability, adversarial constraints, and have link failures, thereby diminishing their accuracy in

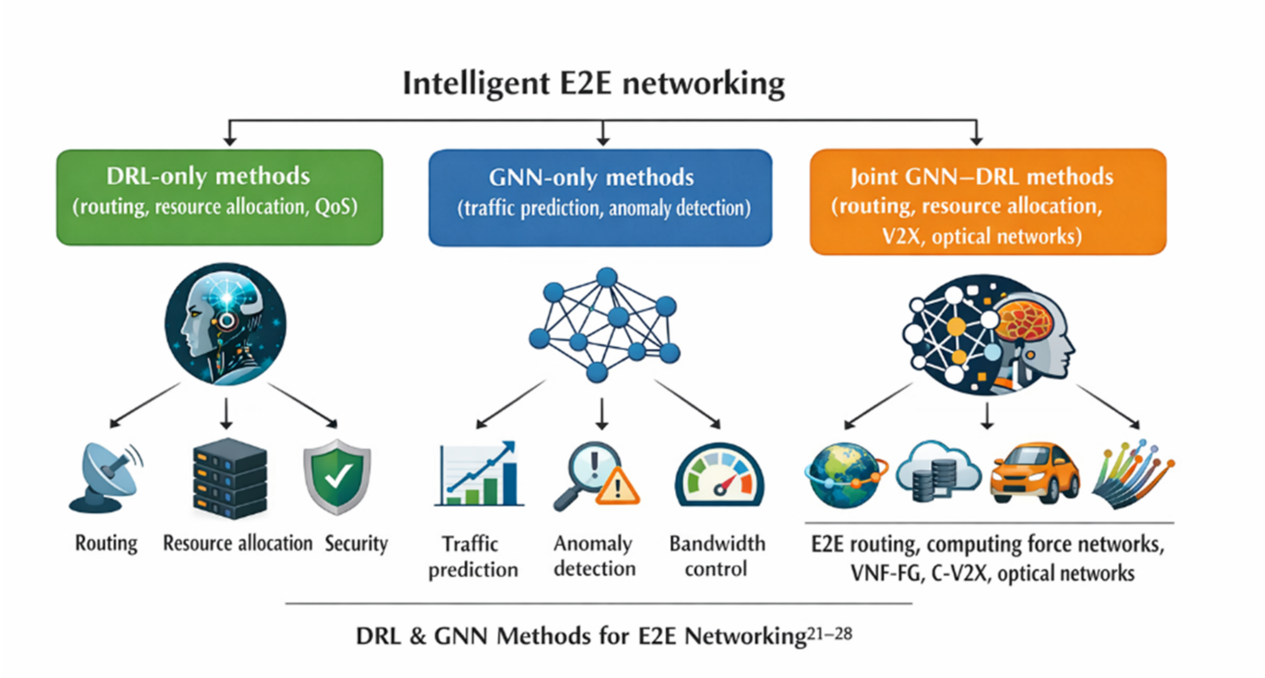

Figure 2. Taxonomy of intelligent networking solutions using GNNs and DRL
Abbreviations: C-V2X: Cellular vehicle-to-everything; DRL: Deep reinforcement learning; E2E: End-to-end; GNN: Graph neural network; IoT: Internet-of-Things; QoS: Quality of service; SDN: Software-defined networking; VNF-FG: Virtual network function-forwarding graph.

realistic settings.[29] The diversity of evaluation practices across studies remains disparate, with inconsistent datasets, topologies, performance measures, and comparisons to other studies, hindering reproducibility.[30]

Table 1 presents a synthesis of representative GNN–DRL-based networking solutions reported in the recent literature, including their applications, learning components, and contributions. This comparison shows how various studies used graph representations with reinforcement learning policies to solve routing, resource allocation, congestion management, and multi-domain optimization.[31] It is also evident in the variety of tasks and architectures employed in the field, as the table shows a clear foundation for understanding the design trends and performance properties of hybrid GNN–DRL models, as shown in **Figure** 2.

The current research literature can be grouped into two basic dimensions. The first dimension concerns the learning paradigm and comprises three main model types: DRL-only, GNN-only, and hybrid GNN–DRL models, which can be combined to achieve end-to-end intelligent control. The second dimension is associated with the networking task, wherein the previous research has mainly focused on other core operational functions, including routing optimization, congestion management, resource allocation, traffic prediction, and network security.[32,33] This two-sided taxonomy can offer a systematic perspective on the relationships among various methodological decisions with specific networking goals, thereby allowing one to see more clearly the advantages and weaknesses within the scope of current research. **Figure** 2 presents the conceptual representation of this taxonomy.

3. Intelligent networking frameworks

This section provides the methodological background for analyzing, interpreting, and unifying the available DRL-based GNN-based intelligent networking solutions. The methodology provides a systematic framework for consistently analyzing a variety of architectures, learning processes, and networking applications. It establishes the network model, outlines the architectures of GNN and DRL components, explains how the two paradigms interact within the hybrid architecture, and provides the end-to-end workflow for the analysis throughout the paper.

3.1. Network model and problem definition

The methodological framework governed how models were assessed in the following sections. Uniform criteria for scalability, latency, generalization, and robustness were used to evaluate all GNN-only, DRL-only, and hybrid GNN–DRL studies. A communication network could

Table 1. Representative hybrid GNN–DRL solutions for end-to-end networking

Reference	Scenario/Network type	Task	Learning components	Key contribution
[13]	Packet-switched networks	Routing optimization	GNN state encoder + DRL policy network	First routing use case where GNN embeddings feed a DRL agent, improving generalization across topologies
[14]	Multi-objective communication networks	Generic routing optimization	Modular GNN–DRL framework	Unified method (GROM) that supports multiple routing objectives and network settings
[15]	Software-defined optical networks	Routing & spectrum assignment	GNN topology encoder + actor-critic DRL	End-to-end GNN–DRL routing method (GDRL-SFCR) that improves blocking probability and delay
[16]	Computing force networks	Joint compute–network allocation	GNN resource graph + DRL agent	GNN–DRL scheme that jointly manages traffic and computing resources under topology variation
[17]	Network function virtualization/Service function chaining	Virtual network function-forwarding graph embedding	GNN representation + DRL policy	Considers structural, node, and link constraints in embedding service chains into substrate networks
[18]	Cellular vehicle-to-everything vehicular networks	Radio resource allocation	Connectivity graph GNN + DRL scheduler	Meets ultra-reliable, low-latency requirements for V2V/V2I links using a GNN–DRL controller
[19]	Multi-domain communication networks	Traffic prediction & optimization	GNN predictors + DRL-ready models	Thesis arguing for integrating GNN traffic prediction with DRL for closed-loop optimization
[20]	Optical communication networks	Routing & spectrum assignment	GNN-assisted DRL	Shows that GNN-encoded topology improves DRL sample efficiency and robustness across network instances

Abbreviations: DRL: Deep reinforcement learning; GDRL: Graph Deep Reinforcement Learning; GNN: Graph neural network; GROM: Graph Representation of Optical Networks ModelSFCR: Service Function Chain Routing; V2I: Vehicle-to-infrastructure; V2V: Vehicle-to-vehicle.

be mathematically represented as a graph $G(V, E)$, where V denotes the nodes (routers, base stations, switches, or servers) and E denotes the communication links between them. The feature attributes of each node ?? $\in V$ are the length of its queue, traffic load, delay, central processing unit (CPU) utilization, buffer occupancy, and the bandwidth capacity, propagation delay, jitter, level of congestion, or spectrum availability of each link ?? $\in E$. By modeling the network as a graph, these features allow it to better capture spatial relationships, structural dependencies, and dynamic state variations than traditional flat feature representations.

The objective of the intelligent controller in this context is to maximize the performance of the end-to-end network decision through the choice of action, including path update, resource assignment, flow scheduling, or congestion control, dependent on its perceived graph state. To be effective in decision-making, the learning model should capture spatial relationships, temporal variations, and structural constraints within the network. The graphical representation of the network can be used as a topology-aware backbone to facilitate more advanced learning processes, especially when combined with GNN-based state encoding and DRL-based adaptive control. Such a formulation allows the controller to decipher the conditions of the global network and the rationale for relational patterns, and to make coordinated decisions that align with QoS goals.

3.2. Topology encoding using graph neural networks

Graph neural networks are used to model the structural perception of the system. A typical GNN layer performs:

$$hi(k) = AGGREGATE(hj^{(k-1)} : j \in N(I)) \tag{1}$$

Graph neural networks can compute both local and global structural information via the message-passing mechanism, thus allowing the downstream DRL agent to typically operate on a rich, topologically structured state representation rather than raw metrics. This improves generalization to unseen topologies and reduces the need for handcrafted network features.

3.3. Adaptive control using deep reinforcement learning

The DRL component serves as a decision-making engine. At each time step, the agent observes the GNN-encoded network state s_t, chooses an action (a_t), and receives a reward, reflecting performance goals such as throughput maximization, latency minimization, or fairness improvement. A DRL policy is often expressed as:

$$\pi\theta(a_t|s_t) \tag{2}$$

Depending on the task, various DRL algorithms—DQN, DDPG, PPO, and A3C—can be used. The incentive function is designed to ensure quality of service objectives and stability, for example:

$$rt{=}-\alpha \cdot Delay{+}\beta \cdot Throughput{-}\gamma \cdot Packetloss \tag{3}$$

This adaptive learning approach allows the controller to optimize network behavior over time without relying on predetermined heuristics or explicit traffic models.

3.4. End-to-end intelligence multi-layer integration

The multi-layer integration end-to-end intelligence in current communication networks, in which the various layers of learning are tightly coupled to work together in the perception of network conditions, rationalize structural dependencies, and implement optimized control decisions. The suggested multi-layer architecture integrates graph-based representation learning, reinforcement-based policy optimization, and real-time environmental feedback, forming a single intelligence loop capable of adapting to a changing, large-scale, and heterogeneous network environment.

3.4.1. Layer 1: Graph construction and feature extraction

At the first layer, the graph construction and feature extraction processes convert the raw network state into a structured graph G $= (V, E)$ augmented with node-level and link-level properties such as queue occupancy, latency, load distribution, and bandwidth availability. The GNN perception layer processes these features and, via iterative message passing, generates topology-aware embeddings that reflect local neighborhood interactions and global structural patterns. These embeddings offer a succinct yet descriptive view of the context around the operation of the network.

3.4.2. Layer 2: Graph neural network-based perception

The second layer, the DRL policy optimization module, uses the GNN-encoded embeddings as the state input. Using value-based, policy-based, or actor-critic learning methods, the DRL agent identifies the best control policies, e.g., route adjustments, flow scheduling, congestion reduction, or resource distribution, after optimizing long-term reward. The state representation obtained with the GNN is much more effective at enabling the agent to generalize across different topologies and various traffic conditions.

3.4.3. Layer 3: Deep reinforcement learning policy optimization

The third layer defines environment interaction and feedback, in which the chosen actions are implemented within the network, leading to updates to states and performance metrics (e.g., throughput, delay, jitter, or loss). These results enable giving instant feedback on the policy to improve and update the GNN–DRL layers. The structure is closed-loop, allowing it to be continuously learned, corrected, and adapted without human intervention.

3.4.4. Layer 4: Environment Interaction and Feedback

A multi-layer intelligence pipeline is generated, in which GNNs provide structural insights, DRL makes adaptive decisions, and the feedback loop between the environment and the system guarantees continuous optimization. This combined

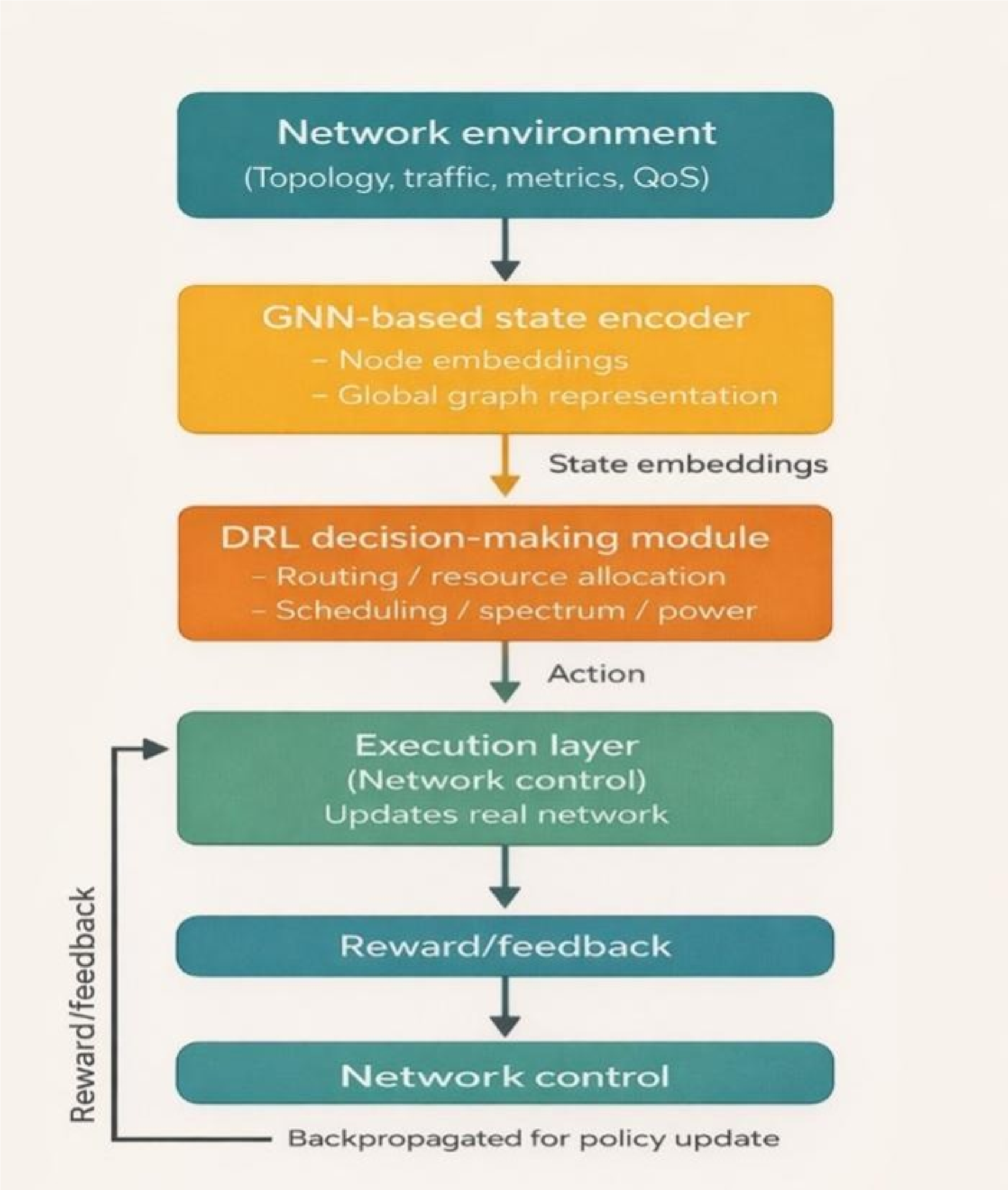

Figure 3. Unified graph neural network (GNN)–deep reinforcement learning (DRL) framework
Abbreviation: QoS: Quality of service.

method will enable real-time and autonomous networking that can respond to unexpected changes, scale effectively across large networks, and remain resilient to failures and shifting traffic patterns. The resulting architecture can be used as a broadly applicable methodological framework that facilitates comparative analysis and subsequent findings in this paper.

Figure 3 describes the process of transforming real network data (topology, traffic, metrics, and QoS) into a graph, then using a GNN to create state embeddings for the graph, including nodes, edges, and the entire network. This is then sent to the DRL module, which applies these embeddings to make intelligent decisions, such as routing, resource allocation, scheduling, spectrum allocation, or power control. The action submitted is then im-

Table 2. Key components of intelligent GNN–DRL networking frameworks

Layer	Role	Techniques used	Main outcome
Network graph construction	Converts the network to a graph	Topology modeling, feature extraction	Structured graph input
GNN-based encoder	Learns topology-aware representations	GCN, GAT, Graphs AGE, GIN	Node/edge embeddings
DRL controller	Decision-making	DQN, PPO, DDPG, SAC	Optimized routing, allocation
Closed-loop interaction	Continuous learning	Reward-based updates	Adaptive self-optimization
Multi-layer integration	End-to-end intelligence	Cross-layer coordination	Holistic network optimization

plemented in the Execution Layer, which modifies the actual network. The output or feedback of the network is relayed to the DRL agent, thereby refining the policy over time. This approach renders the entire pipeline a closed-loop, intelligent learning system that continuously optimizes network performance, as shown in **Table** 2.

4. Application domains (routing, congestion, service function chaining, vehicle-to-everything, and optical networks)

This section outlines key application areas where integrated GNN–DRL frameworks have achieved substantial gains in network intelligence. The preceding chapters have explained the foundations of architecture; here, the practical implications are discussed. The following insights are associated with representative tasks, including routing optimization, resource allocation, congestion control, traffic prediction, and SFC.

Table 4 presents a qualitative coverage comparison of GNN-based, DRL-based, and hybrid GNN–DRL models in intelligent end-to-end networking. The table highlights that GNN-based models provide complete topology awareness and strong scalability by capturing structural dependencies within network graphs; however, they lack sequential decision-making capabilities and offer only partial end-to-end control and adaptability. In contrast, DRL-based models excel at sequential decision-making and real-time adaptability through reward-driven policy learning, but exhibit only partial topology awareness and moderate scalability due to limitations in their structural representation. Hybrid GNN–DRL models integrate the strengths of both paradigms, achieving complete coverage across topology awareness, sequential control, scalability, real-time adaptability, and holistic end-to-end optimization. This comparison clearly demonstrates why hybrid GNN–DRL architectures are better suited to large-scale, dynamic, and autonomous networking environments.[27,29]

4.1. Comparative terminology landscape

The nomenclature for GNN-based, DRL-based, and hybrid GNN–DRL models is inconsistent in intelligent end-to-end networking studies. Similar concepts are defined by different names across studies; this is one of the causes of ambiguity when comparing methodologies, comprehending architectures, and interpreting results. Thus, a comparative terminology map is needed to standardize vocabulary, define conceptual frames, and harmonize definitions across studies. In this subsection, a systematic comparison between the key terms employed in the three broad methodological categories is presented: (i) DRL-based, (ii) GNN-based, and (iii) hybrid GNN–DRL models. By matching terms, researchers can see how various communities (networking, machine learning, wireless communications, and graph learning) share a common core component. This helps prevent confusion during interpretation and facilitates communication across fields.

These acronyms encompass a wide range of concepts, technologies, learning models, optimization algorithms, and networking architectures that underpin modern intelligent communication systems, as shown in **Table** 5. These include core machine learning terms, such as artificial intelligence (AI), deep neural networks, convolutional neural networks, GNNs, and DRL, as well as specific variants, such as GAT, GCN, and message passing neural networks (MPNNs), which represent different graph-learning approaches. It also lists widely used DRL algorithms—including DDPG, DQN, PPO, A3C, and multi-armed bandit—that support autonomous decision-making in network control. In addition, essential networking and

Table 3. Summary of related work and contributions to GNN–DRL-based end-to-end networking

Domain	GNN	DRL	End-to-end networking	Network optimization	Network management/ contr	Reference	Year
Intelligent wireless networking using GNNs (resource optimization & new trends)	✓	?	✓	✓	?	[22]	2021
Graph-based learning models for modern communication networks	✓	?	?	?	✓	[23]	2022
GNN-enabled automation for network control and orchestration	✓	?	?	?	?	[24]	2024
DRL–GNN integration: challenges, insights, and open research issues	✓	✓	×	×	×	[25]	2023
DRL applications in communication and networking	×	✓	×	×	✓	[26]	2022
Graph-driven reinforcement learning for networking: applications & future scope	✓	✓	?	×	×	[27]	2025

Notes: ? indicates partial support.
Abbreviations: DRL: Deep reinforcement learning; GNN: Graph neural network.?

Table 4. Qualitative coverage comparison of learning paradigms in intelligent end-to-end networking

Learning paradigm	Topology awareness	Sequential decision-making	End-to-end control	Scalability	Real-time adaptability
GNN-based models	Complete	No	Partial	Complete	Partial
DRL-based models	Partial	Complete	Partial	Partial	Complete
Hybrid GNN–DRL models	Complete	Complete	Complete	Complete	Complete

Abbreviations: DRL: Deep reinforcement learning; GNN: Graph neural network.?

communication concepts, such as QoS, quality of experience, service level agreement, radio access networks, mobile edge computing (MEC), SDN, network function virtualization (NFV), VNF, software-defined wide-area network (SD-WAN), and SFC, are included to contextualize system-level behavior in end-to-end networking environments. Several emerging domains and specialized technologies—such as unmanned aerial vehicle networks, vehicular edge computing, non-orthogonal multiple access, channel state information, and advanced frameworks like topology-aware graph neural networks and graph-enhanced DRL optimization—are also represented, indicating the breadth of interdisciplinary research.

4.2. Contribution-level comparison

The rapid development of intelligent networking research has led to a vast and varied body of literature that advances research across various areas, including representation learning, routing performance, traffic prediction, congestion control, spectrum assignment, VNF placement, and end-to-end autonomy. These, however, exhibit wide variation in their scope, methodology, learning elements, and learning performance goals. Thus, comparison at the contribution level is also necessary to clearly define the progress each technique makes in the field and to position hybrid GNN–DRL architectures relative to existing methods. Although most of the studies reviewed report

Table 5. Acronyms and descriptions relevant to intelligent end-to-end networking with GNN and DRL

Acronym	Description
AI	Artificial intelligence
CNN	Convolutional neural networks
CSI	Channel state information
DDPG	Deep deterministic policy gradient
DNN	Deep neural networks
DQN	Deep Q-networks
DRL	Deep reinforcement learning
GAT	Graph attention networks
GCN	Graph convolutional networks
GNN	Graph neural networks
ITU	International Telecommunication Union
MEC	Mobile edge computing
MLP	Multi-layer perceptron
MPNN	Message passing neural networks
MPTCP	Multi-path transmission control protocol
NFV	Network function virtualization
NOMA	Non-orthogonal multiple access
PPO	Proximal policy optimization
QoS	Quality of service
QoE	Quality of experience
RAN	Radio access networks
SDN	Software-defined networking
SFC	Service function chaining
SLA	Service level agreement
UAV	Unmanned aerial vehicles
VEC	Vehicular edge computing
VL	Virtual link
E2E	End-to-end networking
SD-WAN	Software-defined wide-area network
VNF	Virtual network function
MAB	Multi-armed bandit
A3C	Asynchronous advantage actor-critic
TopoGNN	Topology-aware graph neural networks
GEO-DRL	Graph-enhanced DRL optimization

quantitative improvements in latency, throughput, and scalability, a direct numerical comparison is not feasible due to heterogeneous benchmarking configurations.

Table 6 presents a refined and structured summary of existing studies on intelligent end-to-end networking using GNNs, DRL, and their integrated approaches. The table categorizes prior work into four major areas—GNN-based networking, DRL-based decision-making, integrated GNN–DRL models, and network-wide management frameworks, allowing a clearer understanding of how each technique advances modern communication networks. Within GNN-based solutions, studies applying GCN, GAT, and MPNN architectures highlight the strengths of graph learning for routing, link prediction, and traffic modeling, emphasizing improved structural awareness and predictive capability. DRL-focused works demonstrate how DQN, DDPG, and PPO enable autonomous optimization of routing, resource allocation, and congestion management in uncertain or dynamic environments. The integrated category showcases hybrid models in which GNNs enhance DRL agents by embedding topological features directly into the learning process, thereby improving routing performance, adaptability, and multi-agent coordination. Finally, the network-wide management segment covers SDN-, MEC-, and NFV-enabled implementations that deploy or support these intelligent learning mechanisms, ensuring low-latency inference, scalable optimization, and flexible control. Overall, **Table** 6 synthesizes the state of the art by illustrating how each technique, individually and collectively, contributes to building intelligent, fully optimized end-to-end networking solutions.

Table 6. Refined summary of existing studies and their contributions to intelligent end-to-end networking using GNN and DRL

Category	**Proposed work** 22 – 25	**Key technique/approach**	**Refined contribution**
GNN-based networking solutions	GCN-driven routing optimization	Graph convolutional models for learning network structures	Demonstrates that GCNs can effectively capture topological dependencies, enabling more reliable and efficient end-to-end routing decisions
	GAT-enhanced link prediction	Attention-based message aggregation	Provides finer-grained structural learning, resulting in more adaptive and accurate link-state estimation in dynamic networks
	MPNN for traffic modeling	Message passing neural operations on network graphs	Offers robust traffic forecasting capabilities that support proactive and intelligent resource provisioning
DRL for network control & decision-making	DQN-based path selection	Value-based learning for routing actions	Enables autonomous route optimization through real-time policy learning in complex and uncertain network environments
	DDPG-oriented resource assignment	Actor-critic DRL for continuous action control	Achieves optimal resource utilization through continuous decision refinement for bandwidth, power, and compute distribution
	PPO-supported congestion regulation	Policy-gradient DRL with stable updates	Enhances congestion control by providing stable, high-performance flow regulation across varying network conditions
Integrated GNN–DRL models	GNN-guided DRL routing	GNN-generated features embedded in DRL policy networks	Significantly boosts routing quality by integrating topological awareness into the decision-making process of DRL agents
	DRL with graph-based embeddings	Topology-informed embeddings generated via GNNs	Improves end-to-end adaptability by enabling DRL agents to operate with a richer, structure-aware representation of the network
	Hybrid multi-agent GNN–DRL architecture	GNN-enabled relational modeling within multi-agent DRL	Facilitates intelligent coordination and scalable control across distributed network nodes
Network-wide management & optimization	SDN-integrated DRL controller	Programmable SDN environment combined with DRL logic	Enhances global network management through real-time, data-driven decision-making and flexible control policies
	MEC-powered GNN/DRL inference	Deployment of learning models at edge servers	Reduces latency and improves the responsiveness of intelligent networking tasks via computation offloading
	NFV-enabled learning framework	Virtualized function-executing GNN/DRL modules	Ensures agile deployment, rapid updating, and scalable operation of learning-driven network optimization mechanisms

4.3. Domain-specific comparative insights

Domain-specific comparative knowledge illustrates the behavior of DRL-only, GNN-only, and hybrid GNN–DRL systems across various network settings, including data-center networks, internet service provider (ISP)-scale topologies, vehicular networks, optical networks, and IoT-based resource-constrained systems. One of the most important findings in cross-domain research is that DRL-only algorithms are effective for dynamically constrained decision-making problems, such as adaptive routing in fast-evolving V2X networks or congestion control in data-center networks, but they do not generalize when network structure changes. On the other hand, GNN-only models prove effective for tasks that depend on structure, such as traffic forecasting, topology inference, and link failure localization, because they can encode graph dependencies. Nevertheless, GNNs cannot offer real-time control or policy adaptation. The hybrid GNN–DRL models outperform both paradigms across all the studied areas. In data center and ISP networks, the hybrid architecture provides topology-aware control, enabling the DRL agent to respond more quickly and accurately to changing loads. GNN-enhanced representations in V2X and IoT topologies, with topology dynamics and resource constraints, rank high in reducing ambiguity in state representation and improving policy stability. The hybrid method enhances QoS-conscious resource provisioning in optical networks through structural reasoning, in addition to long-term optimization. Generally, the domain-specific analysis establishes that the hybrid GNN–DRL system is the most transferable, robust, and performance-efficient solution, and it applies to the complex, large-scale, and latency-sensitive networking environments.

Table 7 provides an overview of influential and domain-specific studies that apply GNN and DRL techniques to intelligent end-to-end networking. Each work is classified by network domain, with corresponding insights into how GNNs help extract structural or contextual information and how DRL agents leverage this knowledge to make optimal decisions. In end-to-end routing, for instance, the GNN model dynamically topologies while DRL selects high-throughput paths, demonstrating the synergy between topological learning and adaptive decision-making. In access networks, GCN-based feature extraction enables more accurate channel-condition prediction, supporting DRL-driven resource allocation to improve spectral efficiency. Transport-layer studies blend MPNN-based traffic forecasting or GAT-enabled topology encoding with DRL scheduling or routing policies to reduce delay and boost throughput. Core network efforts integrate DRL logic into SDN controllers enhanced by GNN-derived global states, enabling efficient rule configuration and virtual link management. Edge computing and vehicular networks further showcase hybrid GNN–DRL frameworks that manage offloading, mobility, and cooperative V2X decision-making with reduced latency and improved QoS. IoT network studies extend these ideas to dense, dynamic environments by using GNN-informed states to help DRL agents allocate channels and time slots effectively. Altogether, **Table** 7 demonstrates the scope of applicability of GNN-based state modeling and DRL-based optimization. The allocation shows that this paradigm has been steadily embraced at various levels of end-to-end networking, including inaccessible access, transport, core, and new vehicular and IoT networks, highlighting its applicability as an architecture and not a supremacy in intelligent network design.

4.4. Algorithmic and architectural comparison

Algorithms: Dynamically Random Walk (DRA).

Models Architectures: Graph Neural Network (GNN), Fully Connected (FC) and Hybrid GNN.

Deep Reinforcement Learning has also considered hand-engineered models of network states and sequential learning as a common feature of Deep Reinforcement Learning (DRL)-only systems. There are value-based algorithms (Deep Q-Network (DQN) and Double Deep Q-Network (Double DQN)) and policy-based algorithms (Proximal Policy Optimization (PPO) and REINFORCE) or actor-critic models that are typically trained using these systems.

Despite the high adaptability of these DRL algorithms, they tend to be unstable in convergence and have a poor generalization in the case of network topology and traffic distribution variations. Graph Neural Networks (GNNs), on the other hand, are useful in the graph-structured network setting since they are able to capture graphical connections between nodes and links. Nevertheless, GNNs do not necessarily contain a decision-making process and thus cannot independently optimize the long-term control problems without combining it with reinforcement learning techniques. Hybrid GNN–DRL models leverage these advantages because GNN modules serve as structural encoders, producing topology-sensitive embeddings that are passed to the policy or value networks of the DRL. This offers a more discriminative representation of states, thus providing a larger sample, less training variability, and greater cross-topology generalization. Architecturally, hybrid models introduce hierarchical

Table 7. Selected comprehensive works on intelligent end-to-end networking using GNN and DRL

Network domain	Key remarks	State	Actions	Reward	Reference	Year
End-to-end routing	Integration of GNN-based topology learning with DRL for adaptive path selection	GNN-extracted node/link embeddings representing dynamic topology	Selection of optimal routing path	Maximizing end-to-end throughput	[1]	2024
Access network optimization	DRL agent supported by GCN for channel quality prediction and resource assignment	GCN states encoding CSI, interference, and connectivity	Allocation of transmission power and channel blocks	Improved spectral efficiency and reduced delay	[3]	2022
Transport network scheduling	GNN-assisted traffic forecasting enables DRL to make proactive scheduling decisions	Temporal–spatial traffic states learned via MPNN	Adjusting flow scheduling and queue priority	Average delay reduction across flows	[4]	2022
Transport network routing	Attention-based GNN generates topology-aware embeddings for DRL routing policy	GAT modes capturing the local/global structure of the network graph	Selection of next-hop routing decisions	Path-level throughput improvement	[11]	2024
Core network (SDN)	DRL embedded in the SDN controller with GNN modeling the full network state	SDN global view + GNN graph-state extraction	Optimal configuration of routing rules and virtual links	Maximizing network utility with low overhead	[12]	2023
Edge computing & offloading	Hybrid GNN–DRL framework for computational task offloading at edge nodes	GNN state capturing server load, mobility, link quality	Selecting the offloading node and resource allocation	Reduced energy consumption and latency	[13]	2023
Vehicular networks (V2X)	Multi-agent DRL with GNN interaction graph for cooperative decision-making	Vehicle interaction graph and neighbor influence encoded by GNN	Joint routing, speed adjustment, and lane switching	Stabilized V2X communication and improved QoS	[15]	2023
IoT network management	GNN-enhanced DRL for handling dense IoT deployments	State includes topology evolution, device density, and link status	Allocation of channels and transmission slots	Maximizing QoS under device mobility	[18]	2024

processing pipelines in which spatial reasoning is performed by the GNN and temporal optimization by the DRL module. In algorithms and architectures, the hybrid paradigm consistently shows better robustness, improved learning stability, and more efficient performance, making it the most appropriate choice for next-generation intelligent networking systems.

Table 8 provides a detailed and topic-oriented synthesis of representative studies that apply

Table 8. Representative studies on intelligent end-to-end networking using GNN and DRL

Reference	Contribution type	Key method	Algorithm(s)	Network domain	Summary of contribution
1	Routing optimization	GCN-based topology learning integrated with DRL	GCN + DQN	End-to-end routing	Proposes a GCN-assisted DRL model that extracts topology embeddings and enables intelligent route selection to maximize throughput under dynamic conditions.
2	Resource allocation	Graph-aware channel quality prediction	GNN + PPO	Wireless access networks	Utilizes graph representations of interference and CSI to guide a PPO agent for optimal allocation of power, bandwidth, and channel blocks.
3	Traffic prediction	MPNN for spatio-temporal forecasting	MPNN	Transport networks	Introduces an MPNN model capable of forecasting traffic congestion patterns, enabling proactive end-to-end scheduling decisions.
4	Routing enhancement	Attention-based graph learning	GAT + DRL	Transport layer	Develops a GAT-driven DRL policy that leverages attention-weighted topology information to improve hop-by-hop routing accuracy.
5	SDN control	Global network-state extraction via GNN	GNN + Actor-critic	Core SDN networks	Integrates GNN state modeling with DRL to optimize SDN routing rules, achieving low latency and improved network adaptability.
6	Task offloading	Topology-aware DRL offloading framework	GNN + DDPG	Edge computing (MEC)	Presents a GNN–DDPG architecture to intelligently offload tasks to edge nodes by learning mobility patterns, load distribution, and wireless quality.
7	Cooperative control	Multi-agent learning with GNN interaction graphs	MA-DRL + GNN	Vehicular networks (V2X)	Uses GNN to capture vehicle-to-vehicle relationships and improve coordinated routing, lane switching, and communication reliability.
8	Channel assignment	Structural graph features for dense IoT	GNN + Q-learning	IoT networks	Applies GNN-assisted Q-learning to assign channels efficiently in dense IoT deployments, reducing collisions and enhancing QoS.
9	End-to-end congestion control	Hybrid traffic representation model	GNN + PPO	Multi-domain end-to-end networks	Enhances congestion control by leveraging GNN-derived graph embeddings that represent queue states and path congestion levels.
10	Virtual network mapping	Graph mapping and resource placement using DRL	GCN + DDPG	Virtualized networks (NFV)	Employs GCN embeddings combined with DRL to improve virtual network placement decisions under resource constraints.
11	QoS optimization	GNN support for multi-objective DRL	GNN + A3C	Core networks	Improves QoS metrics (latency, jitter, throughput) through an A3C policy guided by GNN-based topology abstraction.
12	Link failure recovery	GNN-aware DRL for resilience	GAT + DQN	Backbone networks	Uses GAT-generated structural scores to help the DRL agent reroute traffic quickly after link failures.
13	Power control	Graph-based state distribution modeling	MPNN + PPO	Wireless access	Enables energy-efficient power allocation by modeling interference patterns with MPNN and optimizing decisions via PPO.
14	Path computation	Hierarchical GNN encoders	HGNN + DQN	Multi-tier networks	Introduces hierarchical graph encoders that capture core-edge topology for more accurate DRL-based routing.
15	Intelligent load balancing	DRL with neighborhood-aware GNN states	GNN + DDPG	SD-WAN	Uses GNN to capture WAN link dependencies and supports DRL in balancing traffic intelligently across multiple paths.

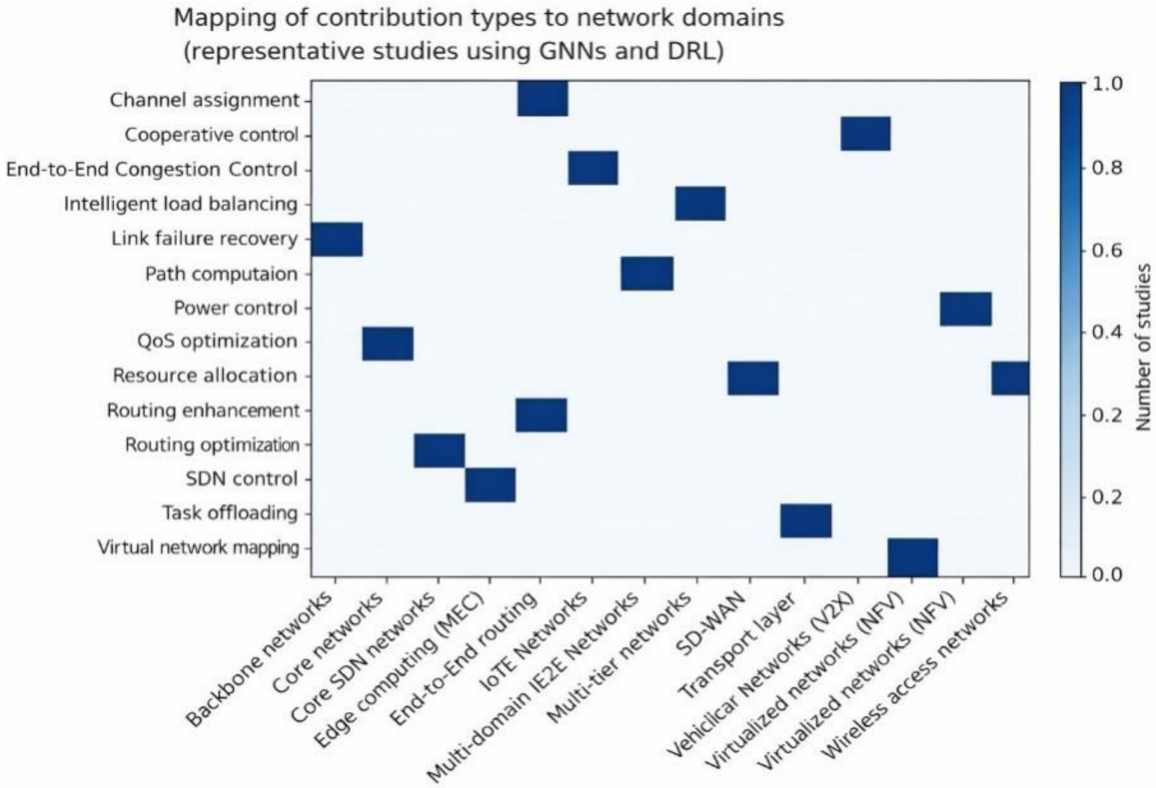

Figure 4. Mapping of contribution types
Abbreviations: DRL: Deep reinforcement learning; E2E: End-to-end; GNN: Graph neural network; IoT: Internet-of-things; MEC: Mobile edge computing; NFV: Network function virtualization; QoS: Quality of service; SD-WAN: Software-defined wide-area network; SDN: Software-defined networking; V2X: Vehicle-to-everything.

GNN and DRL techniques to intelligent end-to-end networking across multiple network domains. Collectively, the table showcases how GNN models—such as GCN, GAT, MPNN, and heterogeneous graph neural network—are used to extract topology-aware, interference-aware, or congestion-aware state representations, while DRL algorithms—such as DQN, PPO, DDPG, actor-critic, A3C, and Q-learning—perform optimal decision-making over these learned representations.

The summarized contributions shown in **Figure** 4 span a wide range of networking challenges: routing optimization through GCN-enhanced DRL; resource allocation in wireless access networks using graph-informed PPO; spatio-temporal traffic prediction via MPNN; transport-layer routing improved by GAT-driven policies; SDN rule optimization using GNN-based global state extraction; intelligent MEC task offloading via GNN–DDPG; cooperative vehicular control using multi-agent DRL with graph representations; IoT channel assignment through GNN-assisted Q-learning; end-to-end congestion control using GNN–PPO embeddings; virtual network mapping using GCN–DDPG; QoS optimization using GNN-supported A3C; resilient rerouting after link failures with GAT–DQN; energy-efficient power control via MPNN–PPO; hierarchical GNN encoders for multi-tier path computation; and intelligent SD-WAN load balancing through GNN–DDPG integration.

4.5. Comparative strengths and limitations

Table 9 provides a comparative evaluation of different model categories—GNN-based, DRL-based, hybrid GNN–DRL, multi-agent GNN–DRL, and hierarchical topology-aware GNN frameworks—used in intelligent end-to-end networking. The table highlights each model type's core capabilities, major strengths, and critical limitations, offering a clear understanding of where each approach excels and where challenges remain.

Table 9 outlines the relative advantages and disadvantages of various learning models for network intelligence applications. GNN-based networks, such as GCN, GAT, and MPNN, are effective models that can capture both structural and relational dependencies within network graphs,

Table 9. Comparative analysis of GNN-, DRL-, and hybrid GNN–DRL-based approaches for intelligent end-to-end networking

Model type	Core capabilities	Scalability	Adaptability	Applications	Strengths	Limitations
GNN-based models (GCN, GAT, MPNN)	Graph representation learning that captures topology structure and node–link relationships	High (handles large graphs efficiently)	Moderate (requires topology updates)	Transport networks, vehicular networks, graph modeling, backbone routing, traffic forecasting	Strong for topology modeling, traffic prediction, and link-state estimation	Lacks autonomous decision-making capability Requires retraining when network topology changes
DRL-based models (DQN, DDPG, PPO, A3C)	Autonomous decision-making through policy learning	Medium (state-space explosion problem)	High (fast real-time adaptation)	IoT scheduling, SDN routing, MEC resource allocation, end-to-end congestion control	Effective for routing, scheduling, resource allocation, and congestion control	Limited structural awareness of network topology Training instability in complex environments
Hybrid GNN–DRL models	Integrates graph-aware state representation with DRL policy optimization	Very high	Very high (handles dynamic and heterogeneous networks)	End-to-end routing, SD-WAN optimization, distributed multi-hop networks, vehicular networks (V2X)	High accuracy for dynamic routing, QoS control, and load balancing	Increased model complexity High computational requirements
Multi-agent GNN–DRL	Models interactions among multiple agents with cooperative control	High in decentralized environments	High (supports multi-node cooperation)	Cooperative V2X, UAV swarms, IoT clusters, distributed SDN controllers	Effective for distributed decision-making and coordination	High communication overhead Difficult training in non-stationary environments Synchronization challenges across agents
Topology-aware hierarchical GNN frameworks	Learns hierarchical structures across core–edge–cloud layers	Very high (supports multi-layer networks)	High (operates across heterogeneous domains)	Core–edge–cloud networks, enterprise WANs, hierarchical IoT, multi-tier operator systems	Improved interpretability, reduced DRL state space, and accurate multi-tier modeling	Error propagation across layers Increased memory consumption

Abbreviations: A3C: Asynchronous advantage actor-critic; DDPG: Deep deterministic policy gradient; DQN: Deep Q-network; DRL: Deep reinforcement learning; GAT: Graph attention network; GCN: Graph convolutional network; GNN: Graph neural network; IoT: Internet-of-things; MEC: Mobile edge computing; MPNN: Message passing neural network; PPO: Proximal policy optimization; QoS: Quality of service; SD-WAN: Software-defined wide-area network; SDN: Software-defined networking; UAV: Unmanned aerial vehicles; V2X: Vehicle-to-everything; WAN: Wide area network.

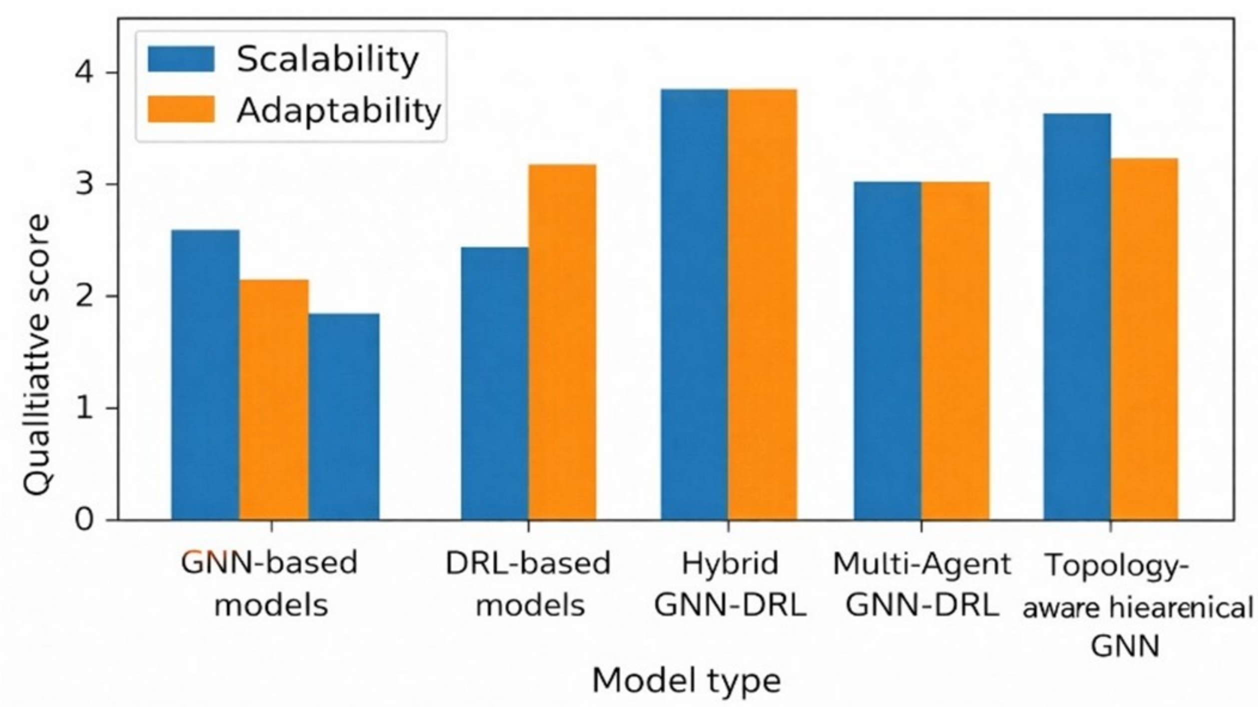

Figure 5. Comparative scalability and adaptability of GNN, DRL, and hybrid-based network intelligence models
Abbreviations: DRL: Deep reinforcement learning; GNN: Graph neural network.

thereby properly modeling topology, predicting traffic, and estimating link states. However, they cannot make sequential decisions independently and need retraining to keep up with changes in network topology. Conversely, the use of DRA-based methods, including DQN, DDPG, PPO, and A3C, is highly efficient for real-time decision-making, including routing, scheduling, resource allocation, and congestion control, with minimal structural awareness and potential instability in multi-dimensional or complex state spaces.

Figure 5 presents a qualitative analysis of the scalability and adaptability of various categories of intelligent networking models: GNN-based, DRL-based, hybrid GNN–DRL, multi-agent GNN–DRL, and topology-aware hierarchical GNN models. The findings indicate that GNN-based approaches have moderate scalability but low adaptability, as they are more predictive than adaptive, exhibit weaker topological trends, and require larger networks. In contrast, DRL-based approaches offer greater adaptability but lower scalability for large and dynamic networks. Hybrid GNN–DRL models exhibit more balanced performance, with significantly higher adaptability and greater scalability, combining the strong capabilities of topology-sensitive representation learning and reward algorithms. Multi-agent GNN–DRL systems offer good configurability and moderate scalability, but the coordination overhead grows with network size. Topology-aware hierarchical GNN architectures, in turn, are the most scalable, offering multi-level graph abstraction and high adaptability. In general, the comparison shows that hierarchical and hybrid architectures are the better-performing approaches for large-scale, dynamic, and autonomous end-to-end networking.

5. Future research directions

Some future research directions are listed as follows:

(i) Develop standardized and open benchmarking suites that enable fair comparisons across routing, congestion, vehicular networks, IoT, and optical domains.
(ii) Train lightweight GNN controllers for real-time operation with lower inference cost and latency, ready for deployment.
(iii) Advance generalization mechanisms across topologies to enable model transfer between networks without retraining.
(iv) Combine robustness to ensure partial observability, dynamic traffic, link failures, and sensor noise.
(v) Add privacy-sensitive and safe learning levels to mitigate adversarial manipulation, bias, and data leakage in decision-making processes.

(vi) Make GNN models interpretable and introduce decision modules composed of DRL components that give clear policy explanations rather than control pipelines that are not interpretable.
(vii) Standardize benchmarking models to facilitate comparability based on latency, throughput, and scalability between intelligent networking solutions on a common quantitative basis.

6. Conclusion

This review provides a solid and thorough overview of intelligent end-to-end networking systems enabled by both GNNs and DRL. It collected the existing literature, the developmental trends in architecture, and empirical findings, and showed the interrelations between these two complementary paradigms in creating autonomous and adaptive networking. The review shows that hybrid GNN–DRL models achieve superior long-term results compared to single-paradigm solutions by leveraging GNNs to provide topology-sensitive representations and DRL to execute real-time, reward-based control across a variety of tasks, including routing, congestion reduction, resource allocation, and service coordination. The findings also indicate that GNN–DRL models are better in terms of generalization, scalability, robustness, better latency-throughput trade-offs, better resilience to congestion, and faster convergence. All of these benefits demonstrate the fundamental importance of state encoding in graph form to improve the stability, sample efficiency, and flexibility of DRL-based controllers. According to the review, a communication framework for homogeneous closed-loop architectures that coordinate GNN-based perception and DRL-based control in a feedback loop is needed.

Despite the development, some difficulties remain, including scaling hybrid models to ultra-large topologies, achieving real-time inference under very strict latency constraints, integrating cross-layer and multi-domain information, addressing security vulnerabilities, and establishing a standard benchmarking infrastructure. Addressing these drawbacks will be important for deploying GNN–DRL systems across emerging 5G/6G networks, IoT systems, optical systems, vehicular networks, and edge computing. Finally, the combination of GNNs with DRL represents a paradigm shift toward completely self-contained, self-optimizing, and self-protecting networking ecosystems. The knowledge, categories, and study lines outlined in the presented review provide a solid foundation for future work to develop scalable and reliable AI-based networking systems that could meet the requirements of next-generation communication systems.

Acknowledgments

None.

Funding

This work was funded by the Faculty of Engineering and Quantity Surveying, INTI International University Colleges, Nilai, Negeri Sembilan, Malaysia.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Muhammad Zain ul Abideen, Junaid Asghar
Formal analysis: Muhammad Kamran, Muhammad Farman, Salwa Muhammad Akhtar, Aseel Smerat
Investigation: Muhammad Zain ul Abideen, Junaid Asghar, Aseel Smerat
Methodology: Muhammad Zain ul Abideen, Junaid Asghar
Writing–original draft: Muhammad Zain ul Abideen, Junaid Asghar
Writing–review & editing: Muhammad Kamran, Muhammad Farman, Mohamad Hafez, Salwa Muhammad Akhtar

Availability of data

Not applicable.

AI tools statement

All authors confirm that no AI tools were used in the preparation of this manuscript.

References

1. Ksentini A, Nikaein N. Toward enforcing network slicing on RAN: flexibility and resource abstraction. *IEEE Commun Mag.* 2017;55(6):102-108. https://www.doi.org/10.1109/MCOM.2017.1601119
2. Giordani M, Polese M, Mezzavilla M, Rangan S, Zorzi M. Toward 6G networks: use cases and technologies. *IEEE Commun Mag.* 2020;58(3):55-61.

https://www.doi.org/10.1109/MCOM.001.1900411

3. Chen Z, Ma X, Zhang C, Wen Z, Li L. Terahertz wireless communications for 2030 and beyond: a cutting-edge frontier. *IEEE Commun Mag.* 2021;59(11):66-72. https://www.doi.org/10.1109/MCOM.011.2100195
4. Gong S, Lu X, Hoang DT, et al. Toward smart wireless communications via intelligent reflecting surfaces: a contemporary survey. *IEEE Commun Surv Tutor.* 2020;22(4):2283-2314. https://www.doi.org/10.1109/COMST.2020.3004197
5. Gkarmpounis G, Vranis C, Vretos N, Daras P. Survey on graph neural networks. *IEEE Access.* 2024;12:128816-128832. https://www.doi.org/10.1109/ACCESS.2024.3456913
6. Akyildiz IF, Kak A, Nie S. 6G and beyond: the future of wireless communications systems. *IEEE Access.* 2020;8:133995-134030. https://www.doi.org/10.1109/ACCESS.2020.3010896
7. Tam P, Ros S, Song I, Kang S, Kim S. A survey of intelligent end-to-end networking solutions: integrating graph neural networks and deep reinforcement learning approaches. *Electronics.* 2024;13(5):994. https://www.doi.org/10.3390/electronics13050994
8. Li X, Chen M, Liu Y, Wang L. Federated multi-agent deep reinforcement learning for resource allocation of vehicle-to-vehicle communications. *IEEE Trans Veh Technol.* 2022;71(8):8810-8824. https://www.doi.org/10.1109/TVT.2022.3173057
9. Chen H, Wang J, Li D, Zhang Z. A tutorial on terahertz-band localization for 6G communication systems. *IEEE Commun Surv Tutor.* 2022;24(3):1780-1815. https://www.doi.org/10.1109/COMST.2022.3178209
10. Yang X, Chen J, Wang H, Liu X. A survey on smart agriculture: development modes, technologies, and security and privacy challenges. *IEEE/CAA J Autom Sinica.* 2021;8(2):273-302. https://www.doi.org/10.1109/JAS.2020.1003536
11. Żarski M, Wysocki T, Kulesza J. Computer vision-based inspection on post-earthquake with UAV synthetic dataset. *IEEE Access.* 2022;10:108134-108144. https://www.doi.org/10.1109/ACCESS.2022.3212918
12. Alencar D, Barreto R, Santos A. Dynamic microservice allocation for virtual reality distribution with QoE support. *IEEE Trans Netw Serv Manag.* 2022;19(1):729-740. https://www.doi.org/10.1109/TNSM.2021.3076922
13. Xiao L, Zhang Y, Li H. A segmented variable-parameter ZNN for dynamic quadratic minimization with improved convergence and robustness. *IEEE Trans Neural Netw Learn Syst.* 2023;34(5):2413-2424. https://www.doi.org/10.1109/TNNLS.2021.3106640
14. Salam MA, Azar AT, Hussien R. Swarm-based extreme learning machine models for global optimization. *Comput Mater Contin.* 2022;70(3). https://www.doi.org/10.32604/cmc.2022.020583
15. Ding Z, Feng B, Jiang C. COIN: a container workload prediction model focusing on common and individual changes in workloads. *IEEE Trans Parallel Distrib Syst.* 2022;33(12):4738-4751. https://www.doi.org/10.1109/TPDS.2022.3202833
16. Naderializadeh N, Avestimehr AS, Jafar SA. Resource management in wireless networks via multi-agent deep reinforcement learning. *IEEE Trans Wireless Commun.* 2021;20(6):3507-3523. https://www.doi.org/10.1109/TWC.2021.3051163
17. Tam P, Song I, Kang S, Ros S, Kim S. Graph neural networks for intelligent modelling in network management and orchestration: a survey on communications. *Electronics.* 2022;11(20):3371. https://www.doi.org/10.3390/electronics11203371
18. Hu W, Chen J, Zhang S. Graph signal processing for geometric data and beyond: theory and applications. *IEEE Trans Multimed.* 2022;24:3961-3977. https://www.doi.org/10.1109/TMM.2021.3111440
19. He Q, Liu L, Zhang J. Routing optimization with deep reinforcement learning in knowledge-defined networking. *IEEE Trans Mob Comput.* 2024;23(2):1444-1455. https://www.doi.org/10.1109/TMC.2023.3235446
20. Li M, Li H. Application of deep neural network and deep reinforcement learning in wireless communication. *PLoS One.* 2020;15(7):e0235447. https://www.doi.org/10.1371/journal.pone.0235447
21. Casas-Velasco DM, Rendon OMC, da Fonseca NLS. Intelligent routing based on reinforcement learning for software-defined networking. *IEEE Trans Netw Serv Manag.* 2021;18(1):870-881. https://www.doi.org/10.1109/TNSM.2020.3036911
22. Zhao Z, Wang X, Liu Y. A transmission-reliable topology control framework based on deep reinforcement learning for UWSNs. *IEEE Internet Things J.* 2023;10(15):13317-13332. https://www.doi.org/10.1109/JIOT.2023.3262690.
23. Suárez-Varela J, Ferriol-Galmés M, López A, et al. The graph neural networking challenge: a worldwide competition for education in AI/ML for networks. *ACM SIGCOMM Comput Commun Rev.* 2021;51(3):9-16. https://www.doi.org/10.1145/3477482.3477485

24. Wu Z, Pan S, Chen F, et al. A comprehensive survey on graph neural networks. *IEEE Trans Neural Netw Learn Syst*. 2021;32(1):4-24. https://www.doi.org/10.1109/TNNLS.2020.2978386
25. Ji M, Zhang Y, Li X. Graph neural networks and deep reinforcement learning based resource allocation for V2X communications. *IEEE Internet Things J*. 2024;12(4):3613-3628. https://www.doi.org/10.1109/JIOT.2024.3469547
26. Khemani B, Patil S, Kotecha K, Tanwar S. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. *J Big Data*. 2024;11(1):18. https://www.doi.org/10.1186/s40537-023-00876-4
27. Munikoti S, Nunes I, Rao A. Challenges and opportunities in deep reinforcement learning with graph neural networks: a comprehensive review of algorithms and applications. *IEEE Trans Neural Netw Learn Syst*. 2024;35(11):15051-15071. https://www.doi.org/10.1109/TNNLS.2023.3283523
28. Jian C, Wang Y, Zhang L. Online-learning task scheduling with GNN-RL scheduler in collaborative edge computing. *Cluster Comput*. 2024;27(1):589-605. https://www.doi.org/10.1007/s10586-022-03957-w
29. Lai Y, Liu H, Zhang Q. Toward adversarially robust recommendation from adaptive fraudster detection. *IEEE Trans Inf Forensics Secur*. 2024;19:907-919. https://www.doi.org/10.1109/TIFS.2023.3327876
30. Han J, Cen J, Wu L, et al. A survey of geometric graph neural networks: data structures, models and applications. *Front Comput Sci*. 2025;19(11):1911375. https://www.doi.org/10.1007/s11704-025-41426-w
31. Idris NF, Ismail MA, Kasim S, et al. A review of feature selection methods on diabetes mellitus classification. *Int J Adv Sci Eng Inf Technol*. 2025;15(3):686-692. https://www.doi.org/10.18517/ijaseit.15.3.12652
32. Yogeesh N, Mohammad SI, Raja N, et al. From crisp to fuzzy: a comparative review of statistical and fuzzy approaches to problem solving. *Appl Math Inf Sci*. 2025;19(3):647-658. https://www.doi.org/10.18576/amis/190313
33. Al-Daoud KI, Yogeesh N, Mohammad SI, et al. Explainability in AI using fuzzy inference systems for the regression problem. *Appl Math Inf Sci*. 2025;19(5):973-987. https://www.doi.org/10.18576/amis/190501

Muhammad Kamran *holds a PhD in Mathematics, with research focusing on computational and applied mathematics, as well as computational intelligence. He completed a postdoctoral fellowship at Uludağ University, Turkey. Dr. Kamran has also received fellowships from INTI International University, Malaysia, and Near East University, Turkey. His research experience spans fuzzy systems, intelligent systems, artificial intelligence, and graph networks for transportation.*
https://orcid.org/0009-0000-5467-0497

Salwa Muhammad Akhtar *is a faculty member in the Department of Information Systems at the University of Management and Technology, Lahore, Punjab, Pakistan. She is actively involved in research on machine learning models for various types of networks.*
https://orcid.org/0000-0002-2068-9939

Muhammad Zain ul Abideen *is an active researcher and faculty member in the Department of Mechanical Engineering, Faculty of Engineering, at the University of Central Punjab, Lahore, Punjab, Pakistan. His work focuses on machine learning and artificial intelligence models for optimization and graphs.*
https://orcid.org/0009-0008-8557-2628

Junaid Asghar *is a Computer Science professional and researcher specializing in Machine Learning, Deep Learning, and AI-driven Computer Vision. He currently serves as a Lecturer at The University of Lahore, balancing teaching with high-impact research on suspicious activity detection and agricultural AI frameworks. He holds an MS in Computer Science from UMT, where his thesis focused on deep learning for skin disease detection. Junaid was named "Best Teacher of the Year" at NCBA&E and has extensive experience managing MIS and SAP systems. Proficient in WordPress and Magento, and exploring Reinforcement Learning, he is dedicated to developing innovative algorithms for real-world problems while mentoring future tech talent.*
https://orcid.org/0009-0006-3207-3310

Muhammad Farman *earned his Ph.D. in Mathematics from the University of Lahore, Pakistan, in 2019. He completed postdoctoral fellowships at Universitas Airlangga Indonesia, Lebanese American University, and Near East University (North Cyprus, Turkey). He serves as an Adjunct Professor (Research) at Universitas Airlangga and the University of Lahore, and is currently an Associate Professor at Near East University. With over eight years of teaching and research experience, his work focuses on mathematical biology, control theory, numerical analysis, and fractional calculus applied to epidemic models. He has published over 300 papers in reputed SCI and Scopus journals.*

Dr. Farman has received multiple awards, including Research Excellence Awards (2019, 2020), Stanford's Top 2% Scientist listing (2023, 2024), and the Best Academic Performance, Publication Honorary, and Most Impactful Researcher Awards (2024) from Near East University.
https://orcid.org/0000-0001-7616-0500

Aseel Smerat *is a professor in the Department of Mathematics, Faculty of Educational Sciences, at Al-Ahliyya Amman University, Amman, Jordan. He also serves as adjunct faculty in the Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India.*
https://orcid.org/0009-0008-4600-509X

Mohamad Hafez *is a faculty member in the Department of Mathematics, Faculty of Engineering and Quantity Surveying, at INTI International University Colleges, Nilai, Negeri Sembilan, Malaysia. He is an active researcher involved in various top-tier research projects.*
https://orcid.org/xxxx-xxxx-xxxx-xxxx

An International Journal of Optimization and Control: Theories & Applications (https://accscience.com/journal/ijocta)