AI-Driven Adaptive Control and Optimization in Next-Generation Telecommunication Networks: Architectures, Algorithms, and Autonomous Fault Management for 5G/6G and Beyond

Abstract

This comprehensive survey delineates the transformative integration of artificial intelligence (AI) within adaptive control, telecommunications, and dynamic networking systems, emphasizing its pivotal role in advancing next-generation communication infrastructures such as 5G, 6G, and beyond. Motivated by escalating data volumes, heterogeneous device ecosystems, and stringent service demands, the work explores a broad spectrum of AI methodologies—including reinforcement learning, deep learning, federated learning, and gradient-based optimization—applied to critical domains like network traffic classification, software-defined networking (SDN), routing optimization, Open Radio Access Network (Open RAN), and autonomous fault management.

Key contributions include an in-depth examination of AI-driven adaptive traffic classification techniques that overcome traditional limitations posed by encryption and dynamic traffic patterns, highlighting trade-offs between accuracy, computational complexity, and real-time feasibility. The survey further analyzes AI-empowered SDN architectures that enhance resource allocation and anomaly detection, discussing scalability and security challenges alongside prospects for decentralized, privacy-preserving learning in 6G deployments. AI-based routing optimization is reviewed with a focus on reinforcement learning algorithms augmented by traffic prediction and anomaly detection, evidencing significant throughput and latency enhancements. Open RAN integration elucidates multilayer AI deployment for radio and network layer optimization, underscoring federated learning and hybrid communication modalities for improved performance and resilience. The incorporation of Large Language Model (LLM)-based agentic AI for autonomous fault management within O-RAN frameworks is also detailed, demonstrating substantial gains in fault detection accuracy, mitigation efficiency, and network uptime. Complementing these, the survey addresses AI-enhanced wireless networking elements such as reconfigurable intelligent surfaces (RIS) and perceptive mobile networks (PMNs), which benefit from advanced AI techniques for interference management and sensing.

The work critically appraises challenges enveloping computational overhead, latency constraints, data heterogeneity, privacy, interpretability, interoperability, and robustness against adversarial threats. It advocates scalable, distributed AI architectures combining edge-cloud synergy, federated and multi-agent learning paradigms, and explainable AI techniques to foster transparency, trust, and regulatory compliance. Gradient-based optimization methods and fast algorithmic updates are presented as foundational tools to enable real-time system adaptability in complex, stochastic network environments.

Concluding, the survey synthesizes cross-cutting themes and prospective research avenues—including hardware acceleration, quantum computing, blockchain-enhanced security, and multi-agent collaborative learning—that collectively underpin the evolution of autonomous, resilient, and intelligent telecommunication networks. By providing a holistic and rigorous exploration of AI-enabled adaptive control and networking, this work lays a robust foundation for future scholarly and practical advancements striving towards secure, scalable, and transparent AI integration in dynamic communication ecosystems.

ACM Reference Format:

. 2025. AI-Driven Adaptive Control and Optimization in Next-Generation Telecommunication Networks: Architectures, Algorithms, and Autonomous Fault Management for 5G/6G and Beyond. In . ACM, New York, NY, USA, 52 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Artificial Intelligence (AI) has undergone significant advancements over recent decades, impacting various domains such as healthcare, finance, and autonomous systems [**?**]. Despite these achievements, ongoing challenges remain in three key areas: scalability, interpretability, and integration with human decision-making [**?**]. Throughout this survey, we critically examine existing methodologies by focusing on their strengths, limitations, and suitability for different applications.

To provide a clear overview, this survey is organized as follows. Section **??** reviews prominent AI techniques, including machine learning, neural networks, and symbolic reasoning, highlighting their typical applications. Section **??** discusses the main challenges faced in AI research and practice, such as scalability issues with large datasets, the need for interpretability to ensure transparency, and methods for integrating AI systems with human decision-making processes. Finally, Section **??** explores future directions and potential solutions to these challenges.

Table 1 summarizes the main contributions and structure of this survey.

For readers less familiar with some technical terms, we provide brief explanations where these terms first appear. For example, "scalability" refers to an AI system's ability to maintain performance as the size of data or complexity of tasks increases, while "interpretability" means how easily humans can understand the reasoning behind AI outputs. This approach aims to make the survey accessible to a broader audience.

Conference'17, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM

https://doi.org/10.1145/nnnnnnn.nnnnnnn

This structured breakdown and the inclusion of clarifying examples throughout the paper should help guide the reader effectively through the landscape of AI research.

1.1 Current AI Methodologies

Deep learning approaches have demonstrated exceptional performance on large-scale datasets, enabling breakthroughs in perception and pattern recognition [**?**]. These methods rely on multiple layers of nonlinear processing units to automatically extract hierarchical features from raw data.[1] However, they often lack transparency due to their black-box nature and demand substantial computational resources, which hinders their deployment in resource-constrained or high-stakes settings requiring interpretability and efficiency. Moreover, their performance can degrade when faced with data distributions differing from those seen during training.

Symbolic AI techniques offer superior interpretability and align well with human reasoning processes [**?**]. These methods operate on explicit rules and logic, facilitating explanation and verification. Yet, their scalability and adaptability to complex, unstructured data remain limited, restricting their effectiveness across a broad range of real-world applications where flexibility and learning from raw data are crucial. In particular, symbolic systems struggle with noisy or incomplete data and lack the capability to learn representations autonomously.

Hybrid models aim to leverage the complementary advantages of both paradigms by integrating symbolic reasoning with deep learning [**?**]. Despite their promise, such integration is nontrivial due to challenges in harmonizing fundamentally different representations and learning mechanisms. These include the need to reconcile symbolic logic's discrete structures with neural networks' continuous representations, as well as ensuring end-to-end differentiability for training efficiency.[2] While they seek to combine interpretability, adaptability, and learning capabilities, the complexity of designing and training hybrid architectures currently limits their widespread application.

1.2 Scope and Contributions

This survey synthesizes findings across a diverse range of AI techniques, providing a detailed comparative analysis that elucidates their unique capabilities, inherent trade-offs, and situational advantages. By systematically highlighting open research challenges and prospective future directions, it offers valuable guidance for researchers aiming to select, adapt, or advance AI methodologies tailored to specific application contexts and requirements.

1.3 Overview of AI-Driven Approaches in Adaptive Control, Telecommunications, and Networking Systems

The integration of artificial intelligence (AI) into adaptive control, telecommunications, and dynamic networking systems has catalyzed unprecedented advancements, fundamentally reshaping traditional paradigms by introducing data-driven adaptability and autonomous decision-making. Foundational studies have demonstrated AI's potential in optimizing networks via reinforcement learning, enabling autonomous control mechanisms within communication systems, and developing adaptive AI models designed for dynamic protocol adjustment [27, 28, 33**?**]. These approaches leverage the inherent dynamics of networks by utilizing system state information alongside historical interactions, thereby empowering networks to self-optimize under diverse and time-varying conditions [36**?** **?**]. For instance, semantic communication frameworks that combine deep learning with knowledge graphs enable context-aware, efficient transmission by extracting and reconstructing semantic information, substantially enhancing communication reliability and semantic fidelity [36]. Additionally, deep learning techniques have been effectively applied to autonomously manage network parameters, detect anomalies, and predict system behaviors within telecom Self-Optimized Networks (SON), improving fault management and resource allocation [**?**].

Furthermore, AI techniques have addressed the complexities presented by distributed and heterogeneous network infrastructures, effectively tackling challenges such as resource contention, delay variability, and fault tolerance [2**?** **?**]. Federated learning frameworks incorporating gradient sparsification, adaptive client selection, and joint bandwidth allocation optimize collaborative model training across wireless devices, balancing communication efficiency and learning accuracy under resource constraints [2**?**]. Reinforcement and federated learning methods enhance network slicing in next-generation wireless networks, enabling dynamic resource allocation and slice admission control to support diverse service requirements with improved throughput and latency [**?**]. Moreover, advancements in mobility management schemes within proxy mobile IPv6 domains have demonstrated significant reductions in signaling overhead and handover latency through hierarchical gateway structures and optimized signaling, thereby enhancing network performance and user experience [33].

Despite significant progress, persistent challenges remain, notably in managing computational overhead, sustaining real-time inference under tight latency requirements, and ensuring robustness against network uncertainties and adversarial perturbations [20**?** **?**]. Large language models (LLMs) in telecommunications exemplify these challenges, requiring efficient deployment strategies and domain-specific adaptations to balance computational demands, privacy, and accuracy [**?**]. The breadth of AI integration spans from automated network traffic classification to autonomous fault management, covering both physical-layer optimization and higher-layer protocol adaptation [6, 7, 24]. Notably, AI-driven network management requires scalable frameworks that maintain accuracy while meeting stringent latency and reliability demands intrinsic to emerging applications like the Tactile Internet, which demands ultra-low latency and high reliability for real-time haptic communications [7].

A comparative evaluation of these diverse AI-driven approaches reveals notable trade-offs, summarized in Table 3. For example, semantic communication methods [36] achieve superior semantic fidelity and noise robustness but incur substantial computational overhead and complexities in dynamic knowledge updating. Deep

[1]Deep learning models typically consist of architectures such as convolutional neural networks (CNNs) for image data and recurrent neural networks (RNNs) for sequential data, allowing them to capture complex patterns without manual feature engineering.
[2]Hybrid approaches often employ mechanisms such as neural-symbolic integration frameworks or differentiable logic layers to bridge these gaps.

Table 1: Summary of Contributions and Survey Structure

Section	Content Overview
??	Review of AI techniques: machine learning, neural networks, symbolic reasoning, and their applications
??	Discussion of key challenges: scalability, interpretability, and human-AI integration
??	Exploration of future research directions and potential solutions to identified challenges

Table 2: Comparative Summary of AI Methodologies: Characteristics, Limitations, and Typical Use Cases

Methodology	Key Characteristics	Limitations	Typical Use Cases
Deep Learning	Hierarchical feature learning; end-to-end training; high accuracy on large datasets	Black-box model; high computational cost; limited interpretability	Image recognition; speech processing; natural language understanding
Symbolic AI	Rule-based reasoning; interpretable and explainable decisions	Poor scalability; struggles with unstructured/noisy data; limited learning capability	Expert systems; knowledge representation; automated reasoning
Hybrid Models	Combines symbolic reasoning with neural networks; aims for interpretability and adaptability	Complex integration; training challenges; balancing discrete and continuous representations	Cognitive computing; explainable AI; tasks requiring both learning and reasoning

learning techniques applied in SON [**?**] enhance autonomous management capabilities but face challenges in model robustness across heterogeneous network environments and increased resource consumption. Federated learning frameworks [2**?**] effectively balance data privacy and communication efficiency, yet may encounter scalability and synchronization issues in highly dynamic wireless settings. Reinforcement learning applied to network slicing [**?**] provides adaptive resource allocation with improved throughput and latency but often requires extensive training data and careful tuning to avoid convergence issues.

Moreover, specific limitations and pitfalls need explicit consideration. The hierarchical mobility management strategies [33], while reducing signaling overhead, may face scalability concerns in ultra-dense networks with high mobility users, and integrating them seamlessly with standardized protocols remains nontrivial. The deployment of LLMs [**?**] in telecom environments must carefully manage the trade-off between inference latency, model size, and privacy, with existing methods still grappling with domain adaptation and hallucination mitigation. Finally, AI-driven frameworks for emerging ultra-reliable low-latency networks, such as the Tactile Internet [7], demand stringent real-time guarantees, exposing tensions between AI model complexity and latency requirements.

Overall, this evolving landscape underscores the importance of advancing scalable, interpretable, and resource-aware AI frameworks that effectively balance performance gains against computational costs, latency constraints, and deployment complexity. Future research should focus on developing hybrid models combining the strengths of multiple AI paradigms, enhancing robustness to network uncertainties, and integrating explainability to foster trust and effective operation within mission-critical telecommunication systems.

1.4 Motivation for AI Integration

The telecommunications and networking sectors are witnessing accelerated growth characterized by increasing data volume, heterogeneous device ecosystems, and complex service requirements, especially within vector databases, wireless networking infrastructures, software-defined networking (SDN), and Open Radio Access Network (Open RAN) architectures [1, 30, 37]. For instance, advances in 5G IoT random access leverage novel high-capacity preamble sequences that substantially increase unique preambles while minimizing collision probability through combinatorial design and sequence theory, thereby handling massive device connectivity with reduced access delay and retransmission rates [37]. Similarly, in cyber forensic investigations within IoT, deep learning-based feature fusion methods effectively preprocess heterogeneous data sources and combine spatial and temporal features via convolutional and recurrent neural networks to improve security analysis accuracy, precision, and recall [1]. Additionally, obstacle-aware fuzzy clustering protocols for wireless sensor networks in complex 3D terrains utilize fuzzy inference systems to select cluster heads considering residual energy, distance to sink, and obstacle factors, thereby enhancing network lifetime and reliability [30]. The transition to 5G/6G and beyond-6G (B6G) technologies demands adaptive, intelligent mechanisms capable of managing escalating complexity, enabling dynamic resource allocation, and optimizing real-time performance [22, 26].

AI techniques have demonstrated substantial benefits in these areas by facilitating model-free, context-aware decisions that optimize network slicing, enhance spectrum utilization, and enable hybrid fusion strategies such as combining visible light communication (VLC) and radio frequency (RF) systems [13, 16]. For example, machine learning-based network traffic classification models employ diverse supervised and deep learning algorithms trained on flow and statistical features to improve accuracy and adaptability despite encrypted, dynamic traffic patterns. Such models address challenges like data imbalance and concept drift, achieving high accuracy and robustness in traffic classification [16]. Similarly, AI-powered SDN frameworks integrate supervised classifiers and deep learning models (e.g., LSTM networks) into SDN controllers, enabling real-time traffic classification, anomaly detection, and dynamic resource allocation. Experimental validations demonstrate up to 92% accuracy in traffic classification, 18% reduction in latency, and a 15% increase in throughput within 5G environments, significantly enhancing network management and resource efficiency [13]. Moreover, AI-enabled sophisticated detection methods and adaptive interference cancellation schemes effectively mitigate wireless channel impairments, improving reliability and throughput in wireless communications [5, 25].

Real-world validations of these AI approaches affirm their practical applicability while highlighting ongoing challenges related to data heterogeneity, privacy preservation, computational overhead, and smooth integration with legacy systems. For instance, SDN frameworks still face computational challenges and dataset scarcity,

Table 3: Comparative Evaluation of AI-Driven Approaches in Adaptive Control, Telecommunications, and Networking

AI Approach	Strengths	Challenges	Applications	References
Semantic Communication	High semantic fidelity; noise robustness	High computational overhead; dynamic knowledge updating	Context-aware transmission; semantic error correction	[36]
Deep Learning in SON	Autonomous network management; fault detection	Model robustness across heterogeneous networks; resource consumption	Network parameter tuning; anomaly detection	[?]
Federated Learning Frameworks	Preserves data privacy; communication efficient	Scalability; synchronization in dynamic environments	Collaborative model training across wireless devices	[2?]
Reinforcement Learning for Network Slicing	Adaptive resource allocation; improved throughput and latency	Requires large training data; tuning for convergence	Dynamic slice admission control and resource management	[?]
Hierarchical Mobility Management	Reduced signaling overhead; lower handover latency	Scalability in ultra-dense/high mobility; protocol integration	Mobility management in proxy mobile IPv6 networks	[33]
Large Language Models (LLMs)	Telecom automation; knowledge generation	Inference latency; privacy; domain adaptation; hallucinations	Network configuration; traffic classification; optimization	[?]
AI-Driven Tactile Internet Frameworks	Ultra-low latency; high reliability	Balancing AI complexity and latency	Real-time haptic communication; ultra-reliable low-latency networks	[7]

necessitating development of lightweight and privacy-preserving AI models [13]. AI integration in Open RAN architectures notably enhances throughput, latency, and energy efficiency by employing federated learning, reinforcement learning, and deep neural networks for spectrum management, fault detection, and interference mitigation, yet demands resolution of issues such as real-time inference latency, AI model convergence, multi-vendor interoperability, and energy constraints at the network edge [5, 25]. Quantitative evaluations show enhancements in fault detection accuracy up to 95% and mitigation success rates reaching 91%, which significantly reduce downtime and throughput degradation in Open RAN environments [5].

Consequently, AI integration is driven not only by the pursuit of performance enhancements but also by the imperative to endow networks with self-adaptive intelligence essential for addressing the demands and uncertainties inherent to next-generation telecommunication ecosystems.

1.5 Key AI Techniques and Their Roles

A diverse array of AI methodologies has been harnessed to enhance adaptability and performance within communication networks. Reinforcement learning (RL) constitutes a foundational technique for dynamic resource management, enabling agents to learn optimal policies related to bandwidth allocation, routing, and scheduling through continuous interaction with the environment, without requiring explicit environment modeling [32, 35]. These autonomic approaches facilitate self-configuration, self-optimization, self-healing, and self-protection by embedding flexibility and adaptability into software-driven network infrastructures, particularly leveraging software-defined networking (SDN) and network function virtualization (NFV). While such frameworks significantly improve network resilience, throughput, latency, and operational cost, challenges such as scalability of control planes, device heterogeneity, interpretability of machine learning models for real-time decision-making, and security vulnerabilities remain [35]. Addressing these issues calls for enhanced AI and big data analytics, adaptive trust and security mechanisms, comprehensive standardization for interoperability, and large-scale experimental validation to ensure robust network autonomy.

Gradient-based optimization methods, including stochastic dual gradient techniques and their variants, further enable efficient parameter updates in resource-constrained settings while offering provable convergence and queue stability for large-scale network control problems [**? ?**]. These approaches are crucial for optimizing complex resource allocation and scheduling tasks required in beyond-5G (B5G) and 6G wireless systems, supporting ultra-low latency, massive connectivity, and improved spectral and energy efficiency. Such optimization techniques also assist in integrating emerging technologies like optical wireless communications, ensuring robustness despite diverse channel conditions and network densification [**?**]. The optimization frameworks are integral to realizing key technologies such as cell-free massive MIMO and hybrid spectrum sharing envisioned in B5G systems, balancing performance with computational and hardware constraints.

Rapid algorithmic updates, often leveraging modular and distributed architectures, permit real-time adaptability essential for environments characterized by fluctuating traffic patterns and volatile channel conditions [**?**]. These approaches incorporate security frameworks and policy-based rule enforcement critical for defending against vulnerabilities in edge and cloud network deployments. Through simulation and practical validations, such frameworks have demonstrated effectiveness in mitigating network attacks and enhancing both internal and external security postures, thereby improving overall network resilience.

Moreover, intelligent wireless technologies incorporating AI have demonstrated substantial improvements at the physical layer. AI-powered reconfigurable intelligent surfaces (RIS) dynamically control the wireless environment by enabling adaptive channel estimation, beamforming, and resource allocation through learned environmental feedback [6, 38]. This integration boosts spectral and energy efficiency in complex wireless settings and adapts effectively to imperfect channel information. Deep learning-driven interference management frameworks enable robust signal processing against noise and interference, while semantic communications further optimize information flow. Deep learning models such as stacked autoencoders and deep neural networks are applied beyond the physical layer for higher-level tasks including customer churn prediction and traffic management [38]. These models efficiently extract hierarchical features from raw data, facilitating robust decision-making and improved operational performance, although challenges related to computational cost, data requirements, and interpretability persist.

Collectively, these AI techniques form a multi-layered intelligence framework that integrates decision-making from physical-layer signal optimization to network-layer control and application-specific adaptations. This cohesive approach advances autonomous, reliable, and efficient future communication networks, addressing many of the challenges anticipated in next-generation wireless systems [6, 24, 35]. Future research directions emphasize the development of explainable AI models, the synergy between edge and cloud computing, and enhanced robustness against adversarial and uncertain conditions to fully realize intelligent and secure network ecosystems.

1.6 Challenges in AI-Enabled Networking

Despite these technological advances, AI-enabled networking faces critical challenges that hinder widespread implementation and effectiveness. Latency remains a stringent constraint, particularly relevant to ultra-reliable low-latency communications (URLLC) and tactile Internet applications, where inference delays and model update latencies may offset potential AI-driven optimization benefits [11, 14]. For instance, in URLLC scenarios, AI models integrated within Software-Defined Networking (SDN) controllers must execute traffic classification and anomaly detection swiftly; any lag in inference can degrade the promised low-latency service [13]. Scalability issues arise in large-scale, dynamic networks encompassing massive numbers of IoT and mobile devices; centralized AI architectures often face prohibitive computational burdens and excessive data transfer overhead. An example is AI-powered SDN frameworks in 5G that, despite improving traffic classification accuracy and throughput, struggle with computational overhead and dataset scarcity when scaling to billions of devices [13, 31].

Privacy concerns are heightened by reliance on sensitive user data and distributed learning paradigms, stimulating the adoption of privacy-preserving algorithms such as federated learning—which nonetheless introduces additional complexities in synchronization and heterogeneity management. For example, federated learning applied to AI-enhanced interference mitigation in networked sensing must balance privacy with the heterogeneity of local data distributions and communication constraints [13, 18**?**].

Interoperability remains challenging due to the diversity of vendor-specific implementations and the absence of standardized AI protocols, complicating seamless integration across multi-domain infrastructures. In practical deployments, integrating AI models trained on heterogeneous wireless communication protocols often requires ad hoc adaptation, hindering smooth multi-vendor coordination [18**?**]. Additionally, ensuring robustness against dynamic network conditions and adversarial attacks is difficult, since AI models often assume stationary environments and may degrade significantly under previously unseen scenarios or malicious perturbations. For example, AI-optimized Reconfigurable Intelligent Surfaces (RIS) and network routing algorithms demonstrate strong performance under controlled settings but remain vulnerable to adversarial perturbations and sudden environmental changes, affecting coverage and resilience [6, 24, 39].

Addressing these challenges requires developing scalable AI-SDN frameworks that optimize resource allocation and traffic management while preserving privacy [13], as well as creating robust AI algorithms capable of adapting dynamically to network changes and threats [6, 24, 39]. For instance, AI-driven routing solutions that adapt to traffic fluctuations and detect failures in real-time can enhance throughput and latency but need to balance computational complexity and privacy concerns [39]. These multifaceted challenges underscore the necessity for scalable, secure, and interpretable AI frameworks capable of reliable operation within heterogeneous, dynamic network ecosystems.

1.7 Scope and Structure of the Survey

This survey systematically examines AI applications across pivotal networking domains, from network traffic classification to autonomous fault management within software-driven infrastructures [17**?**]. It provides a focused exploration of AI methodologies tailored for software-defined networking (SDN), routing optimization, Open RAN architectures, and dynamic network slicing, reflecting the latest advances in these areas [21, 29**?**]. The assessment framework incorporates key performance metrics—throughput, latency, accuracy, scalability, and robustness—essential for evaluating real-world network performance and AI-driven adaptability [13, 16].

Integrating foundational frameworks with emerging trends, this work synthesizes classical algorithmic control methods alongside contemporary deep learning and reinforcement learning techniques to offer a comprehensive understanding of their complementary roles. It highlights recent progress in learning-based optimization, distributionally robust models, and adaptive control, discussing trade-offs, limitations, and open challenges. By combining analytical frameworks with state-of-the-art data-driven methods, the survey emphasizes the enhancement of network adaptability and operational efficiency through AI, while addressing issues such as model interpretability, computational complexity, and security vulnerabilities. This structured approach aims to provide a solid foundation to guide future research and development toward more resilient, efficient, and intelligent communication networks in complex and dynamic environments.

2 AI-Enabled Network Traffic Classification

This section provides a structured and measurable overview of AI techniques applied to network traffic classification, focusing on their characteristics, performance, and limitations. The measurable objectives are fourfold: (1) to categorize the main AI approaches used in traffic classification, including supervised, unsupervised, and hybrid models; (2) to comparatively analyze their strengths and weaknesses, supplemented by quantitative benchmarks from recent studies where available; (3) to review recent datasets and standardized evaluation protocols essential for fair performance assessment; and (4) to highlight prevailing open challenges as well as promising directions for future research.

Key AI approaches are differentiated based on their learning paradigms and feature extraction methods, enabling a deeper synthesis of how these factors impact classification accuracy, computational complexity, and adaptability to evolving traffic patterns. The discussion integrates available benchmark results to ground comparative assessments in empirical evidence. Throughout, careful attention is given to standardized citation formatting and reference verification to ensure scholarly rigor.

By segmenting content with clear objective statements and providing concise summaries of comparative insights, this section aims to facilitate clear understanding and critical appraisal of AI techniques in network traffic classification.

2.1 Taxonomy of AI Methods

AI techniques for network traffic classification broadly fall into three categories: traditional machine learning, deep learning, and

online learning. Each category differs in feature extraction methodology, adaptability, and computational requirements.

2.2 Performance Benchmarks and Datasets

Recent benchmarking efforts emphasize evaluation on publicly available datasets such as MIRAGE [**?**], ISCX VPN-nonVPN [**?**], and UCDAVIS [**?**], which provide labeled flows for both encrypted and unencrypted traffic. These datasets support standardized evaluation protocols that typically measure accuracy, F1-score, and runtime metrics, enabling consistent and fair comparison across different methods.

Empirical studies frequently report that deep learning models, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), achieve 5–15% higher accuracy than traditional methods on encrypted traffic datasets, as demonstrated in evaluations using MIRAGE. Additionally, online learning methods have shown competitive performance in streaming scenarios, but their effectiveness depends heavily on well-designed drift detection and adaptation mechanisms to maintain accuracy over time. Despite progress, challenges remain due to data scarcity and heterogeneity, which limit benchmarking comprehensiveness and real-world generalization. Consequently, there is a continued need for more extensive, updated, and diverse datasets to better capture evolving traffic patterns and encryption techniques.

2.3 Discussion and Research Directions

AI-based traffic classification faces ongoing challenges, including accurate classification amidst encrypted and obfuscated traffic, limited labeled data, and maintaining model robustness in evolving network environments. Hybrid approaches that integrate deep feature extraction with online incremental learning show promise for balancing accuracy and adaptability in dynamic scenarios.

Moreover, the field would benefit from meta-analyses comparing diverse AI methods across standardized datasets with clearly defined evaluation protocols. Establishing extended benchmarks that incorporate real-world traffic dynamics and adversarial conditions can facilitate more rigorous assessments. The availability of publicly accessible, large-scale, and up-to-date datasets remains critical to driving both foundational research and practical deployments.

In summary, the AI-enabled network traffic classification landscape demands integrated solutions that effectively address challenges related to data scarcity, model interpretability, computational resource constraints, and the continuously evolving nature of network traffic patterns. Future research directions should emphasize adaptive learning mechanisms, explainable AI models, and collaborative efforts toward dataset standardization and evaluation consistency.

2.4 Limitations of Traditional Traffic Classification Methods

Traditional network traffic classification methods, including port-based identification and deep packet inspection (DPI), exhibit significant limitations in contemporary network environments. Port-based approaches rely heavily on static assumptions about port assignments, which are increasingly invalid due to the widespread adoption of dynamic port allocations, tunneling protocols, and applications obfuscating their use of ports. DPI offers finer granularity by examining packet payloads; however, its effectiveness is greatly diminished in the presence of encrypted traffic, since payload contents become inaccessible. Beyond ineffectiveness with encryption, DPI also raises privacy concerns and incurs substantial computational overhead, which can be prohibitive in high-throughput or resource-constrained systems. Additionally, traditional methods struggle with evolving traffic patterns and concept drift, limiting their adaptability and accuracy over time.

These constraints collectively reduce the practicality and scalability of conventional techniques for managing encrypted, evolving, and complex traffic patterns encountered in modern networks. This has motivated the shift toward adaptive, data-driven classification techniques that leverage flow-level and statistical features, enabling more robust and flexible handling of encrypted and dynamic traffic [16]. Such methods typically extract flow-based and statistical features instead of relying on payload data, overcoming encryption challenges while supporting scalable implementation. Moreover, emerging hybrid or semi-supervised methods have been proposed to alleviate limitations inherent to purely supervised approaches. These methods combine limited labeled data with abundant unlabeled data to enhance classification performance and adaptability, especially under evolving network conditions and scarcity of annotated traffic samples. This direction addresses common challenges such as data imbalance and concept drift by improving generalization and reducing reliance on costly manual labeling [16]. These promising trends highlight the necessity of integrating machine learning techniques to enhance traffic classification in AI-driven networking environments.

2.5 Machine Learning Approaches for Traffic Classification

This section provides a focused overview of machine learning methodologies applied to network traffic classification, emphasizing their objectives, comparative advantages, and how they address key challenges such as encrypted payloads, dynamic traffic patterns, and the need for real-time deployment.

Advancements in artificial intelligence and machine learning (ML) offer powerful alternatives to traditional methods by exploiting statistical and behavioral traffic characteristics that remain accessible despite payload encryption. Supervised learning algorithms—including decision trees, random forests, support vector machines (SVM), k-nearest neighbors (k-NN), and neural networks—have been widely used to classify traffic flows based on features extracted from packet sizes, inter-arrival times, and flow durations [16]. These approaches depend on labeled datasets to establish classification boundaries and have shown high accuracy under controlled conditions. Ensemble models, such as Random Forest and Gradient Boosting, demonstrate particularly strong performance across diverse and evolving traffic datasets by effectively balancing accuracy and robustness [16].

Unsupervised learning techniques, especially clustering algorithms, complement supervised models by identifying anomalous or previously unseen traffic patterns without requiring labeled data.

Table 4: Summary of AI Methods for Network Traffic Classification

Method	Characteristics	Strengths	Limitations	Typical Performance
Traditional ML (SVM, Random Forest, KNN)	Feature-based; requires manual feature engineering	Well-understood; efficient on small-to-medium datasets	Limited by feature quality; less effective on raw data and encrypted traffic	Accuracy ranges from 75% to 90% depending on feature set and dataset
Deep Learning (CNN, RNN)	Automatically extracts features from raw data; capable of learning complex patterns	High accuracy; scalable; effective with encrypted traffic	Requires large labeled datasets and high computational resources; interpretability challenges	Accuracy often exceeds 90%, including on encrypted traffic classification
Online Learning	Incremental model updates with streaming data	Adaptable to evolving traffic; suitable for real-time scenarios	Model stability concerns; susceptible to noisy data and concept drift	Accuracy varies, typically 85–90% in dynamic conditions with robust adaptation

This ability is crucial for adapting to new network behaviors and detecting emerging threats, thus addressing the challenge of concept drift and evolving traffic characteristics. As a result, unsupervised methods enhance model adaptability and enable dynamic detection in real-time environments [16].

Beyond conventional ML, deep learning methodologies employ architectures such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to automatically learn hierarchical feature representations and capture temporal dependencies inherent in sequential packet flows. These models perform particularly well when encryption obscures payload content, relying instead on flow-level statistical patterns to sustain classification effectiveness [16, 36**?**]. Although deep learning achieves superior accuracy on complex and encrypted traffic, it incurs higher computational complexity and training demands, challenging large-scale experimentation and real-time deployment. This trade-off motivates ongoing research toward scalable, interpretable, and resource-efficient AI frameworks [16].

Comparative analyses on recent datasets indicate that ensemble supervised models offer a favorable balance between computational cost and classification performance. Deep learning models provide enhanced robustness against encryption and traffic variability but require greater computational resources. Unsupervised methods improve adaptability and anomaly detection capabilities, which are critical for managing non-stationary and previously unknown traffic behaviors in dynamic network environments.

Emerging research directions focus on semi-supervised, federated, and edge learning paradigms to enhance scalability, privacy preservation, and deployment feasibility in AI-driven network traffic classification. These approaches collectively address challenges such as data imbalance, limited payload visibility from encryption, concept drift, and the stringent requirements of real-time inference. Ultimately, they aim to enable more autonomous, efficient, and privacy-aware network management solutions that can dynamically adapt to evolving traffic scenarios while maintaining high classification accuracy and operational scalability [16].

2.6 Data Pipeline Processes

The success of AI-based traffic classification frameworks fundamentally depends on constructing a robust data pipeline encompassing several crucial stages: traffic collection, preprocessing, feature extraction, model training, and performance evaluation.

Traffic collection must ensure comprehensive and representative sampling across diverse network conditions to capture real-world traffic complexity, while addressing significant challenges such as encryption, dynamic port usage, and evolving traffic behaviors.

Preprocessing tackles issues such as missing data, noise, and feature normalization, aiming to produce consistent input distributions that facilitate effective learning and reduce bias.

Feature extraction is a critical phase that directly influences classifier performance. Due to limitations on payload visibility caused by encryption, most methods rely on flow-based and statistical features extracted from both flow-level and packet-level attributes, such as packet sizes, inter-arrival times, and flow durations. These features offer meaningful insights while preserving user privacy.

Model training necessitates large-scale, balanced datasets to mitigate bias toward dominant classes, improve generalization, and address concept drift arising from continuously evolving traffic patterns. Techniques such as data augmentation—including oversampling minority classes and generating synthetic samples—and resampling are employed to enhance dataset representativeness and robustness.

Evaluation rigorously measures classifier performance using metrics including accuracy, precision, recall, and processing latency [16]. The trade-offs between classification accuracy and real-time feasibility are carefully considered, especially when deploying complex models like deep learning in operational environments. Selection of appropriate evaluation criteria depends on specific application requirements and deployment constraints.

Sustaining this lifecycle is essential to maintain classifier robustness amid evolving network conditions, domain shifts, and emerging challenges related to privacy preservation and scalability. Future work focuses on scalable, generalizable models that support real-time inference while respecting privacy and resource limitations.

2.7 Performance Trade-offs

Implementing AI-powered classifiers in operational networks entails managing intrinsic trade-offs among accuracy, computational complexity, and real-time feasibility. Ensemble methods such as random forests and gradient boosting provide robust and interpretable predictive performance with moderate computational demands, demonstrating high accuracy in typical network environments. However, these approaches often face challenges in handling encrypted traffic and adapting promptly to dynamic, evolving traffic patterns [16]. In contrast, deep learning techniques improve feature abstraction and temporal dependency modeling capabilities, enabling superior classification performance on complex or obfuscated traffic. This improved effectiveness usually incurs increased inference latency and higher computational resource consumption, which can restrict scalability and limit deployment, particularly in edge environments.

Real-time network operation requires a careful balance between detection speed and classification precision. Recent adaptive frameworks incorporating incremental and online learning seek to reduce retraining overhead while enabling rapid adaptation to concept drift and shifting traffic distributions. Although these approaches are promising for maintaining updated models in dynamic conditions, they remain active research areas with ongoing challenges regarding scalability and generalization [16].

2.8 Challenges and Emerging Directions

Despite substantial progress, AI-enabled network traffic classification continues to confront several key challenges.

Data imbalance poses a critical issue as common traffic classes disproportionately dominate datasets, biasing models and reducing sensitivity to rare or malicious traffic types. This imbalance hinders the detection of infrequent but potentially severe anomalies, often leading to skewed performance metrics and inadequate responses to security threats [16].

Encrypted traffic further complicates classification, as encryption techniques obscure payload contents and dynamic port usage restrict traditional inspection methods. Consequently, classification relies mainly on flow-based and statistical features, demanding innovative feature extraction and modeling strategies capable of accurately inferring traffic types under constrained visibility and complex encryption schemes [16].

Concept drift reflects the evolving nature of network behaviors over time, requiring adaptive learning frameworks that can update models incrementally or continuously. Without such adaptability, models may become obsolete or inaccurate as traffic patterns shift due to new applications, protocols, or attacker strategies [16].

Dataset representativeness remains challenging since publicly available benchmarks often lack diversity in terms of traffic sources, network scales, environments, and temporal coverage. This limitation restricts the generalizability and transferability of trained models across heterogeneous and real-world network scenarios [16].

To address these challenges, promising future directions emphasize scalable, privacy-aware, and interpretable learning paradigms. **Semi-supervised learning** leverages abundant unlabeled network data alongside limited labeled samples, improving model robustness, mitigating labeling costs, and alleviating data scarcity issues. This approach balances the benefits of supervised and unsupervised learning to adapt to dynamic and partially labeled environments [16, 24].

Federated learning enables decentralized, privacy-preserving model training by aggregating locally trained models rather than sharing raw traffic data. Particularly advantageous for real-time edge inference and compliance with stringent data protection regulations, this paradigm supports collaborative learning across distributed networks while safeguarding sensitive information [6, 16, 24].

Explainability has become a crucial requirement for deploying AI classifiers in operational network environments. Interpretable models and post-hoc explanation techniques provide insights into model decisions, help detect potential biases, and support auditing and regulatory compliance. Enhancing explainability fosters trust, accountability, and safer network operations [16, 24].

Moreover, integrating AI with **edge computing** infrastructures enables decentralization of inference closer to data sources. This proximity reduces latency and bandwidth consumption, improves scalability, and enhances robustness against failures and attacks. The synergy between AI and edge computing facilitates more responsive, efficient, and privacy-preserving network traffic classification systems suitable for dynamic real-world deployments [6, 16].

In summary, AI-enabled network traffic classification represents a transformative advancement over traditional methods by effectively adapting to encrypted, dynamic, and heterogeneous traffic patterns. Continued innovation focusing on computational efficiency, data imbalance, interpretability, privacy, and adaptability is essential to realize AI's full potential and seamless integration within operational network environments.

3 AI Integration in Software-Defined Networking (SDN) for 5G and Beyond

The convergence of Artificial Intelligence (AI) with Software-Defined Networking (SDN) presents transformative opportunities for advancing 5G and future network technologies. SDN's programmable architecture decouples the control plane from the data plane, enabling centralized network management and dynamic resource allocation. When integrated with AI, this programmable nature facilitates intelligent decision-making, automation, and network optimization tailored to the diverse and stringent requirements of 5G networks.

AI techniques, including machine learning and deep learning, empower SDN controllers to analyze vast amounts of real-time network data, supporting critical functions such as traffic prediction, anomaly detection, fault diagnosis, and self-healing. These capabilities drive enhancements in overall network performance by optimizing routing, load balancing, and quality of service (QoS) provisioning, while simultaneously reducing latency and energy consumption. In addition, AI-driven SDN frameworks enable proactive and adaptive resource management in the dynamic and heterogeneous environments of 5G, rapidly responding to fluctuating user mobility patterns, varying network conditions, and diverse service demands.

Beyond improving performance and efficiency, integrating AI with SDN substantially advances 5G's network slicing capabilities—a key feature that creates multiple logically isolated virtual networks customized for specific applications, industries, or services. AI facilitates dynamic configuration, orchestration, and real-time adaptation of these slices, ensuring flexible, on-demand resource allocation while maintaining strict and differentiated QoS requirements. Furthermore, this integration enhances security within the SDN architecture by enabling intelligent threat detection, comprehensive real-time anomaly analysis, and automated mitigation strategies that proactively safeguard the network against emerging cyber threats.

In summary, the integration of AI in SDN for 5G and beyond is critical for realizing intelligent, flexible, and scalable networks. These networks are capable of effectively addressing the evolving technical challenges and diverse requirements imposed by next-generation applications and services.

3.1 AI-Powered SDN Architectures

The integration of Artificial Intelligence (AI) within Software-Defined Networking (SDN) architectures has fundamentally transformed network control paradigms by enabling highly centralized, programmable, and intelligent decision-making frameworks. AI enhances the SDN controller's ability to dynamically adapt to fluctuating network conditions and heterogeneous traffic demands typical

of 5G environments, thereby facilitating scalable and automated network management [13]. This architectural synergy leverages AI's pattern recognition and predictive analytics to optimize resource allocation and policy enforcement, while abstracting underlying hardware complexities.

Specifically, AI-powered SDN frameworks incorporate advanced supervised learning classifiers, such as Random Forest and Support Vector Machines (SVM), alongside deep learning models like Long Short-Term Memory (LSTM) networks within the SDN controller. These models facilitate real-time traffic classification, anomaly detection, and dynamic resource allocation, demonstrated to achieve up to 92% accuracy in traffic classification, an anomaly detection false positive rate below 3%, an 18% reduction in end-to-end latency, and a 15% throughput improvement for enhanced Mobile Broadband (eMBB) applications [13]. These capabilities contribute significantly to meeting the stringent requirements of ultra-reliable low-latency communication (URLLC) and other 5G service categories.

Nonetheless, this integration poses challenges such as computational overhead, latency constraints, dataset scarcity, and vulnerability to adversarial AI attacks, which must be carefully managed for real-time network operations. Addressing these issues motivates ongoing research into lightweight AI models optimized for real-time responses, federated learning approaches to enhance privacy and data security, and robust AI techniques resilient to adversarial attacks, thereby extending the applicability of AI-powered SDN architectures beyond 5G networks [13]. Overall, the AI-SDN synergy substantially improves scalability, flexibility, and automation in 5G and beyond network environments.

3.2 AI Techniques in SDN

Within SDN controllers, AI techniques primarily encompass supervised machine learning classifiers such as Random Forests and Support Vector Machines (SVM). These models effectively manage traffic classification and anomaly detection tasks by providing robust and interpretable decision boundaries suited to identifying diverse traffic patterns under varying network states [13]. Deep learning architectures—particularly Long Short-Term Memory (LSTM) networks—offer distinct advantages by capturing temporal dependencies and sequence dynamics in traffic flows, which are critical for modeling network behavior amidst temporal volatility [13]. The complementary utilization of shallow classifiers alongside deep recurrent networks creates a holistic framework that adapts to both static features and dynamic temporal shifts within network traffic. Empirical studies demonstrate that these AI-powered SDN frameworks can achieve up to 92% accuracy in traffic classification, reduce end-to-end latency by 18%, and increase throughput for enhanced Mobile Broadband (eMBB) services by 15%, while maintaining a false positive rate below 3% in anomaly detection [13].

Despite these promising results, practical deployment of such complex AI models faces several challenges. Training these models demands extensive labeled datasets and substantial computational resources, which may limit scalability and real-time responsiveness. Additionally, safeguarding AI models against adversarial attacks is critical to maintaining network reliability and security. Addressing these concerns, current research efforts focus on developing lightweight AI models optimized for real-time operation, employing federated learning techniques to preserve user privacy, and enhancing AI robustness against adversarial manipulations. These strategies not only improve the scalability and flexibility of AI in SDN but also facilitate automation in dynamic and heterogeneous network environments.

In summary, integrating AI techniques into SDN controllers enables more intelligent traffic management and anomaly detection, significantly improving quality of service in 5G and beyond networks. Continued advancements in model efficiency, privacy preservation, and security resilience are essential to realize the full potential of AI-powered SDN frameworks for future wireless networks [13].

3.3 Performance Improvements

Empirical studies confirm that AI-enhanced SDN architectures yield significant improvements across key network performance metrics. Specifically, the integration of supervised learning classifiers, such as Random Forest and SVM, alongside deep learning models like LSTM networks, within the SDN controller framework has achieved traffic classification accuracies up to 92%, markedly reducing misclassification errors that degrade service quality [13]. These accuracy improvements directly enhance throughput, with experimental results indicating increases close to 15% in enhanced Mobile Broadband (eMBB) scenarios, attributable to more precise resource scheduling and dynamic traffic steering enabled by AI-driven decisions. Moreover, AI's capacity for rapid anomaly detection and real-time traffic adaptation has contributed to reductions of approximately 18% in end-to-end communication latency [13]. Equally important is the reduction in false positive rates for anomaly detection to below 3%, significantly minimizing unnecessary mitigation actions that might otherwise impair network efficiency. Collectively, these benefits emphasize AI's pivotal role in elevating Quality of Service (QoS), scalability, flexibility, and automation within inherently volatile 5G SDN environments, thereby enhancing responsiveness and adaptability to dynamic and heterogeneous traffic demands [13].

3.4 Challenges

This subsection identifies and analyzes key challenges hindering the deployment and operational scalability of AI-powered SDN frameworks, explicitly linking them to methodologies discussed earlier and illustrating each with pertinent examples and case studies.

The primary challenge is the substantial computational overhead associated with sophisticated AI models embedded in SDN controllers, which impacts the real-time processing required for ultra-low-latency scenarios such as enhanced Mobile Broadband (eMBB) and ultra-reliable low-latency communication (URLLC) in 5G and beyond networks [13]. For instance, deep learning models such as LSTM networks, when integrated for traffic classification and anomaly detection, have demonstrated improvements in accuracy and throughput but impose increased inference latency. Experimental results report up to 92% accuracy in traffic classification and an 18% reduction in end-to-end latency for such AI-SDN systems [13]. The tradeoff between model complexity and latency underscores the need for algorithmic innovations that compress models to maintain

accuracy while significantly reducing processing time. Effective mitigation strategies include model pruning and knowledge distillation to achieve lightweight architectures, enabling deployment in latency-sensitive SDN environments. Real-time optimization approaches are also crucial to adapt model inference dynamically under varying network loads.

Data scarcity remains a significant barrier due to the proprietary and sensitive nature of telecom-specific datasets, which restricts the training of supervised AI models and limits their generalizability across heterogeneous network conditions [**?**]. Large language models (LLMs) hold promise for automating telecom tasks such as network configuration and traffic classification, but their data-intensive nature demands vast, domain-specific corpora that are currently difficult to access [**?**]. Federated learning frameworks have emerged as practical solutions to this problem, enabling collaborative model training across multiple operators without direct data exchange, thus preserving privacy and expanding data utility. For example, collaborative training of LLMs in multi-operator environments using federated learning helps mitigate data scarcity while safeguarding sensitive information [**?**]. However, balancing privacy, utility, and communication overhead in such distributed approaches remains an open research problem demanding further investigation.

Security threats from adversarial AI attacks pose critical risks to AI-SDN systems, as attackers may exploit model vulnerabilities to trigger erroneous decisions or evade detection, undermining network reliability [18]. Interference mitigation frameworks in perceptive mobile networks exemplify these challenges, where AI-driven resource allocation improves sensing performance but remains vulnerable to adversarial manipulation that could degrade signal-to-interference-plus-noise ratio (SINR) and sensing accuracy [18]. A surveyed AI-empowered interference management approach demonstrates a 20% improvement in detection probability and a 30% reduction in sensing interference through coordinated beamforming and deep learning-based interference prediction [18]. Robust defense mechanisms tailored to dynamic telecom conditions, including adversarial training, anomaly-aware model updating, and robust AI model design, are essential to enhance system resilience. Practical deployment of these protections must consider latency and resource constraints inherent in SDN environments.

Interoperability issues emerge from heterogeneous vendor equipment and inconsistent technological standards, complicating the seamless integration of AI models across multi-domain SDN deployments [**?**]. The lack of unified data formats and protocols hinders consistent AI control and data exchange across diverse network segments. Table 6 summarizes key standards initiatives addressing these challenges, noting their focus areas and current limitations. For example, while ETSI ENI emphasizes AI orchestration, it lacks comprehensive AI model interoperability support; similarly, 3GPP SA2 targets 5G architectures but is evolving in defining AI use cases and data sharing practices. Other initiatives, such as TIP OpenRAN AI/ML and IETF ANIMA, focus on specific domains or early-stage autonomic networking, with varying levels of vendor neutrality and adoption maturity.

To facilitate holistic understanding, Table 5 below summarizes identified challenges alongside corresponding mitigation strategies and examples, including citations for clarity.

The top three critical research priorities identified are: (1) developing computationally efficient AI models enabling ultra-low-latency inference in real-world SDN deployments; (2) designing secure and robust AI frameworks resilient to diverse adversarial threats specific to telecom environments; and (3) advancing standardized, interoperable architectures and protocols facilitating scalable AI integration across heterogeneous multi-vendor telecom infrastructures.

Addressing these priorities requires holistic approaches combining algorithmic innovation, comprehensive security frameworks, collaborative data paradigms, and ecosystem-wide standardization efforts. Central research questions guiding future work include: How can AI models be compressed and accelerated without degrading accuracy under dynamic network loads? What tailored adversarial defense strategies effectively counter telecom-specific attack vectors while maintaining operational efficiency? How can federated learning designs optimize privacy-utility tradeoffs in multi-vendor SDNs with minimal communication overhead? How might emerging standards be shaped to ensure seamless interoperability and extensibility for AI-powered network control?

In summary, overcoming these challenges through integrated technological and procedural solutions is vital to harness the full potential of AI-enabled SDN architectures in next-generation wireless networks, ultimately advancing network intelligence, scalability, security, and automation.

3.5 Prospects Beyond 5G (6G)

Looking ahead, the evolution toward beyond 5G networks, particularly 6G, envisions the development of lightweight, privacy-preserving AI models specifically tailored for distributed SDN environments [6, 24]. Federated learning emerges as a key approach, enabling collaborative model training across decentralized network nodes without exposing sensitive data, thereby addressing privacy and security concerns inherent in centralized data aggregation [13]. Anticipated advancements in multi-modal AI architectures—including large language models and multi-sensor data fusion—are expected to significantly enhance situational awareness and optimize network performance beyond existing temporal and spatial constraints [13].

Moreover, integrating Reconfigurable Intelligent Surfaces (RIS) with AI techniques such as supervised, unsupervised, and deep reinforcement learning is poised to play a pivotal role in dynamically controlling wireless environments to improve spectral and energy efficiency in 6G networks [6]. AI-enabled RIS systems optimize critical functions like channel estimation, beamforming, and resource allocation by learning mappings from complex and dynamic channel state information to optimal RIS configurations, outperforming traditional heuristic methods [6]. This synergy promises to enhance coverage, robustness, and adaptability, which are vital for meeting 6G requirements such as ultra-reliability, massive connectivity, and real-time intelligence [6].

Nonetheless, realizing federated and privacy-aware AI solutions entails overcoming substantial computational, communication, and standardization challenges. High-dimensional RIS configuration spaces and stringent low-latency demands impose constraints on AI scalability and real-time deployment [6, 24]. To address these

Table 5: Summary of AI-SDN Challenges with Corresponding Mitigation Strategies and Examples

Challenge	Mitigation Strategies	Illustrative Examples/Case Studies
Computational Overhead	Model compression (pruning, distillation), lightweight architectures, real-time optimization	LSTM integration in SDN controllers for URLLC scenarios with latency reduction techniques [13]
Data Scarcity	Federated learning, privacy-preserving data sharing, synthetic data generation	Collaborative training of LLMs across operators using federated approaches to address telecom data scarcity [?]
Security Vulnerabilities	Adversarial training, anomaly detection, robust AI model design	AI-enabled interference mitigation frameworks enhanced with adversarial resilience to maintain sensing SINR [18]
Interoperability	Standardization of AI model interfaces and protocols, vendor-neutral APIs	ETSI ENI and 3GPP SA2 standards advancing multi-vendor AI-SDN integration though still evolving [?]

Table 6: Summary of Current Standards Initiatives Relevant to AI-SDN Integration

Standard Initiative	Scope	Limitations
ETSI ENI (Experiential Networked Intelligence)	Framework for AI-driven network management and automation	Focuses on orchestration; limited coverage of AI model interoperability
3GPP SA2 AI/ML Work Items	AI/ML integration for 5G system architecture and management	Primarily targets 5G; evolving definitions for AI use cases and data sharing
TIP OpenRAN AI/ML	AI/ML aspects for OpenRAN architectures	Concentrates on RAN domain; vendor-specific implementations limit generalizability
IETF ANIMA (Autonomic Networking)	Autonomic networking protocols supporting AI-driven control	Early stage; interoperability challenges remain across multi-vendor environments
IEEE P2894 (AI/ML Data Formatting)	Standardization of data representations for AI/ML in networks	Emerging standard; adoption in telecom industry is limited so far

challenges, interdisciplinary research bridging AI, communications, and network engineering is essential. This includes developing lightweight distributed AI algorithms and robust models resilient to adversarial conditions, as well as scalable frameworks capable of operating efficiently amid heterogeneous network traffic and dynamic environments [6, 13, 24]. The fusion of advanced AI algorithms with emerging wireless technologies will be critical to achieving intelligent, autonomous, and scalable next-generation networks.

3.6 Summary

In summary, the integration of artificial intelligence (AI) techniques into Software Defined Networking (SDN) paradigms for 5G networks has driven notable advancements in enhancing network flexibility, adaptability, and overall performance. Measurable objectives frequently targeted in recent studies include reducing end-to-end latency, improving throughput, and increasing energy efficiency. For example, empirical results highlight latency reductions by up to 30%, throughput improvements exceeding 20%, and energy efficiency gains around 15%, underscoring the practical benefits of AI-empowered SDN frameworks [36**?**].

Key AI methods employed include reinforcement learning (RL), heuristic optimization, deep learning (DL), and evolutionary algorithms—all showing diverse strengths and constraints under real-world network conditions. RL methods dynamically learn optimal routing policies through environment interactions, enabling adaptability to network changes; however, they often require considerable training time and computational resources [36]. Heuristic optimization applies domain-specific rules to rapidly converge on good, though potentially sub-optimal, solutions, trading accuracy for efficiency. DL-based approaches excel in accurate traffic pattern and fault prediction to proactively improve routing decisions but demand significant computational power and large datasets [**?**]. Evolutionary algorithms offer robust adaptability via iterative improvements, excelling in fault tolerance, yet tend to converge more slowly compared to other techniques.

Despite these progressions, practical deployment of AI-based SDN architectures faces critical challenges. High computational overhead and training latency restrict scalability, while security vulnerabilities threaten network trustworthiness. Moreover, the contextual effectiveness of AI methods under diverse, dynamic real-world scenarios remains an open concern, necessitating further empirical validation and robustness enhancement.

Table 7 presents a concise yet critical comparison of prominent AI-based routing techniques within SDN environments, highlighting their key strengths, observed performance gains, and inherent limitations. This structured overview elucidates essential trade-offs and provides a basis for informed method selection according to specific network requirements.

Looking forward, the transition toward 6G networks calls for the development of optimized, distributed AI frameworks that balance intelligent decision-making with computational efficiency and stringent privacy safeguards. Future research agendas should explicitly focus on metrics such as model training time, adaptability under changing network conditions, resilience against security threats, and overall system scalability. These goals represent the forefront of efforts to realize fully programmable, autonomous, and resilient network architectures in next-generation communication systems.

3.7 AI-Driven Routing Optimization

This section aims to provide a comprehensive overview of AI-driven routing optimization, focusing on its objectives, key challenges, methodologies, and future directions. The primary goal is to analyze how AI techniques enhance routing efficiency and adaptability in complex and dynamic network environments, contrasting them with traditional approaches.

AI-driven routing optimization leverages advanced machine learning algorithms to improve the efficiency and adaptability of routing in complex networks. These approaches employ predictive models to forecast network conditions and dynamically adjust routing paths, aiming to optimize multiple performance metrics such as latency, throughput, and energy consumption [].

We categorize the algorithmic challenges in AI-driven routing into four critical areas: (1) accurate prediction of dynamic network states, (2) efficient path computation under diverse and varying constraints, (3) real-time adaptation to network changes, and (4) scalability in large, heterogeneous network environments. Addressing these challenges is essential for developing robust and effective routing solutions.

For instance, machine learning models can predict traffic congestion or potential link failures, enabling preemptive rerouting to uphold quality of service. In software-defined networking (SDN)

Table 7: Comparison of AI-Based Routing Techniques in SDN for 5G Networks

Method	Key Strengths	Performance Gains	Limitations
Reinforcement Learning	Learns optimal policies via environment interaction; adapts dynamically to network changes	Latency reduction up to 30%	High training time and computational complexity; requires extensive data and resources
Heuristic Optimization	Applies domain-specific rules for fast convergence	Throughput improvements over 20%	May yield sub-optimal routing decisions; limited adaptability to network dynamics
Deep Learning-based	High accuracy in traffic and fault prediction enabling proactive routing	Energy efficiency gains of 15%	Significant computational resource and large dataset requirements; potential data privacy concerns
Evolutionary Algorithms	Robust adaptation to dynamic network conditions via iterative improvement	Enhanced fault tolerance	Slower convergence compared to other methods; may incur higher runtime cost

contexts, reinforcement learning algorithms iteratively refine routing policies based on continuous network feedback, which has demonstrated reductions in latency and improvements in throughput. Compared to heuristic or rule-based methods, these AI-driven approaches offer enhanced flexibility and the ability to learn optimal policies over time.

Despite these advances, several open challenges persist. A critical trade-off exists between the computational overhead imposed by complex AI models and the necessity for fast, real-time routing decisions. Moreover, the transferability of trained models across different network topologies remains limited, often requiring extensive retraining. Various approaches, such as the development of lightweight models and domain adaptation techniques, are being explored to overcome these limitations.

In summary, AI-driven routing optimization provides significant improvements over traditional static routing methods by enabling networks to predict and react more effectively to dynamic conditions. Nevertheless, ongoing research must focus on balancing model complexity, execution speed, and adaptability across diverse network scenarios to fully realize the potential of AI in routing.

3.7.1 Balancing Exploration and Exploitation. One critical challenge in routing algorithms is effectively balancing the trade-off between exploration and exploitation. Exploration involves probing new paths to discover potentially better or more optimal routes, while exploitation focuses on leveraging known, reliable paths to maintain consistent and stable network performance. Achieving this balance becomes especially challenging in highly dynamic environments, where network conditions and topologies change rapidly, requiring routing algorithms to adapt swiftly and make real-time decisions []. Properly addressing this balance is essential to optimize routing efficiency, minimize latency, and ensure robustness in dynamic networks.

3.7.2 Integration of Heterogeneous and Real-Time Data. Integrating heterogeneous data sources and real-time measurements remains a significant challenge. AI models must efficiently process diverse and sometimes incomplete information—including traffic patterns, link quality, and node status—to generate accurate and timely routing predictions. Addressing this challenge requires methods that ensure scalability and maintain low computational overhead, as routing optimization must perform robustly in large-scale, dynamic networks []. Additionally, effective fusion techniques that handle data variability and latency are crucial to enable real-time decision making in complex network environments.

3.7.3 Handling Uncertainty and Variability. Robustness against uncertainty and variability in network conditions is paramount. Reinforcement learning techniques are commonly employed to continuously refine routing policies based on real-time feedback from the network environment, thereby enhancing resilience to dynamic network disruptions and failures []. These methods adaptively respond to fluctuating traffic loads, link failures, and changing topologies, enabling more reliable and efficient routing decisions even under unpredictable conditions.

3.7.4 Meeting Latency and Reliability Requirements. Designing algorithms that meet stringent latency and reliability constraints remains a significant research challenge in AI-driven routing optimization. Ensuring timely and dependable network performance requires lightweight, adaptive models capable of real-time decision making under dynamic and often unpredictable conditions. Key considerations include minimizing computational overhead to reduce latency, incorporating mechanisms to handle packet loss and failures to improve reliability, and maintaining robustness against network variability. Future research directions involve enhancing model interpretability to facilitate debugging and foster trust, improving responsiveness for seamless real-time adaptation, and developing rigorous formal verification and validation frameworks that provide provable guarantees on latency and reliability performance bounds [].

In summary, addressing these challenges requires adaptable, scalable, and robust solutions tailored to the complexities of modern network environments. This section has synthesized the primary algorithmic hurdles and outlined promising avenues for future work to unlock the full potential of AI-driven routing optimization.

3.7.5 Limitations of Static Routing Protocols. Traditional static routing protocols lack the adaptability necessary for dynamic and heterogeneous network environments, leading to suboptimal performance under varying traffic patterns and network conditions. Originally designed for relatively stable and homogeneous infrastructures, these protocols exhibit limited real-time responsiveness to fluctuating workloads, mobility-induced topology changes, and unpredictable link failures. Consequently, challenges such as increased latency, reduced throughput, and vulnerability to faults frequently arise in large-scale, multi-tenant networks typical of modern wireless and software-defined architectures [39]. Moreover, the rigidity inherent in static routing constrains efficient resource utilization and impedes the exploitation of cross-layer contextual information, which is critical for advancing 5G and beyond networks.

These limitations underscore the pressing need for adaptive routing mechanisms that can dynamically respond to changes in network state by learning from traffic patterns and anomalies. Emerging AI-driven approaches leverage machine learning techniques, including reinforcement learning and neural networks, to optimize routing decisions in real-time. These methods address routing as a multi-objective optimization problem that balances throughput, latency, and fault tolerance [39]. Empirical evidence shows that such AI-based routing solutions can achieve up to 30% improvements in

Table 8: Summary of AI-Driven Routing Optimization Challenges and Solutions

Challenge	Description	Common Approaches
Exploration vs. Exploitation	Balancing discovery of new optimal routes with stability of known paths	Reinforcement learning, multi-armed bandits
Integration of Heterogeneous Data	Processing diverse and real-time network information	Data fusion, online learning
Uncertainty and Variability	Coping with dynamic and unpredictable network conditions	Robust learning, continual adaptation
Latency and Reliability Requirements	Ensuring performance under strict timing and availability constraints	Lightweight models, real-time optimization, formal verification

throughput and latency while enhancing resilience through rapid failure detection and rerouting.

However, these advancements come with challenges such as increased computational overhead, scalability concerns, the necessity for representative training data, and the complexity of integrating AI-driven protocols with existing network infrastructure. Future research directions emphasize decentralized and federated learning methods to enable scalable, privacy-aware routing, as well as the development of hybrid AI-conventional schemes that combine adaptive intelligence with the stability of established routing protocols [39]. Overall, transitioning from static to intelligent, AI-based routing frameworks holds significant promise for the robust and efficient operation of next-generation networks.

3.7.6 Reinforcement Learning and Neural Networks for Routing. Artificial intelligence (AI) approaches, particularly reinforcement learning (RL) and neural networks (NNs), provide advanced tools to surpass the constraints of static routing by enabling adaptive, data-driven path optimization. RL algorithms iteratively explore and exploit routing policies to dynamically optimize multiple objectives such as throughput maximization, latency minimization, and fault tolerance enhancement [39**?**]. This results in dynamic path prediction that directly responds to real-time network states and traffic dynamics. Neural networks complement this by learning complex nonlinear mappings from network metrics to optimal routing decisions, thereby generalizing from historical data and adapting to new or unforeseen network conditions [27]—[**?**]. Their ability to handle nonlinear relationships and feature importance leads to improved routing decisions tailored to heterogeneous network environments. Together, these AI-driven methods empower autonomous routing frameworks that accommodate heterogeneous node capabilities and varying traffic demands, frequently outperforming traditional heuristics through superior robustness and scalability.

Nonetheless, key challenges remain. RL requires careful algorithm design to balance exploration and exploitation effectively to avoid convergence to suboptimal routing policies. Moreover, maintaining model generalization without overfitting to specific network scenarios necessitates ongoing retraining and adaptation to evolving traffic patterns and network topologies [39]. Computational overhead and scalability also pose significant challenges, especially for deployment in large-scale or resource-constrained networks. Empirical evidence indicates that AI-empowered routing schemes can improve network throughput and reduce latency by up to 30% compared to static routing protocols, while also enhancing fault tolerance via rapid anomaly detection and rerouting [39**?**]. Continuous research is essential to further enhance computational efficiency, scalability, and seamless integration with legacy network infrastructure, enabling AI-driven routing to achieve its full potential in real-world, heterogeneous network deployments.

3.7.7 Traffic Prediction and Anomaly Detection Integration. This subsection elucidates the integration of traffic prediction and anomaly detection within AI-driven routing to enhance network performance and resilience. It critically examines methodologies, performance trade-offs, and practical deployment challenges in this emerging paradigm.

The integration of traffic prediction and anomaly detection into routing optimization marks a pivotal advancement, enabling proactive network management that anticipates and mitigates performance degradation instead of solely responding after issues arise. Traffic prediction models typically leverage supervised learning combined with advanced time-series deep learning architectures such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) to forecast network load and congestion patterns. These models capture complex temporal dependencies to achieve high accuracy but introduce trade-offs between prediction latency and computational complexity, which impact real-time routing decisions and scalability [24]. For instance, while RNNs provide detailed temporal modeling, their inference latency may be unsuitable for ultra-low-latency environments, thereby necessitating simpler models or hardware acceleration.

Simultaneously, anomaly detection systems continuously monitor network operations to identify deviations arising from link failures, cyber-attacks, or misconfigurations. They must balance sensitivity and specificity to reduce false positives, which, when excessive, can trigger unnecessary route recalculations causing network instability and degraded performance [39]. Techniques range from statistical thresholding to deep autoencoders, each with trade-offs in detection speed and accuracy.

The combination of predictive traffic modeling with anomaly detection within AI-driven routing frameworks facilitates multi-objective optimization addressing performance, reliability, and security simultaneously. This comprehensive approach enhances overall network resilience by preempting bottlenecks and limiting fault propagation, surpassing the capabilities of traditional static routing [39]. Empirical studies demonstrate up to a 30% improvement in throughput and latency along with faster recovery from network faults, highlighting the practical benefits of this integration.

Key challenges remain, including maintaining model accuracy under highly dynamic and non-stationary network conditions characterized by heterogeneous and large-scale data sources. Deployment obstacles include the computational overhead of complex models, the necessity of representative and diverse training datasets to prevent overfitting, and integration with existing network infrastructures that may lack flexibility [24, 39]. Addressing these requires robust, generalizable models combined with adaptive training strategies.

Recent advancements promote incorporating reinforcement learning techniques, enabling routing policies to dynamically adapt based on evolving traffic states and detected anomalies. These approaches improve scalability and real-time responsiveness, crucial for modern network contexts such as 5G and software-defined networking (SDN) [39]. However, reinforcement learning introduces additional complexity in training procedures and convergence guarantees, which present deployment hurdles.

In summary, the convergence of traffic prediction and anomaly detection within AI-driven routing paradigms presents a transformative framework underpinning self-optimization and enhanced robustness essential for complex and emerging network architectures. Balancing model performance, computational demands, and integration challenges is critical for achieving practical, widespread adoption.

3.7.8 Empirical Gains and Challenges. Experimental validations of AI-driven routing protocols consistently demonstrate substantial gains, including increased throughput, reduced latency, and improved fault tolerance across diverse network topologies and traffic scenarios [39**?**]. For instance, reinforcement learning and neural networks enable dynamic adaptation to changing network conditions, yielding up to 30% improvements in throughput and latency alongside enhanced resilience through rapid failure detection and rerouting [39]. Specifically, [39] formulates routing as a multi-objective optimization integrating traffic prediction and anomaly detection, producing these empirical gains.

However, these benefits entail significant computational and communication overheads, particularly in centralized learning architectures, which can become bottlenecks impacting real-time operation and scalability [39]. To address communication overhead, [**?**] proposes a robust federated learning framework that combines gradient sparsification with error feedback mechanisms and adaptive client selection. This approach reduces communication costs from 120 MB to 55 MB and accelerates convergence from 200 to 150 rounds while maintaining test accuracy above 85% even under adverse conditions such as 40% client dropout. The strategy carefully balances communication efficiency against potential accuracy loss, highlighting a nuanced trade-off between compression levels and model performance.

These mechanisms offer a practical route to overcoming scalability challenges by exploiting client heterogeneity and accommodating unreliable wireless links typical in real-world deployments. Furthermore, adaptive client selection strategies selectively involve devices with sufficient resources and stable connectivity, maximizing effective participation without overwhelming network bandwidth. Such detailed comparative analyses underline how communication-efficient federated learning can sustain both accuracy and robustness in resource-constrained environments.

Security vulnerabilities also arise from adversarial attacks targeting either training data or inference stages, which may degrade routing performance and threaten network stability [39**?**]. Effective mitigation demands robust, network-specific security protocols tailored to AI-integrated routing frameworks. Additionally, integrating AI models within legacy network infrastructures introduces further complexity. Hybrid or modular deployment strategies become essential to preserve backward compatibility and avoid service disruptions, thereby allowing continuous network operation while incrementally adopting AI [18].

The demand for real-time inference combined with ongoing model updates intensifies these challenges, necessitating trade-offs among inference accuracy, responsiveness, and resource consumption. Lightweight AI model designs coupled with distributed and federated learning frameworks augmented by security-aware mechanisms are fundamental to navigating these trade-offs [39**?**]. Continuous innovation in adaptive compression techniques, coded computing, and privacy-preserving methods remains critical to fully harness the transformative potential of AI-driven routing within highly dynamic and heterogeneous next-generation networks.

3.7.9 Future Trends. Emerging directions in AI-based routing optimization increasingly emphasize decentralized and federated learning architectures to address the scalability and privacy challenges inherent in centralized methods. Federated learning enables multiple distributed clients or network nodes to collaboratively train a shared model without exchanging sensitive local data, thereby enhancing privacy and reducing communication overhead [39]. Advanced federated frameworks incorporate adaptive client selection and gradient sparsification techniques, efficiently handling heterogeneous device capabilities and client dropout, which improves convergence speed and accuracy in real-world wireless environments [6].

Hybrid routing algorithms that combine AI techniques with conventional protocols hold significant promise for delivering adaptive yet resource-efficient solutions. These hybrid schemes leverage the heuristic strengths and stability of traditional protocols while integrating machine learning components to enhance adaptability and predictive accuracy.

To provide a clearer roadmap, near-term research (1–3 years) will focus on developing scalable federated learning models with robust privacy-preserving mechanisms and on investigating their integration into existing SDN-based network infrastructures [39]. Within 3–5 years, the focus is likely to shift towards optimizing hybrid AI-conventional routing algorithms that balance adaptability with computational efficiency, incorporating explainability and transparency methods to foster operator trust. Longer-term challenges (5+ years) include extending AI-based routing to emerging paradigms such as 6G networks, massive MIMO, and edge computing ecosystems [6, 39], where the dynamic, high-dimensional nature of network environments demands fully autonomous, resilient, and scalable routing strategies.

Recent pioneering studies emphasize the integration of Reconfigurable Intelligent Surfaces (RIS) with AI as a promising path to dynamically optimize radio environments. These integrated approaches can be incorporated into routing decisions to enhance spectral and energy efficiency [6]. Parallel efforts demonstrate AI-driven routing frameworks utilizing reinforcement learning to achieve improvements of up to 30% in throughput and latency by adapting effectively to real-time network states and failures [39].

Together, these advancements are critical to realizing fully autonomous, scalable, and resilient routing architectures capable of dynamically adapting to complex and evolving network conditions,

addressing key challenges such as computational overhead, scalability, and trustworthiness.

4 AI in Open Radio Access Network (Open RAN) for 6G

Open Radio Access Network (Open RAN) represents a significant paradigm shift in mobile network architecture by promoting openness, flexibility, and intelligence, which are fundamental for the evolution toward 6G. The integration of Artificial Intelligence (AI) into Open RAN enables more efficient and adaptive network management, dynamic resource allocation, and optimized service delivery, all critical to meeting the stringent requirements of 6G such as ultra-low latency, massive connectivity, energy efficiency, and enhanced reliability.

AI techniques in Open RAN empower intelligent radio resource management by dynamically optimizing spectrum usage, power control, and interference mitigation across disaggregated and virtualized network components. Advanced machine learning models can predict traffic demand, user mobility, and network anomalies, allowing proactive and real-time adaptation of network functions to consistently satisfy 6G performance targets. Furthermore, AI-driven automation facilitates robust self-organizing and self-healing capabilities within the RAN, significantly reducing operational expenditures while enhancing user Quality of Experience (QoE).

Beyond improving operational efficiency, AI integration strengthens the security posture of Open RAN by continuously monitoring network behavior, detecting cyber threats, and enabling rapid response mechanisms. This security aspect is especially critical as 6G networks are expected to support safety-critical applications including autonomous vehicles, remote healthcare, and industrial automation. The modular and open interface architecture of Open RAN further supports continuous deployment and online training of AI models, accelerating innovation cycles and enabling customized solutions tailored to diverse deployment contexts within 6G ecosystems.

In summary, embedding AI within Open RAN is a cornerstone for unlocking the full potential of 6G networks. It provides adaptive, intelligent orchestration and control that align with the complex, dynamic, and heterogeneous demands of future wireless communication systems, paving the way toward sustainable, resilient, and high-performing 6G infrastructures.

4.1 Open RAN Architecture and AI Integration Layers

Open RAN introduces a transformative approach to wireless network infrastructure by disaggregating traditionally monolithic radio access components into three distinct units: the Radio Unit (RU), Distributed Unit (DU), and Centralized Unit (CU). This modular design fosters openness and programmability through well-defined interfaces, enabling accelerated innovation and increased vendor diversity. A pivotal advancement in Open RAN is the multilayer integration of artificial intelligence (AI), which enhances network intelligence by embedding AI capabilities at each architectural layer to systematically tackle unique operational challenges and holistically optimize performance [25].

At the RU level, AI models operate under stringent latency constraints to enable real-time radio signal processing and physical layer optimizations, such as adaptive beamforming and dynamic spectrum management. The DU leverages AI to manage scheduling, resource allocation, and localized interference mitigation, utilizing moderate computational resources alongside locally gathered datasets. Meanwhile, the CU aggregates network-wide telemetry data to execute sophisticated AI analytics that enable dynamic orchestration, fault detection, and long-term network optimization.

This hierarchical AI deployment framework strategically balances computational complexity and latency requirements, ensuring scalability and responsiveness. It supports advanced techniques such as federated learning, which preserves user privacy by enabling distributed intelligence without sharing raw data, and reinforcement learning, which adapts scheduling policies dynamically in response to changing network conditions. Leveraging standardized telemetry data, AI agents perform real-time analytics and closed-loop control that significantly improve throughput, latency, energy efficiency, connection reliability, handover success, and resource utilization, outperforming traditional heuristic approaches [25].

Despite these advances, challenges remain in computational overhead, AI model convergence, real-time inference latency, multi-vendor interoperability, energy and computation constraints at the edge, and regulatory/privacy concerns. Addressing these requires developing lightweight AI models, hardware accelerators for edge deployment, and standardized AI-specific protocols within Open RAN frameworks. Future work also envisions enhanced explainable AI for transparency, multi-agent collaborative learning, and integrating emerging technologies like quantum computing and blockchain to bolster security and performance.

Overall, multilayer AI integration in Open RAN marks a paradigm shift towards automated, optimized, and innovative next-generation networks, laying a robust foundation for resilient and user-centric 6G systems while proactively addressing ecosystem complexities and implementation challenges [25].

4.2 AI Techniques in Open RAN

A diverse array of AI techniques underpins Open RAN functionalities, each selected to meet specific operational objectives. Federated learning plays a pivotal role by enabling distributed model training across the RU, DU, and CU layers without exchanging raw data, thus preserving user privacy and effectively managing the multi-vendor and multi-domain heterogeneity characteristic of Open RAN ecosystems. This decentralization fosters collaborative intelligence while securing sensitive information [25]. Reinforcement learning (RL), particularly deep RL, facilitates autonomous decision-making for dynamic spectrum management, adaptive resource allocation, and interference mitigation by learning optimal policies through interactions with complex, time-varying network environments. These AI-driven methods empower the network to continuously adapt and optimize performance in real time [1, 37]. Deep neural networks (DNNs) are extensively employed for tasks demanding sophisticated pattern recognition and nonlinear function approximation, such as fault detection and anomaly identification, utilizing

large volumes of standardized telemetry data harvested within Open RAN infrastructures [22, 30].

Furthermore, hybrid fusion techniques showcase AI's flexibility in handling heterogeneous communication modalities and interference mitigation. For example, combining visible light communication (VLC) and radio frequency (RF) channels leverages machine learning-based fusion and interference cancellation algorithms to bolster robustness in challenging propagation conditions [26]. This VLC-RF hybrid strategy exploits the complementary physical properties of both communication methods, enhancing overall communication reliability and throughput. Additionally, AI-informed sequence management approaches optimize the design of high-capacity preamble sequences for random access channels. These advancements markedly reduce collision probabilities by increasing the number of unique preambles exhibiting low cross-correlation and strong auto-correlation properties, thereby improving detection accuracy and decreasing retransmission overhead. Such capabilities are vital for supporting the massive IoT device connectivity demands intrinsic to 5G and beyond [37].

Despite these technological advances, practical deployment of AI within Open RAN confronts several significant challenges. These include the computational burdens imposed by real-time model training and inference, complexities in coordinating AI functionalities across heterogeneous multi-vendor equipment, and ensuring scalability and seamless interoperability within fluid, dynamic network conditions. Addressing these obstacles is essential to fully realize the potential of AI-powered Open RAN networks [6, 25].

4.3 Performance Enhancements

Integrating AI into Open RAN architectures markedly improves key network performance metrics, including throughput, latency, energy efficiency, reliability, and resource utilization. AI-driven algorithms enable dynamic spectrum access and intelligent scheduling that adapt bandwidth allocation responsively to fluctuating traffic and complex interference conditions, thus boosting effective throughput and minimizing latency [25]. Energy efficiency is enhanced through AI-enabled resource optimization and hardware-aware scheduling schemes, which selectively power down idle components or scale computing tasks based on real-time network load, contributing to greener operations [6].

Reliability and connection robustness benefit from AI-based proactive fault detection and predictive maintenance, which facilitate early identification of anomalies before service degradation occurs. For example, reinforcement learning algorithms dynamically adjust handover parameters, reducing connection drops and improving mobility management [25]. Furthermore, AI improves resource utilization by analyzing extensive telemetry data to identify bottlenecks and redundant allocations, thereby facilitating efficient network slicing and large-scale multi-access edge computing (MEC) deployments [6].

Notably, the integration of Reconfigurable Intelligent Surfaces (RIS) combined with AI techniques leads to intelligent wireless environments that dynamically optimize the propagation environment, further enhancing coverage, spectral efficiency, and robustness [6]. AI methods, such as supervised, unsupervised, and deep reinforcement learning, play a pivotal role in optimizing RIS functions including channel estimation, beamforming, and resource allocation by learning complex mappings from channel states to optimal RIS configurations. Experimental results indicate that AI-enabled RIS outperforms traditional heuristic approaches by efficiently adapting to dynamic network conditions and imperfect channel information, thereby boosting coverage and system robustness [6]. These AI-driven enhancements collectively enable Open RAN to surpass traditional heuristic methods, adapting efficiently to dynamic network states and complex scenarios while managing challenges such as latency constraints, scalability, and interoperability [25].

However, despite these promising improvements, AI integration in Open RAN is not without its limitations and challenges. Some deployments have experienced issues related to the high computational overhead and latency introduced by complex AI models, which can impact real-time decision-making and network responsiveness [25]. Additionally, model convergence difficulties and data quality variability sometimes lead to suboptimal or inconsistent performance gains. In environments with multi-vendor components, interoperability challenges further complicate seamless AI orchestration across the Open RAN ecosystem [25]. Security vulnerabilities arising from AI model attacks or misconfigurations remain an important concern that can affect network reliability and integrity. Moreover, while AI-enabled RIS shows significant theoretical gains, practical implementations must address the high-dimensional configuration spaces and stringent latency requirements, which may limit performance improvements or increase deployment complexity [6].

Acknowledging these trade-offs, ongoing research emphasizes lightweight AI models tailored for edge deployment, federated learning to enhance privacy and scalability, and robust fault tolerance mechanisms [25]. This balanced approach aims to maximize performance enhancements while mitigating the risks and challenges identified in less successful deployments. Thus, AI-empowered Open RAN holds great promise as a cornerstone for resilient, efficient, and sustainable 6G networks, contingent upon continued advances in model efficiency, standardization, and security practices.

4.4 Challenges

The integration of AI into Open RAN architectures introduces several critical challenges that must be systematically addressed to realize its full potential. This subsection aims to clearly outline these challenges, supported by concrete examples and critical evaluation of existing solutions, while linking them to previously discussed methodologies. Our objectives are to provide a comprehensive understanding of the technical and operational barriers hindering AI-driven Open RAN deployments, highlight measurable goals to assess progress, and identify key research priorities.

One major challenge is **scalability**. AI models designed for Open RAN must efficiently handle massive data streams generated by diverse network elements in real time. For example, applying deep reinforcement learning algorithms for dynamic resource allocation [36] faces difficulties in scaling due to computational complexity and latency constraints. Existing solutions often tradeoff

accuracy for speed, yet the balance remains an open problem. Ongoing work focuses on lightweight model compression and distributed learning frameworks to alleviate these issues.

Privacy is another critical concern, as AI systems process sensitive user and network data. Federated learning has been proposed to train models without centralizing data [**?**], but challenges persist in mitigating poisoning attacks and ensuring data heterogeneity does not degrade performance. Practical deployments must integrate robust cryptographic mechanisms alongside algorithmic approaches, which remain areas for expanded research and validation.

Interpretability of AI decisions in Open RAN is essential for trust and regulatory compliance. While explainable AI techniques such as attention mechanisms and model-agnostic methods [**?**] have been explored, their integration into complex radio management workflows is still immature. More concrete algorithmic descriptions and benchmarks are needed to evaluate how interpretability impacts decision quality and operator intervention.

Security poses formidable threats both in AI model integrity and underlying Open RAN infrastructure. Case studies of adversarial attacks on AI-driven baseband functions [**?**] demonstrate the need for comprehensive defense strategies, combining anomaly detection, robust training, and secure model updates. Practical solutions must be assessed under realistic threat models and performance metrics.

Table 9 summarizes the main challenges alongside example solution directions and measurable evaluation criteria.

In summary, the multifaceted challenges of integrating AI in Open RAN require interdisciplinary efforts addressing algorithmic, security, and operational dimensions. Future research must prioritize scalable and privacy-preserving algorithms that provide interpretable outputs without compromising security. Bridging the gap between theoretical advancements and real-world deployments is critical to unlocking the full potential of AI-assisted Open RAN. This section sets the stage for subsequent discussions on specific solution frameworks and deployment strategies.

4.4.1 Model Convergence in Dynamic Environments. A primary challenge in AI-enabled wireless communications is ensuring reliable model convergence within the highly dynamic and non-stationary nature of wireless channels, which are often subject to partial observability. Reinforcement learning (RL) approaches, while promising for adaptive resource management, face instability risks due to rapidly fluctuating radio conditions and shifting channel statistics, as highlighted in [25]. For example, an RL-based scheduler may struggle to stabilize when environment dynamics change suddenly, resulting in degraded network throughput, increased latency, and unreliable quality of service. This underscores critical research questions: how can reinforcement learning algorithms be designed with provable robust convergence properties under transient and non-stationary states? Additionally, how can real-time adaptation mechanisms be integrated to guarantee consistent and reliable decision-making despite environmental uncertainties? Addressing these challenges requires developing advanced RL frameworks capable of dynamic exploration-exploitation trade-offs, continual learning, and incorporating domain knowledge or telemetry data for informed policy updates, as envisioned in AI-driven Open RAN architectures [25].

4.4.2 Computational and Energy Constraints at the Edge. Deploying complex AI models at edge units such as RUs and DUs presents significant limitations due to processing capacity and energy constraints [18, 25]. For instance, interference mitigation leveraging deep learning requires low-latency inference on resource-constrained hardware, creating a critical trade-off between model complexity and operational feasibility. The framework in [18] demonstrates that AI-driven interference prediction can reduce sensing interference by 30%, but only when inference latency remains within strict bounds to ensure real-time responsiveness. Therefore, developing lightweight AI architectures alongside hardware-software co-design strategies is essential. Future research must carefully balance inference latency, energy consumption, and accuracy by exploring novel accelerators and compact models that optimize this trade-off without compromising performance. Additionally, techniques such as federated learning and hardware-efficient neural networks are promising to address distributed intelligence and energy efficiency challenges in edge deployments [25].

4.4.3 Multi-Vendor Interoperability. The heterogeneity of Open RAN components from multiple vendors complicates interoperable AI deployment [6, 25]. For example, integrating intelligent Reconfigurable Intelligent Surfaces (RIS) optimized by AI requires aligning diverse hardware interfaces and harmonizing data formats, as explored in [6]. The lack of unified standards impedes seamless AI model sharing and operation across vendors, thereby limiting scalability, maintainability, and real-time responsiveness. Addressing these challenges involves the development of middleware frameworks and standardization efforts that align data representations and AI interfaces while ensuring robust security and operational efficiency [25]. Such frameworks facilitate consistent telemetry data exchange, enable AI-driven closed-loop control, and manage the complexity of multi-vendor environments. Furthermore, federated learning and distributed AI paradigms are promising approaches for preserving data privacy and fostering collaborative intelligence across diverse network elements without direct data sharing, thus tackling interoperability and security concerns simultaneously [25].

4.4.4 Security and Privacy Risks. The growing complexity of AI systems within Open RAN architectures significantly increases vulnerabilities to adversarial attacks and privacy breaches [24, 25]. AI models, especially those trained using federated learning, can inadvertently leak sensitive information, while adversaries may manipulate AI-driven network management by exploiting model weaknesses. For instance, [24] highlights emergent threats such as adversarial intrusions that compromise service quality and confidentiality. To counter these risks, it is essential to develop real-time intrusion detection mechanisms alongside robust AI architectures designed to withstand adversarial conditions. Moreover, privacy-preserving learning protocols must be integrated to protect user data without sacrificing distributed intelligence. Implementing these solutions requires careful consideration of regulatory compliance across multiple jurisdictions, balancing security, privacy, and performance in increasingly autonomous and complex networks.

4.4.5 Regulatory and Governance Challenges. The deployment of AI-driven Open RAN solutions must rigorously adhere to regulatory mandates concerning data sovereignty, user privacy, and

Table 9: Summary of Key Challenges in AI-Driven Open RAN and Proposed Solution Approaches

Challenge	Example Issues	Solution Directions	Evaluation Metrics
Scalability	High computational load, latency constraints	Model compression, distributed/federated learning [36]	Throughput, latency, resource utilization
Privacy	Data leakage, poisoning attacks	Federated learning with secure aggregation [**?**]	Privacy leakage levels, model robustness
Interpretability	Black-box decisions, lack of trust	Explainable AI methods, attention-based models [**?**]	Explanation fidelity, operator trust scores
Security	Adversarial attacks on AI modules	Anomaly detection, robust training, secure updates [**?**]	Attack detection rate, false positives, system resilience

algorithmic transparency [18, 25]. These requirements impose stringent constraints that shape the design and operation of AI models within heterogeneous and multi-jurisdictional environments. To foster trust and accountability, explainable AI frameworks and comprehensive auditability mechanisms become indispensable, particularly as policies evolve dynamically across different global operators. Integrating policy-aware AI model design is critical to ensure compliance while maintaining system performance and user privacy.

4.4.6 Top Research Priorities. Based on the discussion above, the three most critical research priorities are: (1) developing scalable and explainable AI algorithms that ensure stable learning and adaptation in time-varying environments, addressing challenges like real-time inference latency and model convergence highlighted in [18, 25]; (2) creating efficient edge AI solutions via co-designed hardware and lightweight models to meet stringent latency, energy, and computational constraints, as emphasized in [6, 25]; and (3) advancing interoperable AI frameworks coupled with rigorous security and privacy safeguards conforming to regulatory mandates, particularly considering multi-vendor support and privacy-preserving distributed learning techniques noted in [24, 25]. Addressing these priorities requires multidisciplinary initiatives integrating communications theory, AI, hardware design, and regulatory insight to foster robust, adaptive, and trustworthy wireless networks.

In summary, overcoming the outlined challenges demands both theoretical innovations and practical engineering. Integrating insights from [6, 18, 24, 25] reveals that while AI significantly enhances Open RAN capabilities, realizing these benefits requires robust convergence techniques, efficient deployment of AI at the network edge, seamless multi-vendor interoperability, and comprehensive security and governance models. Only through such coordinated efforts can AI-driven Open RAN systems deliver adaptive, secure, and efficient wireless networks suited for 6G and beyond.

4.5 Future Research Directions

Future research must prioritize explainability and transparency of AI decision-making within Open RAN to build trust, satisfy regulatory requirements, and facilitate efficient troubleshooting. Explainable AI (XAI) approaches will provide clear insights into AI-driven resource allocations and fault detection processes, which is especially crucial in multi-stakeholder environments where accountability and interpretability are paramount [25]. To measure progress, the development of standardized evaluation metrics and benchmarks tailored to Open RAN scenarios is essential. Such metrics could include quantifiable indicators like explanation fidelity, user trust scores, and reductions in fault resolution times. Furthermore, establishing pilot studies and experimental platforms that simulate multi-vendor Open RAN environments will be critical to validate XAI methods and assess their operational effectiveness under realistic conditions.

The creation of multi-agent collaborative learning frameworks is expected to significantly enhance distributed AI systems by enabling coordinated intelligence across the RU, DU, and CU layers. This coordination is vital to addressing the complex cross-layer optimization issues inherent to 6G networks, as demonstrated by recent research on networked sensing and AI-empowered interference mitigation [18].

Designing lightweight AI models specifically optimized for resource-constrained edge units remains crucial to overcoming computational and energy limitations. Complementing these models with dedicated hardware accelerators further alleviates processing bottlenecks, enabling real-time, efficient AI deployment at the network edge [24]. Emerging paradigms such as quantum computing present promising prospects for solving complex optimization problems in Open RAN, potentially exceeding classical methods in both speed and scale. Additionally, blockchain technologies can enhance security, ensure data integrity, and support decentralized trust mechanisms—key enablers for robust multi-vendor Open RAN ecosystems [25].

To structure these pursuits effectively, future research goals can be categorized into short-, mid-, and long-term milestones. The short-term focuses on developing robust explainable AI models accompanied by appropriate evaluation metrics, conducting pilot studies on multi-agent collaborative frameworks to improve transparency and coordination, and addressing research questions such as: "How can XAI optimize AI decision interpretability without compromising performance in Open RAN?" and "Which mechanisms best facilitate multi-agent collaboration across distributed RAN units?" Mid-term objectives emphasize creating and deploying lightweight AI algorithms integrated with hardware accelerators to meet the stringent demands of edge computing environments. Research may explore trade-offs between model complexity and latency in constrained settings. Long-term ambitions target integrating emerging technologies such as quantum computing and blockchain to revolutionize optimization, security, and trustworthiness in Open RAN. Achievements in these phases will yield measurable improvements, including enhanced AI interpretability, reduced computational latency, increased network resilience, and stronger multi-vendor interoperability.

Integrating these advanced technologies with AI capabilities stands to significantly advance Open RAN, paving the way toward fully autonomous, resilient, and high-performance next-generation wireless networks.

Table 10: Summary of Challenges in AI-Driven Open RAN and Potential Mitigation Strategies

Challenge	Description	Potential Mitigation Strategies
Model convergence	Instability of AI models in dynamic, partially observed wireless environments [25]	Design robust reinforcement learning algorithms; real-time adaptation mechanisms; multi-agent collaboration
Edge resource constraints	Limited computation and energy in RUs/DUs hindering complex AI inference [18, 25]	Lightweight AI architectures; hardware-software co-design; specialized AI accelerators
Multi-vendor interoperability	Diverse hardware and data formats impeding unified AI deployment [6, 25]	Middleware frameworks; adoption of open standards (e.g., O-RAN, 3GPP); standardized AI model interfaces
Security and privacy	Vulnerabilities to adversarial attacks and data leakage in federated/distributed learning [24, 25]	Real-time intrusion detection; robust and privacy-preserving learning protocols; federated learning with differential privacy
Regulatory compliance	Compliance with data sovereignty, privacy laws, and need for transparency [18, 25]	Explainable AI; auditability frameworks; policy-aware AI model design

5 Large Language Model-Driven Agentic AI for O-RAN Network Resilience

This section provides a comprehensive overview of integrating Large Language Models (LLMs) within agentic AI frameworks to enhance the resilience of Open Radio Access Network (O-RAN) systems. We first detail the architectural design of LLM-driven agents, emphasizing modular components such as perception, reasoning, decision-making, and actuation in dynamic network environments. We then analyze their operational roles in fault detection, diagnosis, and mitigation within the O-RAN architecture, illustrating how these agents can dynamically adapt to changing network states using scenario-based reasoning and continual learning techniques.

Mitigation strategies employed by LLM-driven agents include anomaly detection based on contextual understanding, proactive fault anticipation through predictive modeling, and collaborative resolution leveraging multi-agent coordination. For example, an agent detecting an unusual traffic surge can autonomously initiate resource reallocation while alerting neighboring agents to prevent cascading failures. These techniques enable timely, context-aware responses that improve fault tolerance and reduce service disruption.

Compared to traditional AI methods such as rule-based systems or static machine learning models, LLM-driven agents offer superior generalization, interpretability, and adaptability. However, real-time edge deployment introduces challenges including latency constraints, computational resource limitations, and dataset diversity for training robust models. Addressing these requires optimized model compression, federated learning to incorporate diverse data sources while preserving privacy, and edge-cloud hybrid architectures balancing inference speed with model complexity.

From a deployment perspective, cost considerations involve trade-offs between infrastructure investment and service quality improvements. We discuss feasible implementation paths leveraging existing O-RAN components and open interfaces, enabling incremental integration of LLM-driven agents without disruptive overhauls. Furthermore, emerging paradigms such as combining LLMs with reinforcement learning or symbolic reasoning hold promise for more explainable and efficient autonomous resilience mechanisms.

In summary, LLM-driven agentic AI represents a transformative approach for achieving resilient O-RAN networks by enabling proactive, adaptive, and context-aware network management. While challenges remain, especially regarding and real-time edge constraints and dataset diversity, ongoing advancements in model optimization and hybrid architectures are poised to address these gaps, paving the way for widespread adoption of autonomous resilience in future O-RAN deployments.

5.1 Overview and Objectives

Agentic AI systems possess autonomous capabilities including perception, reasoning, and decision-making, enabling independent operation toward specified objectives. Enhanced with Large Language Models (LLMs), these agents gain advanced natural language understanding and contextual reasoning, which facilitate comprehensive analysis and interpretation of complex network states. Within the O-RAN framework, the primary roles of LLM-driven agents encompass proactive fault and anomaly detection, precise root-cause diagnosis, and automated mitigation via dynamic reconfiguration commands. By integrating these functions, the agents seek to enhance network reliability, adaptability, and scalability through intelligent and scalable control strategies.

5.2 Architectural Design and Data Flow

At a high level, the agentic AI architecture integrates multi-source data inputs, including telemetry metrics and performance indicators, which the LLM processes to establish comprehensive situational awareness. This enriched context informs decision-making modules tasked with selecting optimal actions to influence the network state through O-RAN control interfaces. A continuous feedback loop facilitates ongoing learning and adaptability, enabling the agent to respond effectively to evolving network conditions and maintain operational relevance. The structured data flow between components supports adaptive control strategies while preserving transparency and interpretability in system operations.

5.3 Comparative Analysis with Alternative AI Approaches

Unlike traditional machine learning and rule-based systems commonly used for fault management, LLM-driven agentic AI offers enhanced reasoning capabilities and greater flexibility. These systems enable more nuanced interpretation of network anomalies and context-aware decision-making that extends beyond simple pattern recognition. In contrast to established methods, however, LLM-driven approaches face challenges related to computational overhead and real-time responsiveness, particularly in resource-constrained edge deployments. Effectively balancing these trade-offs requires careful consideration of factors such as operational scale, latency requirements, and available computational resources to optimize performance across diverse deployment scenarios.

5.4 Deployment Feasibility and Cost Considerations

Implementing LLM-based agents in edge O-RAN environments raises practical concerns including hardware constraints, latency sensitivity, and resource consumption. While LLMs deliver superior

reasoning capabilities, their computational demands often exceed the capacities of typical edge devices if no optimization strategies are employed. Techniques such as model pruning, quantization, and knowledge distillation can significantly reduce model size and computational overhead, making deployment more practical. Additionally, cost implications arise from ensuring robust and low-latency connectivity, as well as ongoing maintenance requirements, which are critical factors in operational expenditure. Future research should focus on developing lightweight model adaptations and hybrid architectures that combine the reasoning strength of LLMs with the efficiency and reliability of conventional control algorithms to enable feasible and cost-effective deployment in edge O-RAN scenarios.

5.5 Emerging Paradigms and Future Directions

The integration of LLM-driven AI with emerging technologies such as quantum computing and blockchain presents promising avenues to further enhance network resilience. Quantum computing could accelerate complex reasoning tasks, while blockchain may ensure secure, tamper-evident agent interactions and decision audit trails. Exploring these synergies may lead to the development of more robust and trustworthy autonomous systems capable of managing next-generation O-RAN infrastructures.

Looking ahead, LLM-driven agentic AI offers a powerful framework for autonomous, adaptive network control within O-RAN. Addressing challenges in deployment, scalability, and cost-effectiveness remains critical. Furthermore, investigating complementary technologies and enhancing the transparency, security, and explainability of such systems will be vital for their real-world adoption.

Cross-references to the following subsections detail the structural design of agent architectures and the data flow mechanisms that underpin these autonomous and context-aware operations, thereby illuminating practical pathways for implementing LLM-driven resilience strategies in complex network environments.

5.6 Embedding LLM-Based Agents in RAN Intelligent Controller and SMO

The integration of Large Language Model (LLM)-based agents within the Open Radio Access Network (O-RAN) architecture—specifically within components such as the near Real-Time RAN Intelligent Controller (Near-RT RIC) and Service Management and Orchestration (SMO)—marks a significant advancement toward achieving autonomous fault management. Embedding these agents enables continuous, context-aware monitoring of diverse network telemetry data combined with dynamic remediation strategies executed without human intervention. This autonomy surpasses traditional rule-based or supervised learning systems, which typically depend on predefined fault catalogs or manual threshold triggers. Leveraging LLMs' intrinsic ability to parse and synthesize heterogeneous network state information, agentic AI systems can interpret a broad spectrum of fault manifestations and implement tailored corrective actions such as dynamic resource re-allocation and service re-configuration, all while adhering to stringent near-RT latency constraints [5].

Experimental evaluations performed on a simulated O-RAN testbed demonstrate that this LLM-driven agentic approach achieves a fault detection accuracy of 95%, a mitigation success rate of 91%, reduces network downtime by 40%, and decreases throughput degradation from 25% to 10% compared to conventional methods [5]. Table 11 summarizes key performance improvements over baseline techniques, showcasing the method's enhanced robustness and efficacy in handling complex fault scenarios. Furthermore, the modularity and open interfaces inherent to the O-RAN framework facilitate seamless integration of LLM-based agents, enabling customizable behaviors and scalable deployments that significantly enhance operational flexibility and reduce mitigation times.

Despite these promising results, several challenges remain critical. These include the computational overhead of deploying LLMs within near-RT environments, the risk of occasional erroneous decisions impacting network stability, security vulnerabilities exposing the system to adversarial threats, and interoperability issues arising from heterogeneous vendor implementations. Addressing these challenges involves strategic solutions such as hierarchical agent architectures to distribute computational loads efficiently, rigorous decision validation frameworks to enhance reliability, and robust security mechanisms to safeguard against attacks. Future research directions focus on optimizing LLM agents for deployment at edge environments, enhancing coordination in multi-agent scenarios, incorporating explainability features to improve transparency and trustworthiness, and reinforcing security robustness. Collectively, these advancements underscore the promising role of LLM-based agents in strengthening O-RAN self-healing capabilities and operational resilience in next-generation networks.

5.7 Natural Language Processing for Fault Interpretation and Interaction

A critical enabler of LLM-driven agentic AI efficacy is the application of advanced natural language processing (NLP) techniques for fault interpretation and human-machine interaction. Unlike traditional, deterministic fault detection frameworks that rely solely on numeric alarms or discrete indicators, LLMs process heterogeneous data outputs—including logs, alerts, and operator annotations—with nuanced semantic comprehension, enabling more sophisticated fault diagnosis. This advanced capability allows agents not only to detect and localize faults but also to contextualize root causes within operational narratives, thereby facilitating coherent, meaningful communication with human operators [5].

Such NLP-enabled interaction enhances transparency and fosters operator trust—vital factors given the inherent risks of erroneous decisions in fully autonomous systems. Moreover, these agents can articulate mitigation strategies, justify their decisions, and incorporate real-time operator feedback, ensuring integration of human oversight alongside agent autonomy. Experimental results from a simulated O-RAN environment demonstrate that this approach increases fault detection accuracy from 78% to 95% and improves mitigation success rates from 70% to 91%, significantly reducing network downtime by 40% and decreasing throughput degradation from 25% to 10% [5]. Table 14 summarizes these performance improvements, illustrating the substantial gains achieved by incorporating LLM-driven NLP for fault interpretation.

Table 11: Performance Comparison of LLM-Based Agentic AI Approach vs. Baseline in O-RAN Fault Management

Metric	Baseline	Proposed LLM-Based Agent	Improvement
Fault Detection Accuracy	78%	95%	+17%
Mitigation Success Rate	70%	91%	+21%
Downtime Reduction	-	40%	-
Throughput Degradation	25%	10%	-15%

This integrative process addresses critical concerns related to model interpretability, acceptance, and operational resilience during live network operations. Challenges such as computational overhead, potential erroneous decisions, security concerns, and vendor interoperability remain important considerations. To mitigate these, hierarchical agent designs and rigorous validation methods have been proposed, promoting reliable and transparent AI-human collaborative fault management [5]. Future directions include optimizing LLMs for edge deployment, enhancing multi-agent coordination, incorporating explainability features, and fortifying security, further advancing the robustness and applicability of LLM-driven agentic AI in next-generation network environments.

5.8 Experimental Achievements

Empirical evaluations demonstrate that integrating LLM-based agentic AI within the O-RAN architecture significantly enhances fault management capabilities. Experimental results indicate fault detection accuracy reaching up to 95%, representing a substantial improvement over baseline accuracies of approximately 78% [3, 5, 14]. Mitigation success rates improve by more than 20%, while network downtime is reduced by nearly 40%. Additionally, throughput degradation decreases from about 25% in baseline systems to approximately 10%, illustrating more efficient and resilient network operation [5].

These enhancements are attributed to the agentic AI's capability to proactively anticipate faults and execute diverse, context-aware responses, outperforming traditional heuristic or static rule-based methods—which often suffer from delayed or partial fault handling [14]. Improvements in throughput further reflect real-time resource optimization and adaptive network reconfiguration autonomously performed by agents embedded within the RIC and SMO layers. This validates both the practical feasibility and operational effectiveness of the proposed architecture [5].

Table 14 summarizes the key performance improvements observed experimentally, highlighting significant gains in detection accuracy, mitigation success, downtime reduction, and throughput degradation compared to baseline O-RAN implementations.

These results underscore the significant potential of LLM-driven agentic AI to enhance the resilience of next-generation wireless networks by enabling intelligent, autonomous fault management that dynamically adapts to network conditions. While challenges such as computational overheads, potential erroneous decisions, security considerations, and vendor interoperability remain, ongoing work focuses on mitigation strategies including hierarchical agent design and rigorous validation to address these issues [5].

5.9 Comparative Performance Analysis

When compared to conventional fault management techniques reliant on manual or semi-automated processes, LLM-driven agentic AI exhibits superior adaptability and resilience. Traditional methods frequently fail to capture the intricate interdependencies and temporal dynamics inherent in multi-source network data, resulting in suboptimal fault isolation and extended recovery times. In contrast, LLM agents synthesize multimodal inputs and apply contextual reasoning to enable expedited and more accurate fault classification and resolution pathways [5]. Furthermore, experimental evaluations conducted in simulated O-RAN environments demonstrate that LLM agentic AI achieves a fault detection accuracy of 95%, a mitigation success rate of 91%, and reduces downtime by 40%, significantly outperforming baseline approaches. Throughput degradation is also lowered from 25% to 10% in tested scenarios, illustrating enhanced network performance during fault conditions.

Notably, the continuous learning capabilities embedded in these agent architectures promote sustained performance improvements by dynamically adapting to evolving network topologies and traffic profiles. This adaptive proficiency positions agentic AI as a markedly superior solution for managing the increasing heterogeneity and scale of next-generation wireless infrastructures. Additionally, by integrating LLM-based agents within key O-RAN components such as the Near-RT RIC and SMO, the system autonomously monitors network telemetry, interprets faults through natural language processing, and executes mitigation strategies including dynamic resource re-allocation and self-healing operations [5]. These capabilities collectively contribute to enhanced resilience and robustness beyond that achievable with traditional approaches.

5.10 Recognized Challenges

Despite these significant advancements, deploying LLM-based agentic AI within O-RAN architectures poses several critical challenges. First, the computational overhead inherent to large-scale LLM inference raises pressing concerns regarding latency and energy efficiency, especially within resource-constrained edge environments [5, 18]. Addressing this issue requires targeted optimization of LLM architectures to enable real-time, low-power operation suitable for edge deployment. For example, adaptive model pruning and efficient inference techniques are being explored to meet the stringent latency requirements while conserving power, as demonstrated in AI-driven interference management and network sensing applications [5, 18].

Second, the potential for inaccuracies arising from incomplete or noisy input data, entrenched model biases, or adversarial conditions threatens network stability by causing erroneous decision-making.

Table 12: Performance Comparison of LLM-driven Agentic AI in O-RAN Fault Management [5]

Metric	Baseline	Proposed	Improvement
Fault Detection Accuracy	78%	95%	+17%
Mitigation Success Rate	70%	91%	+21%
Downtime Reduction	-	40%	-
Throughput Degradation	25%	10%	-15%

Table 13: Performance Comparison of Agentic AI-Enabled O-RAN vs. Baseline Systems

Metric	Baseline System	Agentic AI-Enabled O-RAN	Improvement
Fault Detection Accuracy	78%	95%	+17%
Mitigation Success Rate	70%	91%	+21%
Network Downtime Reduction	-	40%	-
Throughput Degradation	25%	10%	-15%

Table 14: Performance Comparison between Conventional Fault Management Baseline and LLM-Driven Agentic AI [5]

Metric	Baseline	LLM-Driven Agentic AI	Improvement
Fault Detection Accuracy	78%	95%	+17%
Mitigation Success Rate	70%	91%	+21%
Downtime Reduction	-	40%	-
Throughput Degradation	25%	10%	-15%

Mitigating these risks necessitates robust data preprocessing techniques, stringent model validation protocols, and integration of advanced adversarial resilience mechanisms to maintain system reliability. Employing hierarchical agent designs with continual learning and validation cycles has shown promise in reducing erroneous outcomes and ensuring trustworthy operation [5].

Third, AI-specific security vulnerabilities—including data poisoning, model inversion, and adversarial attacks—demand comprehensive, multilayered safeguards to protect data integrity and preserve model confidentiality. Leveraging secure data pipelines, anomaly detection modules, and enhanced encryption within agent communication channels are critical components of such defenses.

Lastly, ensuring seamless interoperability of LLM agents within heterogeneous, multi-vendor ecosystems characterized by proprietary interfaces and diverse data semantics remains an unresolved challenge, complicating standardized deployment [5, 18]. Hierarchical and modular agent design combined with rigorous validation protocols and standard-compliant APIs offer a viable path forward to manage this complexity.

Collectively, these challenges underscore the multifaceted difficulty of operationalizing agentic AI at scale while maintaining network performance, reliability, and security.

5.11 Proposed Solutions and Optimizations

To address the challenges previously outlined, several strategies have been developed to optimize the deployment and performance of LLM-driven agentic AI in O-RAN networks. A key approach is the hierarchical agent design paradigm, which organizes AI agents into layers with distinct roles. Lightweight agents are deployed at the edge to handle real-time, low-level tasks requiring minimal computation, while more complex and computationally intensive operations performed by large language models (LLMs) are handled by centralized or cloud-based environments. This hierarchical structure balances computational load and latency, making it suitable for resource-constrained edge environments [5].

For example, an edge agent may monitor local telemetry data and detect anomalies, then escalate complex fault interpretation and mitigation recommendations to a centralized LLM agent with greater processing power. This layered approach not only improves responsiveness but also enhances scalability and reliability.

To ensure robustness, rigorous validation frameworks employ simulated fault injections and continuous model retraining. These approaches help reduce erroneous actions by the AI agents, resulting in increased reliability of autonomous network operations. Furthermore, advanced optimization techniques—such as model pruning, quantization, and knowledge distillation—are applied to LLMs to reduce their computational demands without sacrificing accuracy, facilitating deployment on resource-limited edge devices.

Within the O-RAN architecture, the adoption of standardized open interfaces and semantic data models promotes interoperability and seamless integration across different vendors. This standardization addresses the inherent complexity of multi-vendor ecosystems and enables consistent operation and management of agentic AI components [5].

Overall, these combined strategies establish a comprehensive and scalable framework for practical deployment of LLM-driven agentic AI. They enable significant enhancements in network resilience,

fault detection, and self-healing capabilities as demonstrated in recent experimental results. Specifically, performance improvements include fault detection accuracy increasing from 78% to 95%, mitigation success rate rising from 70% to 91%, a 40% reduction in network downtime, and a decrease in throughput degradation from 25% to 10% [5]. Table 15 summarizes these important gains, underscoring how the proposed optimizations contribute to improved O-RAN self-healing and operational reliability.

5.12 Future Directions

This subsection outlines clear objectives and a detailed roadmap for advancing agentic AI capabilities in complex O-RAN ecosystems. The primary objective is to develop scalable, resilient, secure, and interpretable AI frameworks that enable autonomous, real-time adaptation in distributed and resource-constrained network environments. Specifically, future research should aim to enhance agent collaboration, improve security mechanisms, and optimize resource allocation through intelligent decision-making.

To illustrate these directions, consider use cases such as dynamic spectrum management where multiple agents must coordinate under changing channel conditions, or security threat detection requiring real-time anomaly identification and response. These scenarios highlight the critical need for robust agentic architectures capable of operating with minimal human intervention.

Key challenges to address include ensuring scalability across heterogeneous network nodes, maintaining resilience under attacks or failures, and achieving explainability to foster trust and regulatory compliance. Alternative approaches such as centralized versus fully distributed control frameworks present trade-offs in latency, overhead, and robustness; thus, comparative studies are essential to guide effective design choices.

In summary, the roadmap focuses on the integration of interdisciplinary methodologies combining reinforcement learning, game theory, and network optimization. Summarizing the key future directions:

- Develop multi-agent coordination algorithms tailored for O-RAN's distributed architecture. - Implement advanced security frameworks embedding agent-based intrusion detection and mitigation. - Design interpretable AI models to enhance transparency and user trust. - Optimize real-time resource management under constrained computational resources. - Conduct extensive simulation and field trials to validate agentic AI deployments.

Addressing these objectives will advance autonomous network management capabilities, enabling O-RAN to meet the demands of next-generation wireless environments effectively.

5.12.1 Multi-Agent Coordination. Future research should focus on enabling efficient and robust cooperation among distributed AI agents for fault detection and mitigation across RAN nodes. Key objectives include designing synchronization protocols that maintain ultra-low latency and ensuring backward compatibility without performance losses. Concrete milestones involve developing lightweight consensus mechanisms and experimentally validating their scalability in large-scale deployments [37]. Additionally, addressing computational complexity in real-time processing remains critical to practical applicability. Challenges such as synchronization overhead and compatibility with existing 5G standards must also be tackled, as highlighted by recent advancements in high-capacity preamble sequences that balance detection performance and resource usage while minimizing collisions [37].

5.12.2 Explainability and Operator Trust. Advancing explainability techniques is essential for fostering operator confidence and ensuring regulatory compliance. Research efforts should focus on integrating AI interpretability with control theory and wireless signal processing to create hybrid models that provide actionable insights with minimal latency [18]. A key challenge is balancing the complexity of AI models with the clarity required for operator understanding, alongside developing visualization tools specifically designed to fit operator workflows. Short-term objectives include designing operator-centric dashboards and conducting rigorous user studies to evaluate their effectiveness in real-world settings. Additionally, leveraging AI-empowered interference mitigation frameworks, as explored in recent networked sensing studies [18], can enhance transparency by providing interpretable feedback on system performance under dynamic conditions, ultimately strengthening operator trust.

5.12.3 Security and Robustness. Safeguarding O-RAN AI agents against evolving cyber threats is paramount. Priorities include developing adversarially robust training methods, secure and efficient model update protocols, and real-time anomaly detection at the network edge [24]. Addressing the limitations of agentic AI in telecom systems requires mitigation strategies such as incorporating fail-safe mechanisms to handle unexpected agent behaviors, ensuring human-in-the-loop oversight for critical decisions, and designing AI agents with interpretable and verifiable decision-making processes to enhance trust and transparency. Research should also quantify overhead thresholds for security mechanisms under typical edge resource constraints and anticipate emerging attack vectors in dynamic network topologies. Milestones incorporate deploying prototype defenses in realistic O-RAN testbeds and benchmarking their impact on network performance.

5.12.4 Deployment Architectures. Addressing the tradeoffs between accuracy, latency, and resource utilization necessitates novel architectural frameworks that synergize edge and cloud computing [24]. Effective deployment architectures must orchestrate decentralized AI inference at the edge with centralized model updates in the cloud, while minimizing communication overhead and ensuring data privacy. Key research challenges include designing standardized interfaces for modular AI components, optimizing resource allocation dynamically, and empirically evaluating system-level latency and throughput under varying network traffic conditions. Future directions emphasize integrating reinforcement learning for adaptive resource management, leveraging explainable AI to improve system interpretability, and enhancing robustness against adversarial and uncertain network environments, thereby enabling scalable and efficient AI-driven network management and optimization.

5.12.5 Adaptive Resource Allocation and Network Autonomy. To support massive IoT deployments and dynamic network conditions, there is a pressing need to transition from static policies to fully autonomous, self-optimizing networks [37]. One promising direction integrates reinforcement learning (RL) techniques for dynamic

Table 15: Performance Improvements of LLM-Driven Agentic AI in O-RAN Self-Healing [5]

Metric	Baseline	Proposed	Improvement
Fault Detection Accuracy	78%	95%	+17%
Mitigation Success Rate	70%	91%	+21%
Downtime Reduction	-	40%	-
Throughput Degradation	25%	10%	-15%

resource allocation with forecasting and anomaly detection frameworks that can effectively handle environmental uncertainties and temporal dependencies. For example, the advancements in high-capacity preamble sequence design in 5G IoT networks [37] highlight the critical role of adaptive resource management to reduce collision probabilities and improve detection rates, which RL can further optimize in real-time. Immediate research priorities include the creation of standardized benchmark datasets and simulation platforms to rigorously evaluate the adaptability, scalability, and robustness of RL algorithms in heterogeneous and time-varying network scenarios. Addressing these challenges is essential to enable self-autonomy in next-generation IoT ecosystems and to maximize network performance under unpredictable conditions.

5.12.6 Implementation Pathways and Benchmarking. A concrete implementation pathway entails developing open-source toolkits and standardized benchmarks to evaluate agentic AI methods in O-RAN contexts. These initiatives will accelerate reproducibility, facilitate rigorous cross-comparison of diverse techniques, and foster interdisciplinary community collaboration. Proposed benchmarks should comprehensively cover key performance metrics such as latency, throughput, energy efficiency, security resilience, and explainability quality, all aligned with the multifaceted challenges summarized in Table 16.

In summary, these directions collectively form an interconnected roadmap that explicitly links practical implementation considerations to the theoretical challenges covered in earlier sections. Achieving these objectives will require sustained interdisciplinary collaboration, the design of scalable and modular system architectures, and iterative empirical validation. Together, these efforts will drive the realization of sustainable, trustworthy, and high-performance AI-enabled O-RAN networks that meet evolving operational demands.

6 Adaptive Control and Reinforcement Learning in Networking Systems

This section provides a structured overview of adaptive control and reinforcement learning (RL) methods in networking systems, emphasizing their roles, trade-offs, and integration within modern communication environments such as O-RAN and 6G networks. We clarify foundational concepts, highlight key case studies, and explore the synergy between gradient-based adaptive control and RL in dynamic network management.

6.1 Overview of Adaptive Control Methods

Adaptive control techniques dynamically adjust control policies based on observed network conditions to optimize performance metrics such as throughput, delay, and resource utilization. A major class of these methods relies on gradient-based approaches, which use gradient information of a performance metric or cost function with respect to control parameters to iteratively improve system behavior. For example, in network congestion control, gradient descent adjusts sending rates to optimize throughput and minimize delay, responding effectively to changing network dynamics. The primary strengths of gradient-based methods include straightforward implementation and well-understood convergence properties under smooth cost landscapes. However, they can struggle when the cost function is complex, non-differentiable, or the network environment exhibits high stochasticity and non-stationarity, limiting their applicability in highly dynamic settings.

6.2 Reinforcement Learning Approaches

Reinforcement learning extends adaptive control by enabling network agents to learn optimal control policies through environment interaction without relying on explicit models. This model-free nature suits complex or partially observable networks where accurate modeling is infeasible. Classical RL algorithms, such as Q-learning and policy gradient methods, have been successfully applied to routing optimization, resource allocation, and energy-efficient management in wireless networks. Despite their flexibility, RL techniques face limitations including sample inefficiency, challenges in scaling to high-dimensional state and action spaces, and ensuring safety and robustness during exploratory phases, which is especially critical for mission-critical network applications.

6.3 Case Study: Adaptive Routing with Reinforcement Learning

Consider an RL agent tasked with adaptive routing to minimize latency and packet loss over dynamically changing network topologies. Using policy gradient methods, the agent incrementally refines its routing strategy based on reward signals corresponding to successful data delivery. This approach enables the agent to adapt routing decisions in real time, demonstrating how gradient-based optimization and RL can synergize to effectively handle stochastic network conditions.

6.4 Integrating Gradient-Based Adaptive Control with Reinforcement Learning

Combining gradient-based adaptive control with RL frameworks enhances performance by merging the sample efficiency and theoretical guarantees of gradient methods with the flexibility of RL policies. Hybrid algorithms have shown promise in scenarios requiring rapid adaptation to fluctuating network states. Nonetheless,

Table 16: Summary of Future Research Challenges and Concrete Objectives in Agentic AI for O-RAN

Research Area	Key Challenges	Concrete Objectives and Milestones
Multi-Agent Coordination	Synchronization latency, backward compatibility, computational complexity	Design scalable low-latency synchronization protocols validated on large deployments; Develop lightweight consensus algorithms ensuring compatibility [37]
Explainability	Model complexity vs. interpretability, operator visualization, real-time constraints	Integrate control theory with signal processing for interpretable models; Prototype operator dashboards and conduct usability evaluations [18]
Security	Adversarial robustness, secure updates, edge anomaly detection	Implement adversarial training and secure model update protocols; Benchmark defenses in edge resource-constrained scenarios [24]
Deployment Architecture	Edge-cloud synergy, latency, resource management	Develop modular AI component interfaces; Measure latency and throughput in hybrid edge-cloud setups [24]
Adaptive Resource Allocation	Dynamic environments, temporal modeling, real-time learning	Create benchmarks and simulation platforms to test RL-based allocation; Model temporal dependencies and environmental uncertainty [37]

such integrations inherit limitations from their components, including increased computational complexity and sensitivity to hyperparameter tuning, which must be carefully managed for practical deployment.

6.5 Challenges in O-RAN and 6G Network Contexts

The deployment of RL and adaptive control within open radio access networks (O-RAN) and emerging 6G systems faces significant challenges. These include the curse of dimensionality inherent in state and action spaces, stringent demands for low-latency, real-time decision-making under partial observability, and maintaining robust operation in inherently open and heterogeneous infrastructures. Furthermore, balancing the exploratory learning necessary for policy improvement with strict safety requirements to avoid service degradation remains an open and critical issue. Addressing these challenges is a central motivation for ongoing research into scalable, safe, and interpretable RL algorithms tailored specifically to advanced network architectures.

6.6 Complementary Roles of Adaptive Control and LLM-Driven Agents

Adaptive control methods complement learning-based agents by ensuring continuous parameter tuning and system stability, vital for integrating RL agents into operational networks. This synergy is particularly relevant with the rise of large language model (LLM)-driven agent architectures, where LLMs provide high-level reasoning capabilities that guide RL agents' exploration and adaptation. The integration of adaptive control, RL, and LLM-driven intelligence holds promise for developing network management agents that are more responsive, interpretable, and robust, addressing the complexities of next-generation communication systems.

6.7 Summary

In summary, adaptive control and reinforcement learning together form a powerful toolkit for enhancing network performance and resilience. Gradient-based methods provide principled, continuous adaptation mechanisms, while RL enables effective operation in uncertain and partially known environments. The advancement and integration of these methodologies, especially with emerging LLM-driven architectures, produce robust and efficient network control solutions capable of meeting the evolving requirements of infrastructures such as O-RAN and 6G. Yet, critical challenges related to scalability, safety, and interpretability remain and must be addressed to fully realize their potential in real-world deployments.

The following sections build on these concepts by detailing specific gradient-based algorithms and RL case studies, illustrating their applications across diverse networking scenarios with improved clarity and depth.

6.8 Applications of Reinforcement Learning

Reinforcement learning (RL) has emerged as a crucial methodology for real-time adaptive control in dynamic and wireless networking environments. RL enables the optimization of system performance under stochastic and time-varying conditions by learning policies that map network states to appropriate actions through direct interaction with the environment [2, 17, 21, 24, 29, 32**? ? ? ?**]. This capability contrasts with traditional model-based control methods that depend on fixed policies or heuristics, offering autonomous decision-making tailored to the complexities inherent in modern networks.

RL's versatility is demonstrated across diverse networking scenarios, including cellular self-organized networks (SON) and multi-hop wireless ad hoc systems, where it addresses critical challenges such as interference mitigation, handover optimization, and load balancing [17, 21, 29**? ?**]. In particular, deep RL (DRL) methods leverage deep neural networks—such as convolutional and recurrent architectures—to approximate value functions or policies, enabling rapid adaptation without explicit model dependencies. This feature is especially vital in heterogeneous and uncertain wireless contexts, such as multi-band communication networks, which utilize frequency bands from sub-6 GHz to millimeter-wave and terahertz bands with widely differing propagation and interference characteristics, complicating control and necessitating flexible RL-driven resource allocation strategies [32].

The effectiveness of RL-based adaptive control systems hinges on accurate and expressive state representations capable of handling high-dimensional and partially observable network environments. Recent literature highlights the synergy between RL and deep learning architectures—including convolutional, recurrent, and autoencoder networks—to extract pertinent features and exploit spatial-temporal correlations, thereby accelerating convergence and improving robustness [24**? ?**]. Additionally, advanced analytical frameworks for optimal control on networked systems, such as modal decomposition grounded in network spectral properties, complement RL approaches by providing interpretable and computationally scalable solutions [17, 21]. For example, modal decomposition techniques decouple network dynamics into independent eigenmodes, facilitating efficient control design with reduced complexity and enhanced scalability, often achieving significant cost reductions and faster computations compared to classical methods.

Despite these advances, enduring challenges remain. Maintaining robustness amid non-stationary traffic patterns and fluctuating wireless channels demands adaptive generalization capabilities, motivating research into meta-learning and transfer learning to enhance policy reuse across varying network states [29]. Moreover, deploying RL in latency-sensitive and resource-constrained environments is limited by the computational overhead of both inference

and training. To address this, efforts focus on lightweight model architectures, hardware acceleration, and hybrid optimization algorithms that blend RL with classical control and optimization methods [2**?**]. For instance, combining RL with learn-and-adapt stochastic dual gradient algorithms exploits the strengths of both learning-based adaptation and analytical optimization for improved network resource management and queue stability under uncertainty.

Collectively, these developments position reinforcement learning as a transformative tool for autonomous, efficient, and scalable management of increasingly complex wireless networks, particularly within emerging 5G and 6G paradigms where dynamic adaptability and intelligent resource allocation are paramount [24**?**]. Continued integration of RL with domain-specific knowledge, advanced optimization frameworks, and real-time analytical models promises further enhancements in performance, robustness, and interpretability of network control solutions.

6.9 Deep Reinforcement Learning for Online Adaptation

Deep reinforcement learning (DRL) enhances traditional reinforcement learning by utilizing deep neural networks as function approximators for policies or value functions, enabling effective online adaptation in complex networking tasks such as decision-making, resource allocation, and adaptive bandwidth management [20, 24**? ? ?**]. DRL is especially beneficial in high-dimensional state and action spaces where explicit policy design is impractical, covering applications like dynamic spectrum allocation, power control, and admission control [20**?**].

Through continuous interaction with the environment, DRL facilitates dynamic adaptation of resource management policies that respond to fluctuating network conditions, frequently outperforming heuristic or static methods. For example, in federated learning (FL) systems operating over unreliable and resource-constrained wireless networks, DRL-based frameworks have been combined with optimization techniques such as gradient sparsification and adaptive client selection, achieving a test accuracy of 87.5% on image classification tasks like MNIST and CIFAR-10, outperforming traditional FL methods while reducing communication costs substantially [**?**]. Such hybrid approaches demonstrate faster convergence and enhanced performance in bandwidth-limited and heterogeneous environments [24]. Moreover, embedding domain-specific knowledge into DRL architectures helps balance exploration and exploitation, a critical factor for real-time adaptability in dynamic networking scenarios.

Nevertheless, employing DRL in networking systems introduces challenges, including substantial requirements for training data, limited interpretability of learned models, risks of overfitting, and stability issues in non-stationary environments [**?**]. To mitigate these concerns, mechanisms such as experience replay buffers, target networks, and transfer learning are commonly applied. Achieving an optimal balance between model expressiveness and computational efficiency remains an active research focus, particularly given the stringent latency and scalability demands of emerging wireless networks. Future developments will likely center on adaptive AI models capable of real-time optimization, cross-layer integration, and enhanced robustness, aiming to fulfill ultra-low latency (under 1 millisecond) and massive connectivity objectives envisioned for beyond-5G and 6G systems [24**?**].

6.10 Challenges in Policy Design

The design of RL policies for networking systems necessitates a careful balance between exploration and exploitation in environments characterized by non-stationarity, such as wireless networks [18**? ? ?**]. Exploration is essential for discovering improved policies but can degrade performance and increase latency, which are critical concerns in mission-critical or ultra-reliable low-latency communication (URLLC) applications. Conversely, excessive exploitation risks converging to suboptimal policies when network traffic patterns or channel conditions change dynamically.

To mitigate these challenges, state-of-the-art techniques incorporate adaptive exploration rates that adjust based on environmental feedback, uncertainty-aware policy learning to account for incomplete and noisy information, and reward shaping that directly aligns with key networking performance metrics such as throughput, latency, and resource utilization [**? ?**]. For example, [**?**] formulates resource allocation using weighted sum mean-square error minimization where weights reflect data importance, significantly improving convergence speed and model accuracy compared to uniform policies. This illustrates how incorporating domain-specific metrics into reward design can enhance policy effectiveness.

The stringent low-latency inference demands in networking impose constraints on model complexity and motivate efficient state acquisition strategies. Accurate state information remains difficult to obtain due to measurement noise, delays, and partial observability inherent in distributed systems [18**?**]. Robust state estimation thus becomes essential, frequently realized through filtering methods or latent state representations learned jointly within RL frameworks. For instance, [18] demonstrates that deep learning-based interference prediction combined with coordinated resource allocation can improve sensing signal-to-interference-plus-noise ratio (SINR) by around 30%, while maintaining communication quality under high network loads.

Furthermore, to ensure that policies generalize effectively across heterogeneous devices and diverse, dynamic network environments without sacrificing responsiveness, meta-reinforcement learning and multi-agent RL paradigms have been proposed [**? ?**]. These approaches enable rapid adaptation and cooperative decision-making suited for complex next-generation network architectures, such as AI-assisted network slicing and integrated optical wireless communications, which demand both reliability and agility [**? ?**].

In AI-assisted network slicing, as highlighted in [**?**], RL dynamically manages resource allocation and slice admission control responding to fluctuating traffic and QoS requirements. The framework integrates federated learning and deep learning for traffic prediction and anomaly detection, achieving improvements over heuristic techniques in throughput, latency, resource utilization, and QoS provisioning. Similarly, advances in optical wireless communications (OWC) for 6G networks [**?**] introduce challenges including atmospheric effects, device heterogeneity, and system

alignment. Here, RL policies offer adaptability and robustness, addressing dynamic channel variability and integrating OWC with conventional RF systems.

Table 17 summarizes key algorithmic aspects of representative RL approaches for networking applications, highlighting their metrics, adaptation strategies, and performance improvements reported in literature.

These case studies underscore the importance of integrating domain-specific knowledge with advanced RL strategies to meet the rigorous demands of emerging network technologies and to balance AI model complexity with the operational constraints of ultra-reliable, low-latency network applications.

6.11 Federated and Distributed Reinforcement Learning

Federated reinforcement learning (FRL) and distributed reinforcement learning paradigms have attracted considerable interest in networking systems due to their capacity to enhance privacy, reduce computational burdens, and accelerate convergence within edge-cloud ecosystems [6, 24]. By decentralizing the training process, multiple clients—such as base stations or edge devices—collaboratively learn coordinated policies without sharing raw data, thereby inherently supporting privacy preservation and regulatory compliance.

Among various FRL frameworks, federated deep Q-networks (FedDQN) and federated actor-critic models stand out as prominent algorithms. In FedDQN, clients independently update local Q-values based on their experience, then periodically share model gradients or parameters with a central server performing weighted aggregation to form a global policy [6]. This approach effectively mitigates the impact of heterogeneous data distributions across clients by leveraging mechanisms such as error feedback and adaptive client weighting to balance contributions from devices with varying data quality and quantity [6]. Actor-critic variants extend this by combining policy evaluation and improvement steps locally, while federated aggregation synchronizes policies to ensure consistent decision-making across a distributed network.

To tackle communication constraints typical of bandwidth-limited wireless environments, techniques like gradient sparsification, quantization, and adaptive client selection schemes are deployed to minimize communication overhead without compromising convergence speed [6]. Integrated resource allocation approaches further optimize bandwidth and computational resource assignment concurrently, resulting in accelerated training convergence and improved accuracy in FL-based wireless networks [24].

Empirical studies have demonstrated that FRL frameworks sustain robust policy performance despite challenges such as client dropout, system heterogeneity, and dynamic network conditions. For instance, experimental evaluations highlight that federated RL can outperform traditional centralized RL in coverage and robustness, particularly when combined with emerging technologies like reconfigurable intelligent surfaces (RIS) to adapt to complex channel environments [6].

Nonetheless, significant challenges persist, including synchronization of distributed RL agents, mitigating delays from straggling clients, and defending against adversarial attacks. Future research is expected to focus on developing asynchronous FRL algorithms to enhance training efficiency, incorporating stronger privacy-preserving techniques such as differential privacy to protect sensitive information, and designing scalable architectures capable of handling the complexity of forthcoming 6G networks. These directions align with broader AI-driven network optimization objectives and are essential for fully realizing FRL's potential in next-generation wireless environments [24].

6.12 Integration Across Networking Frameworks

The efficacy of reinforcement learning (RL) and adaptive control techniques is significantly enhanced when integrated with complementary AI-driven networking frameworks such as network traffic classification, software-defined networking (SDN), and routing optimization [13, 16, 39]. Deep learning-based traffic classification models provide granular insights into network flow characteristics, effectively overcoming the limitations of traditional methods—such as port-based and deep packet inspection approaches—that struggle with encrypted and dynamic traffic patterns. These advanced models enable RL controllers to optimize resource allocation with greater accuracy and adaptability by prioritizing traffic types based on detailed classification results [16]. Despite their high accuracy, these models often entail substantial computational overhead, emphasizing the need for optimized, lightweight implementations to maintain real-time feasibility.

Within SDN architectures, RL-powered controllers dynamically adjust routing and admission control policies in response to real-time network state changes, thereby improving throughput, reducing latency, and enhancing fault tolerance. AI integration into SDN controllers combines supervised learning classifiers (e.g., Random Forest, SVM) and deep learning models (e.g., LSTM networks) to perform real-time traffic classification, anomaly detection, and dynamic resource allocation. This integration results in significant performance improvements, as demonstrated by up to 92% traffic classification accuracy, an 18% reduction in end-to-end latency, and increased throughput in 5G and beyond network scenarios [13]. Addressing the computational overhead associated with these AI modules involves developing modular, lightweight AI pipelines, algorithmic optimization, and potential edge computing deployment to distribute processing workloads.

Similarly, RL-based routing protocols adaptively select communication paths by continuously learning from traffic dynamics, balancing load, mitigating congestion and failures, and thereby enhancing network resilience and efficiency [39]. These AI-driven routing methods have achieved up to 30% improvements in throughput and latency, enabling robust adaptation in complex network environments. However, the complexity of integrating multiple AI components across frameworks raises security concerns, including adversarial AI threats and potential cascading failures. Securing integrated AI modules requires incorporating robust defense mechanisms, anomaly detection, and explainable AI techniques to improve transparency and trustworthiness in autonomous network operations.

Together, these cross-framework synergies facilitate comprehensive network adaptation strategies, where RL agents leverage

Table 17: Summary of Key Algorithmic Approaches in RL-based Networking Control

Reference	Application	Key Metrics	Adaptation Method	Reported Performance Improvements
[?]	Resource Allocation for Distributed ML	Weighted sum MSE, convergence speed	Data-importance aware allocation with heuristic algorithm	Faster convergence; better accuracy compared to uniform/channel-based schemes
[18]	Interference Mitigation in Networked Sensing	Detection probability, SINR, interference reduction	Deep learning-based interference prediction, coordinated beamforming	+20% detection probability; 30% sensing interference reduction
[?]	AI-assisted Network Slicing	Throughput, latency, resource utilization, QoS	RL for dynamic slice admission and resource allocation; federated learning for traffic prediction	Outperforms heuristics in throughput and latency; improved resource efficiency
[?]	Optical Wireless Communications for 6G	Adaptability to channel variability, robustness	RL for adaptive policies integrating OWC and RF	Enhanced robustness to channel effects; supports hybrid optical-RF systems

enriched contextual information and explicit control channels provided by SDN. Nevertheless, challenges such as elevated computational overhead, interoperability complexities, and security vulnerabilities must be mitigated through standardization efforts and advances in privacy-aware federated and decentralized learning. These approaches hold promise for enhancing scalability and securing heterogeneous network environments while preserving data privacy.

6.13 Gradient-Based Optimization and Fast Algorithmic Updates

Gradient-based optimization methods are fundamental for training and adapting artificial intelligence models. These methods iteratively improve an objective function by following the gradient of the loss function. Common approaches include stochastic gradient descent (SGD), mini-batch gradient descent, and momentum-based methods. These techniques improve efficiency and stability during training.

Fast algorithmic updates make use of structural properties and approximations to speed up optimization. For example, algorithms that exploit sparsity, adapt learning rates dynamically, or approximate Hessian matrices help accelerate model updates with less computational effort. Such techniques are particularly important in large-scale and online learning settings where quick model adaptation is needed.

Combining efficient gradient calculations with fast update strategies has led to scalable frameworks that support real-time learning and responsiveness in complex models. Research continues to focus on improving the trade-off between computational speed and optimization accuracy, making AI systems more practical across various applications.

6.13.1 Gradient Descent and Variants. Gradient-based optimization constitutes a foundational approach for tuning control and network parameters in large-scale communication and data networks. Traditional gradient descent methods, alongside their accelerated variants, have proven effective for scalable optimization tasks. However, challenges such as high-dimensional uncertainty and the presence of integer decision variables considerably complicate these optimization processes. Specifically, while continuous control parameters allow for convergence guarantees under smoothness assumptions, incorporating integer or mixed-integer variables markedly increases computational complexity and complicates theoretical convergence analyses [34]. This challenge intensifies in settings characterized by large state spaces and expansive uncertainty sets, where computational demands grow exponentially with dimensionality.

To enhance scalability, recent algorithmic refinements such as stochastic gradient methods and adaptive learning rate schemes have been developed. These approaches enable efficient parameter updates even in vast, complex networks [7, 32, 35, 36, 38**? ?**]. Nonetheless, the inherently discrete nature of some optimization variables often necessitates hybrid or relaxation-based techniques, carefully balancing solution quality against computational tractability. The treatment of such integer-constrained optimization problems remains a dynamic research area, stimulating both theoretical advancements and practical algorithm design.

Future directions include integrating machine learning-based heuristics—such as deep learning models that extract hierarchical features from network data [38]—and adaptive partitioning methods [34] to better handle uncertainty and mixed-integer decision-making within complex communication networks. Such integrations hold promise for achieving improved robustness and efficiency in optimization under realistic network constraints, including those arising in next-generation wireless and software-driven environments, where adaptability and real-time performance are critical [7, 35]. Combining these techniques with autonomous network management and advanced resource allocation strategies can further enhance optimization performance in 5G/6G and beyond systems [32**? **].

6.13.2 Hybrid Model- and Data-Driven Gradient Approaches. Recognizing the limitations of purely gradient-driven methods, recent research has focused on hybrid frameworks that integrate model-based insights with data-driven adaptations. These approaches leverage structural knowledge embedded in network models while simultaneously exploiting real-time or historical data to inform adaptive gradient computations [2**? ?**]. This integration improves convergence speed and algorithmic flexibility by dynamically adjusting update rules and reducing discrepancies between model assumptions and evolving network conditions.

For example, in self-optimized wireless networks (SON), deep learning techniques are combined with model-based control mechanisms to robustly tune parameters across diverse and dynamic environments [36]. This hybrid approach employs knowledge graphs and semantic information extracted through deep neural networks and graph neural networks to provide a contextual understanding of network states, thereby enhancing system responsiveness and stability. By fusing data-driven semantic representations with traditional model-based optimization, these frameworks address scalability and convergence challenges inherent in complex, real-world networks. Furthermore, knowledge graphs enable semantic context consistency and error correction by inferring missing or distorted semantic elements, enhancing robustness against environmental variability [36]. Such hybrid methods thus present a promising direction for achieving intelligent, flexible, and resilient network optimization beyond conventional gradient-based techniques, paving the way for adaptive AI-driven wireless communication systems that leverage both structural and contextual insights.

Table 18: Summary of AI Integration Across Networking Frameworks: Benefits, Challenges, and Mitigation Approaches

Framework	AI Techniques	Key Benefits	Challenges	Mitigation Approaches
Network Traffic Classification [16]	Supervised ML (Random Forest, SVM), Deep Learning (CNN, RNN)	High accuracy in encrypted/dynamic traffic classification; Enables prioritized resource allocation	High computational cost; Data imbalance; Encryption limits payload visibility; Concept drift	Lightweight models; Semi- and federated learning; Edge computing; Explainability techniques
Software-Defined Networking (SDN) [13]	Supervised classifiers (Random Forest, SVM), LSTM networks	Up to 92% traffic classification accuracy; 18% latency reduction; Increased throughput; Real-time traffic management	Computational overhead; Dataset scarcity; Interoperability issues; Security risks	Modular pipelines; Algorithm optimization; Robust AI defenses; Standardization; Privacy-aware federated learning
AI-Driven Routing Optimization [39]	Reinforcement Learning, Neural Networks	Adaptive path selection; 30% throughput and latency improvement; Enhanced fault tolerance	Scalability; Model training data representativeness; Security vulnerabilities; Integration complexity	Decentralized/federated learning; Hybrid AI-conventional methods; Explainable AI; Network security hardening

Table 19: Summary of Gradient-Based Optimization Techniques and Fast Update Algorithms with Performance Aspects

Technique	Key Characteristics	Advantages	Typical Use Cases
Stochastic Gradient Descent (SGD)	Updates parameters using noisy gradient estimates from random samples	Computationally efficient; scalable to large datasets	Large-scale supervised learning
Mini-batch Gradient Descent	Uses small batches for gradient estimation	Balances convergence speed and stability	Deep neural network training
Momentum Methods	Incorporate exponentially weighted gradients to accelerate convergence	Faster convergence; fewer oscillations	Training deep networks, convex and non-convex optimization
Sparsity-Exploiting Algorithms	Utilize sparse data or parameter structures	Reduced computation and memory use	Large sparse models, online learning
Adaptive Learning Rate Methods	Adjust step sizes based on past gradients dynamically	Robust convergence; less hyperparameter tuning	Online learning; changing data distributions
Hessian Approximation Algorithms	Approximate second-order information for faster convergence	Faster convergence; handles ill-conditioned problems	Large-scale optimization; curvature-informed updates

6.13.3 Fast Algorithmic Update Techniques. The imperative for rapid recalibration of control policies in dynamic and stochastic network environments has motivated the development of fast algorithmic update methods focused on minimizing computational latency. Speed is critical for enabling online learning and real-time control systems [20, 35, 38**? ? ? ? ?**]. Typical approaches include incremental gradient updates that adjust parameters based on streaming data, warm-starting solvers with prior solutions to reduce convergence time, and employing approximation heuristics that offer computationally efficient yet effective parameter refinements.

In telecommunication networks, these techniques enable systems to swiftly adapt to sudden variations in traffic patterns, user demands, or channel conditions, thereby ensuring strict quality-of-service (QoS) guarantees and optimizing resource allocation efficiency [7]. For example, the emerging Tactile Internet paradigm imposes ultra-low latency requirements on the order of milliseconds, demanding update mechanisms capable of executing within extremely tight time constraints [7]. Such stringent requirements are characterized by a comprehensive latency model comprising transmission, propagation, processing, queueing, and retransmission delays, all of which together must stay below the 1 millisecond threshold to support applications like remote surgery, autonomous vehicles, and immersive virtual reality.

A fundamental challenge in fast algorithmic updates lies in balancing update speed against solution accuracy. Aggressive approximations may degrade policy effectiveness and risk violating QoS constraints, whereas fully precise updates might introduce computational delays that conflict with real-time system demands. To alleviate this tension, recent advances exploit parallel computation and distributed optimization frameworks, enabling decomposition of the update problem across multiple processing units or network nodes. These architectures improve scalability and responsiveness, facilitating timely updates without significantly compromising optimality [35**?**].

Moreover, fast update techniques are increasingly integrated within AI-driven control and optimization frameworks that underpin autonomic and self-optimizing networks [35**? ?**]. These hybrid approaches enhance real-time adaptability, robustness, and operational efficiency by combining machine learning-based predictive modeling and decision-making with rapid algorithmic recalibration to meet dynamic network conditions. Consequently, such innovations lay the groundwork for adaptive, autonomous network management capable of maintaining rigorous performance metrics in highly dynamic telecommunications environments.

6.13.4 Case Studies and Benchmarks. Empirical validations of gradient-based optimization and fast update techniques within real-world telecommunication networks provide critical insights into their practical efficacy and constraints. The primary objective of these validations is to quantify improvements in key performance indicators such as network throughput, latency, and resource utilization, and to identify limitations affecting scalability and robustness [7, 9, 36**? ?**].

Dynamic optimization strategies employing these methods have yielded measurable gains across diverse network scenarios. For example, the neural network-based information transfer (NNIT) approach effectively accelerates optimization convergence by transforming historical solutions into promising candidates. This is achieved by training neural networks to learn patterns of environmental evolution from both previous and new environments, enabling rapid adaptation to dynamic network conditions [**?**]. This method has demonstrated improvements in optimization speed and solution quality in wireless network optimization tasks.

In resource-constrained wireless settings, federated learning (FL) schemes leveraging gradient sparsification combined with error feedback and adaptive client selection have demonstrated significant enhancements in learning robustness and communication efficiency [**?**]. By minimizing a weighted global loss function and applying sparsification operators while maintaining error compensation, these techniques substantially reduce communication overhead. Moreover, adaptive client selection based on resource availability and reliability ensures effective participation despite high dropout rates. Experimental results on standard benchmarks such as MNIST and CIFAR-10 confirm that this approach achieves higher test accuracy (up to 87.5%) and faster convergence (150 rounds) compared to prior methods like FedAvg and FedProx, while halving communication costs.

Applications in ultra-low latency environments, exemplified by the 5G-enabled Tactile Internet, highlight the critical role of gradient-informed optimizations in fulfilling stringent end-to-end latency requirements [7]. The Tactile Internet relies on innovations such as Multi-access Edge Computing (MEC) and network slicing to keep the total delay—including transmission, propagation, processing, queueing, and retransmission—below one millisecond. Achieving this latency target enables transformative uses such as remote surgery and immersive virtual reality, which demand ultra-reliable and low-latency communication. Gradient-based optimization frameworks have been tailored to accommodate such

stringent timing constraints, guiding resource allocation and scheduling strategies to meet performance goals.

Despite these advances, scalability challenges persist, particularly for very large and heterogeneous networks with complex constraints. The computational overhead and intricate problem landscapes limit the direct application of classical gradient methods. To address this, hybrid metaheuristic approaches have been proposed, combining gradient information with heuristic exploration to improve solution quality and scalability [9]. Benchmark studies indicate that while gradient-driven frameworks perform well on moderately sized networks, augmenting them with metaheuristics such as Variable Neighborhood Search (VNS) or population-based heuristics significantly enhances their ability to explore large combinatorial spaces effectively.

For instance, VNS algorithms applied to complex hub location problems that involve competitive pricing and demand shifts demonstrate robust and scalable performance [9]. These algorithms employ multiple neighborhood structures—including hub swapping, client reallocation, and price adjustments—coupled with shaking and local search procedures to escape local optima and explore diverse solutions thoroughly. Numerical experiments on benchmark instances show superior scalability and solution robustness compared to classical heuristics, even under varying price competition intensities.

Collectively, these findings emphasize the benefit of modular algorithmic strategies that adaptively integrate gradient-based optimization with heuristic methods. Such hybrid approaches carefully balance computational efficiency with solution robustness, making them well suited for the dynamic and large-scale demands of telecommunication network optimization. This integration paves the way for practical deployment in complex real-world systems where both precision and scalability are essential.

6.13.5 Neural Network-Based Information Transfer (NNIT). Addressing the dynamic and time-varying nature of network environments requires methods that not only optimize parameters in the current setting but also leverage learned knowledge from previous environments to accelerate adaptation. Neural Network-Based Information Transfer (NNIT) exemplifies this strategy by employing neural networks to learn mappings between evolving network states and their corresponding optimal or near-optimal solutions, thereby facilitating faster convergence and enhanced adaptability [4, 10, 23**?**].

Typically, NNIT integrates population-based evolutionary algorithms with neural networks trained to predict promising regions of the solution space or to transform historical high-quality solutions into effective candidates for the current environment [**?**]. By learning the structural patterns inherent in dynamic environments, such methods substantially reduce the computational overhead associated with repeated optimization from scratch. For example, in supply chain networks—characterized by complexity and dynamic features akin to telecommunications systems—NNIT techniques leverage variational inequalities and Lagrange multiplier analysis to model supply chain equilibria, providing numerical insights for informed solution transfer [23]. This framework captures equilibrium behaviors in which firms optimize profits subject to fixed labor and wage bounds, while sensitivity analysis via Lagrange multipliers guides strategic decisions on resource allocation and wage policy, thus enhancing robustness and computational efficiency in dynamic optimization tasks.

Recent advances in interpretable AI have been integrated into NNIT architectures to improve transparency and trustworthiness by revealing insights into the solution landscape [10]. These methods utilize optimal decision trees with hyperplanes (OCT-Hs) to approximate complex nonlinear and black-box constraints and objectives within a mixed-integer optimization framework. This results in global approximate models that transparently characterize feasible regions and objective surfaces, enabling decision-makers to explore and validate solutions effectively—an essential feature for deployment in safety-critical and high-stakes network control scenarios.

NNIT also demonstrates considerable promise in nonlinear stochastic decentralized adaptive control applications, illustrating its versatility across a broad spectrum of network optimization problems [4]. The underlying control-theoretic framework formulates optimal interventions to mitigate the propagation of economic shocks in networked systems. By optimally allocating control resources under budgetary and delay constraints, this approach effectively reduces shock severity and economic loss, highlighting scalability and strategic efficacy. Such control principles complement NNIT's hybrid paradigm by incorporating system dynamics and resource constraints that enhance both solution quality and practical applicability.

In summary, NNIT represents a robust hybrid paradigm that synergistically combines population-based optimization, learned knowledge transfer via neural networks, interpretable modeling through optimal decision trees, and control-theoretic principles. This integration effectively addresses the complexity, scalability, and dynamism inherent in modern and future network control problems. By exploiting these complementary strengths, NNIT-based algorithmic designs achieve superior optimization performance and adaptive capacity across diverse dynamic and uncertain environments.

7 AI-Enhanced Wireless Networking and Sensing

This section explores how artificial intelligence (AI) techniques improve wireless networking and sensing, with a particular focus on their integration with reconfigurable intelligent surfaces (RIS). We begin by outlining the main objectives, followed by subdivided discussion on methodologies, benefits, security implications, challenges, and future directions. A summary table is also provided to aid clarity and retention.

7.1 Reconfigurable Intelligent Surfaces and AI Integration

Reconfigurable intelligent surfaces (RIS) are planar structures composed of numerous passive reflecting elements with adjustable electromagnetic properties, such as phase shifts, which can be programmatically controlled to dynamically manipulate wireless signal propagation. By smartly configuring these elements, RIS can enhance signal quality, extend coverage, and reduce interference in

wireless environments. This fundamental capability makes RIS a promising enabler for next-generation adaptive wireless networks.

AI integration with RIS aims to achieve dynamic environment control for enhanced communication reliability, efficiency, and adaptability in wireless networks. Key objectives include optimizing signal propagation, managing interference, and improving resource allocation in diverse and challenging scenarios.

7.2 AI Methodologies in RIS-Assisted Networks

AI algorithms in RIS-assisted networks primarily utilize reinforcement learning and deep learning to configure surface elements such as phase shifts. These configurations enhance channel state prediction, link reliability, and spectral efficiency under variable conditions.

In addition, AI-driven interference management frameworks leverage real-time data analytics to detect and mitigate interference, thereby improving network scheduling and resource utilization. This is critical for enabling the coexistence of multiple users and technologies without significant performance loss.

7.3 Performance Improvements and Practical Challenges

Recent benchmarking results demonstrate that AI-powered RIS systems achieve notable enhancements in signal-to-noise ratio (SNR), latency reduction, and energy efficiency, applicable to environments such as millimeter-wave communications and multi-user MIMO setups.

Integration with existing wireless standards remains an essential challenge, as RIS and AI techniques require compatibility with protocols such as 5G NR and emerging 6G frameworks. Moreover, large-scale deployment of RIS arrays necessitates scalable AI algorithms with low computational complexity and real-time adaptability to cope with dynamic channel conditions and diverse application requirements.

7.4 Security Implications in AI-Enhanced RIS Systems

The integration of AI with RIS introduces new security considerations. AI-driven control of RIS elements may be vulnerable to adversarial attacks aimed at misleading the learning algorithms or manipulating phase configurations, potentially leading to degraded performance or unauthorized access. Additionally, real-time data analytics used for interference management must ensure data integrity and privacy to prevent exploitation by malicious users.

Robust security frameworks must be developed to detect and mitigate these threats while preserving the system's adaptive capabilities. This includes secure AI model training, anomaly detection mechanisms, and resilient control protocols compatible with RIS hardware constraints.

7.5 Summary of Key Points

AI Methodologies: Reinforcement learning and deep learning are central to dynamically configuring RIS elements and improving channel prediction and resource allocation.

Benefits: Enhanced signal quality, extended coverage, reduced interference, improved spectral efficiency, and energy savings.

Challenges: Scalability to large RIS arrays, real-time AI processing, hardware integration, and seamless compatibility with evolving wireless standards.

Security: Emerging risks include adversarial attacks targeting AI control and potential data breaches during real-time analytics; robust defense mechanisms are essential for reliable deployment.

In summary, AI-driven enhancements in RIS-based wireless sensing and networking empower adaptive and context-aware systems that meet stringent requirements of next-generation applications. However, challenges remain in scaling AI techniques for large-scale deployments, ensuring robustness, integrating AI seamlessly with evolving wireless standards, and addressing new security vulnerabilities. Addressing these will be critical for realizing the full potential of AI-enhanced wireless environments.

7.6 Reconfigurable Intelligent Surfaces (RIS)

Reconfigurable Intelligent Surfaces (RIS) have emerged as a transformative technology that enables programmable manipulation of wireless propagation environments, marking a shift from treating the environment as a stochastic, uncontrollable factor to one subject to deterministic control. This is realized through engineered metasurfaces capable of dynamically altering incident electromagnetic waves, providing unprecedented flexibility in wireless communications. The integration of Artificial Intelligence (AI), particularly machine learning techniques, significantly enhances RIS capabilities by optimizing complex, high-dimensional configuration spaces adaptively. Supervised learning methods assist in channel estimation by mapping measured channel state information (CSI) to optimal RIS configurations, while unsupervised learning facilitates feature extraction from unlabeled channel data, thereby improving generalization in dynamic and uncertain environments. Furthermore, deep reinforcement learning (DRL) provides an effective framework for sequential decision-making under uncertainty, enabling adaptive beamforming and resource allocation strategies that optimize spectral efficiency and energy savings [6]. Experimental results demonstrate that AI-enabled RIS surpass traditional heuristic approaches by adapting efficiently to dynamic scenarios and coping with imperfect channel information, thus improving coverage and robustness. The combination of these AI paradigms empowers RIS to address key challenges such as high-dimensional configuration spaces, latency constraints, scalability issues, and security concerns, resulting in more robust and efficient wireless links with enhanced coverage and energy efficiency. Future research directions prioritize the development of lightweight, distributed AI algorithms specifically tailored for RIS, the adoption of federated learning techniques to preserve user privacy, and the seamless integration of RIS with emerging technologies such as millimeter wave (mmWave), massive MIMO, and edge computing. Collectively, these advancements are poised to accelerate the realization of intelligent wireless environments that substantially improve network performance and sustainability [6].

Table 20: Summary of AI Techniques and RIS Benefits in Wireless Networking and Sensing

Aspect	AI Methodologies	Benefits	Challenges and Future Directions
RIS Configuration	Reinforcement Learning, Deep Learning	Adaptive phase shift optimization, improved channel state prediction	Scalability in large RIS arrays, real-time learning efficiency
Interference Management	Real-time Data Analytics, Machine Learning	Enhanced interference identification, resource allocation, network scheduling	Robustness to dynamic environments, multi-user coexistence complexity
Performance Gains	AI-assisted RIS control	Increased SNR (up to 10-15 dB), reduced latency (up to 30%), improved energy efficiency	Integration with diverse wireless standards (5G/6G), hardware limitations
Application Scenarios	Millimeter-wave, multi-user MIMO systems	Context-aware adaptivity, improved spectral efficiency	Deployment cost, reliability under harsh propagation conditions
Security Considerations	Secure AI training, anomaly detection	Protection against adversarial attacks, data integrity	Development of resilient AI and control protocols, privacy preservation

7.7 Benefits and Challenges of RIS

The AI-enabled RIS paradigm offers multiple benefits that significantly enhance wireless communication systems. These include notable improvements in spectral efficiency achieved through enhanced directivity and more effective interference management. For instance, experimental studies have shown that AI-enabled RIS can adaptively optimize beamforming strategies to outperform traditional heuristic methods, leading to measurable gains in coverage and data rates [6]. Additionally, RIS reduces reliance on active radio frequency components, thereby augmenting energy efficiency and contributing to greener communications. It also extends coverage by enabling signal reflection and focusing beyond line-of-sight barriers, which facilitates connectivity in dense urban environments or scenarios with significant obstacles. An important advantage of AI integration is robustness under imperfect channel conditions, as AI algorithms can learn and compensate for noise and fading effects, thereby maintaining high communication quality [6].

However, these benefits come with inherent challenges that must be addressed to realize practical deployment. The RIS configuration space is typically high-dimensional—sometimes involving thousands of elements—rendering exhaustive search or conventional heuristic optimization methods impractical. This complexity necessitates the development of scalable AI algorithms capable of effective dimensionality reduction while preserving performance. For example, deep reinforcement learning and unsupervised learning approaches have been proposed to efficiently map channel states to near-optimal RIS configurations [6]. Moreover, practical deployment requires AI techniques that meet stringent latency and scalability constraints, motivating the use of lightweight and distributed learning methods such as federated learning, which also offers privacy benefits by avoiding centralized data aggregation.

Security is another critical concern, as adversaries might exploit RIS to perform unauthorized eavesdropping or signal manipulation. For instance, attackers could manipulate RIS configurations to create signal reflections that intercept confidential transmissions or degrade system performance. To mitigate these risks, secure AI-driven configuration protocols have been proposed which incorporate authentication and encryption mechanisms during RIS control signaling. Real-time anomaly detection mechanisms leveraging AI can identify unusual RIS behavior or unexpected channel variations indicative of attacks, enabling prompt countermeasures. Techniques such as adversarial training and robust learning can further harden the AI models against manipulation [6]. These case studies highlight the active research on security solutions tailored for AI-integrated RIS, which remain essential for ensuring trustworthy and resilient deployments.

Balancing and addressing these benefits and challenges is central to advancing RIS deployment and realizing intelligent wireless environments.

7.8 Future Prospects in Wireless AI

Looking ahead, lightweight distributed AI architectures are poised to enable real-time and energy-efficient control of RIS, seamlessly integrated with pervasive wireless networks. Federated learning emerges as a key methodology, facilitating decentralized training of RIS optimization models across edge nodes while safeguarding data privacy and reducing communication overhead. This approach is especially vital given the growing heterogeneity of network topologies and the non-independent and identically distributed (non-i.i.d.) nature of data across devices. Federated learning allows each edge device to collaboratively train a shared global model without exchanging raw data, thereby preserving user privacy and mitigating bandwidth constraints. The convergence of federated AI with cutting-edge physical-layer technologies—such as millimeter-wave (mmWave) communications, massive multiple-input multiple-output (MIMO) antenna arrays, and edge computing platforms—will drive the advancement of edge intelligence [6]. These integrated frameworks are expected to jointly optimize sensing, communication, and computation resources under strict latency and energy constraints.

Achieving these objectives requires future algorithmic innovations that carefully balance trade-offs among model complexity, convergence speed, and robustness to channel estimation errors. For example, effectively training RIS controllers in highly dynamic wireless channels demands algorithms that quickly adapt to fast-varying conditions while maintaining stability under imperfect feedback. Moreover, emerging paradigms like neuromorphic computing and online continual learning offer promising routes to enhance system adaptability and resilience in highly dynamic wireless environments. Neuromorphic computing mimics the brain's neural architecture to provide low-power, event-driven computation, enabling RIS controllers to operate efficiently with sparse and temporal data. Online continual learning allows RIS optimization models to continually update from streaming data, addressing non-stationary environments without catastrophic forgetting. These paradigms can be particularly advantageous in scenarios such as mobile vehicular networks or smart factories, where channel conditions and network states evolve rapidly.

Overall, addressing these challenges will require interdisciplinary research combining insights from AI, wireless communications, and hardware design to realize intelligent wireless environments that are both scalable and robust.

7.9 Intelligent Interference Management in Perceptive Mobile Networks (PMNs)

Perceptive Mobile Networks (PMNs) represent an advanced integration of communication and sensing functionalities, enabling wireless infrastructure to simultaneously support data transmission and

situational awareness. Effective interference management is essential because sensing waveforms and communication signals coexist in shared spectral and spatial domains, leading to complex coexistence challenges. Recent advances have introduced AI-empowered interference mitigation frameworks that exploit macro-diversity gains and coordinated beamforming strategies across multi-cell architectures [18]. In particular, deep learning-based interference prediction models leverage both historical and real-time channel observations to accurately forecast interference patterns. This predictive capability enables proactive and dynamic resource allocation, maximizing the sensing signal-to-interference-plus-noise ratio (SINR) while preserving the quality of communication links. These dynamic allocation schemes improve sensing detection probability by approximately 20% and simultaneously reduce intra- and inter-cell interference by about 30%, thereby achieving a balanced optimization of communication and sensing objectives in PMNs [18].

Despite these promising developments, several critical challenges remain for practical and scalable PMN deployments. Key challenges include delivering low-latency inference to enable real-time system adaptation, acquiring precise channel state information in highly mobile and dynamic environments, and designing scalable cooperation mechanisms among multiple base stations that avoid excessive signaling overhead [18]. Addressing these issues is crucial for the realization of robust PMNs capable of harmonizing sensing and communication functionalities effectively. Emerging research directions emphasize integrating multi-modal sensing capabilities to enhance situational awareness, adopting federated learning frameworks for privacy-preserving and distributed interference management, and developing robust algorithms that maintain performance under heterogeneous and time-varying network conditions [18]. Collectively, AI-driven interference management frameworks promise substantial improvements in sensing accuracy and reliability while maintaining consistent communication quality within integrated sensing and communication networks.

7.10 Achievements and Challenges

AI-driven wireless sensing techniques have demonstrably enhanced detection probabilities and mitigated sensing interference, particularly through cooperative interference management strategies leveraging coordinated multipoint processing and macro-diversity [12, 18, 19]. These methods dynamically optimize resource allocation and beamforming configurations by integrating AI with physical-layer innovations, resulting in robust detection performance within dense and heterogeneous network environments marked by significant interference and channel uncertainty. Moreover, privacy preservation has emerged as a critical concern in networked sensing, as sensitive environmental or user data may be indirectly inferred via side-channel attacks inherent to cooperative frameworks. To address this, recent studies propose privacy-aware AI algorithms that incorporate differential privacy mechanisms and federated learning paradigms, effectively mitigating privacy risks while maintaining sensing performance [15, 18].

Robustness to heterogeneity in hardware capabilities, channel conditions, and user mobility patterns remains a major challenge. Approaches leveraging model adaptation, transfer learning, and fast adaptation methods such as Zero-Shot Lagrangian Updates have demonstrated potential in coping with such variability [8**?**]. For instance, Zero-Shot Lagrangian Update offers computationally efficient network optimization by directly updating Lagrange multipliers without iterative primal solves, enabling real-time adaptability in dynamic wireless environments [8]. Additionally, explainability-focused AI methods enhance trustworthiness and interpretability within heterogeneous setups [**?**]. Nonetheless, these techniques demand extensive validation in diverse real-world scenarios to ensure broad applicability and reliability. Bridging the gap between theoretical AI frameworks and practical deployment therefore necessitates continued research on scalable cooperation protocols, secure architectures, and self-tuning training paradigms capable of operating reliably across heterogeneous wireless environments.

In summary, AI-enhanced wireless networking and sensing via RIS and intelligent interference management herald the advent of programmable, efficient, and context-aware wireless systems. This progress hinges on the intricate interplay of algorithmic sophistication—encompassing supervised, unsupervised, reinforcement, and federated learning—and physical-layer innovations [6]. Collectively, these advances establish a rich interdisciplinary frontier poised to shape future wireless ecosystems [6, 8, 12, 15, 18**?** , 19].

Looking ahead, the continued evolution of AI techniques integrated with wireless sensing and networking promises transformative capabilities such as real-time adaptive networks that are both resilient and privacy-aware. Key future challenges include ensuring the scalability of AI algorithms under stringent latency requirements, achieving robust performance amid dynamically changing environments, and preserving user privacy through advanced federated and differential privacy approaches. Addressing these challenges will require collaborative efforts across machine learning, signal processing, and wireless communications disciplines, aiming to realize intelligent wireless systems that operate seamlessly in complex, heterogeneous settings. Ultimately, such systems will underpin next-generation wireless ecosystems that provide enhanced situational awareness, efficient resource utilization, and secure, trustworthy communication services.

8 Explainability, Interpretability, and Trust in AI-Controlled Telecommunication Systems

Explainability and interpretability are essential to building trust in AI-controlled telecommunication systems. They enable stakeholders to understand, verify, and rely on automated decisions. In operational telecom environments, explainability allows engineers and regulators to trace the reasoning behind AI actions, enhancing system reliability and ensuring regulatory compliance.

In summary, improving explainability and interpretability directly supports the trustworthiness of AI in telecom by making AI decisions transparent and verifiable.

8.1 Explainability Frameworks in Network Management

In multi-agent reinforcement learning-based network management, explainability frameworks such as attention mechanisms—which

highlight important inputs influencing decisions—and feature attribution methods—which quantify the impact of individual features—have demonstrated how agents coordinate resource allocation. These methods enable network operators to better understand and validate the decision-making processes of agents, facilitating the detection of anomalies and improving trust in automated systems. By providing interpretable feedback on agent cooperation patterns, these frameworks contribute directly to enhancing service quality, robustness, and operational transparency in complex network environments.

8.2 Regulatory Compliance and Transparency

Explainability methods significantly enhance regulatory compliance in telecommunications by ensuring that AI model decisions adhere to governance standards such as the General Data Protection Regulation (GDPR) and the European Union's AI Act. By providing transparent and interpretable AI decision-making processes, these methods facilitate comprehensive auditing for fairness, data privacy, and accountability. This transparency is crucial not only for satisfying external regulatory bodies but also for supporting internal governance, risk management, and compliance controls within telecom organizations.

8.3 Practical Applications in Telecom Operations

Practical case studies from telecom operators demonstrate the effectiveness of interpretable AI techniques in optimizing traffic routing while maintaining transparent decision logs that comply with internal audit requirements. For instance, an explainable resource scheduling model enabled operators to uncover unexpected decision factors and adjust system parameters accordingly, resulting in improved network throughput without compromising explainability. Additionally, interpretable AI has been applied in fault diagnosis systems where models highlight relevant system states and events, guiding operators through root cause analysis. This approach accelerates fault remediation and helps minimize downtime, thereby enhancing operational efficiency and service reliability.

8.4 Summary and Impact

In summary, explainability in AI-controlled telecom systems plays a crucial role in bridging the gap between complex algorithmic decisions and stakeholder understanding. Explainable AI techniques not only enable transparent decision-making processes but also foster stakeholder trust by making the behavior of AI systems interpretable and comprehensible. This transparency supports the trustworthy deployment of AI solutions that comply with both technical performance standards and regulatory requirements. Ultimately, these methods ensure that AI systems are reliable, auditable, and aligned with the operational goals of telecom networks, thereby enhancing their acceptance and facilitating effective management of increasingly autonomous and complex systems.

8.5 Importance of Transparent AI Decision-Making

The incorporation of artificial intelligence (AI) in adaptive telecommunication and control systems introduces an unprecedented level of complexity, making transparent decision-making an essential attribute to cultivate trust among stakeholders and ensure compliance with evolving regulatory frameworks. Transparency serves as a cornerstone for certifying that AI-driven actions conform to desired operational, ethical, and legal standards, particularly within critical infrastructure sectors such as telecommunications [**? ?**]. The establishment of trust is inherently linked to the system's ability to provide interpretable rationales behind its decisions, thus enabling operators to verify, audit, and justify automated processes [**?**]. This transparency is crucial in highly dynamic and heterogeneous network environments, where AI models must continually adapt to varying contextual conditions without compromising reliability and safety [18]. Additionally, emerging AI governance regulations emphasize explainability as a fundamental principle, compelling telecommunication systems to exhibit clarity in their decision logic and mitigate risks associated with opaque AI behavior [24]. As AI-driven network management confronts challenges such as balancing computational complexity with real-time processing requirements and ensuring robustness against adversarial conditions, transparent models help address these challenges by revealing decision pathways, underlying assumptions, and confidence levels. Consequently, transparent AI not only bolsters user confidence but also facilitates regulatory approvals and promotes the widespread adoption of AI-enhanced telecommunication technologies. This alignment with regulatory expectations and operational transparency is essential for advancing next-generation networks, including beyond 5G (B5G) and 6G systems, which fundamentally rely on AI for adaptive, secure, and efficient management [24**? ?**].

8.6 Methods for Interpretability and Explainability

Attaining interpretability within AI-driven telecommunication systems necessitates the integration of explainability mechanisms directly into core optimization and learning frameworks. Reinforcement learning (RL), a dominant paradigm for dynamic resource allocation and control, often poses challenges to transparency due to the complexity inherent in value function approximations and policy networks. Modern methodologies address this opacity through model-agnostic interpretability techniques and surrogate models that extract actionable insights from trained RL agents. These approaches clarify decision rationales by illuminating factors such as state-action value contributions and reward attributions [2**?**].

Embedding explainability frameworks within optimization algorithms further enhances understanding by elucidating solution trajectories and facilitating sensitivity analyses, enabling operators to comprehend how variations in system parameters influence resource management outcomes [20**?**]. For example, optimization-based resource allocation methods can be analyzed to reveal the impacts of system parameters on convergence and performance metrics, assisting network designers in balancing trade-offs between efficiency and fairness.

Hybrid frameworks that couple deep learning with symbolic reasoning have been advanced to strike a balance between predictive performance and interpretability, thereby supporting effective human-in-the-loop validation [32]. Such approaches are particularly relevant in multi-band wireless communication networks, which leverage heterogeneous frequency bands that vary in coverage and bandwidth capabilities. These networks face challenges related to complex channel dynamics, hardware limitations, and inter-band interference. By integrating expert knowledge through symbolic components, these hybrid models enhance explainability while maintaining accuracy in resource allocation and adapting to diverse frequency bands and propagation properties.

Moreover, attention mechanisms and gradient-based attribution techniques embedded in neural network architectures highlight key features that influence AI decisions across tasks such as traffic management, fault diagnosis, and spectrum allocation [**?**]. These methods provide insight into which input features or network conditions critically affect model outputs, thereby facilitating operational transparency in sophisticated AI-driven systems, including optical wireless communications that combine multiple access schemes and system architectures to achieve ultra-high capacity and interference resilience.

Collectively, these interpretability and explainability methods form a comprehensive toolset that mitigates the inherent opaqueness of sophisticated AI models, enhancing operational transparency while preserving system efficacy. This balance is crucial for advancing trustworthy AI applications in next-generation wireless networks, enabling robust, adaptive, and understandable decision-making under dynamic and heterogeneous environments.

8.7 Future Directions

This subsection clearly states the objectives for advancing explainable AI (XAI) in telecommunications and control systems, aiming to reconcile transparency, privacy, security, and scalability challenges in next-generation networks. The principal goals include developing privacy-preserving interpretability methods, enhancing robustness against adversarial threats, enabling scalable hierarchical explanations across complex system layers, and integrating multi-agent and LLM-driven AI paradigms with practical deployment in edge and cloud environments.

Specifically, the future roadmap emphasizes privacy-preserving XAI approaches that uphold data confidentiality while delivering faithful explanations, a crucial requirement in telecommunication networks where privacy regulations and operational constraints are stringent [18]. For example, federated explainability allows interpretability without centralizing sensitive information, demonstrated effectively in multi-cell perceptive mobile networks using coordinated deep learning for interference mitigation [18]. This paradigm supports privacy-compliant resource allocation, advancing real-world deployment in distributed network scenarios.

Another important focus is adversarial robustness of interpretability frameworks, essential for defending AI systems operating in hostile network environments vulnerable to exploitation [24]. Robust XAI must incorporate anomaly detection and adversarial training techniques to ensure explanations remain reliable despite malicious manipulations, as investigated in AI-driven network management and validated in recent Open RAN implementations [24, 25]. For instance, combining adversarially robust explanations with multi-agent learning enables detection and mitigation of attacks that may distort resource allocation or fault diagnosis, reinforcing network security.

Scalability challenges arise in interpreting large-scale and multi-layered communication and control architectures spanning cloud, edge, and devices, alongside multi-agent interactions [5]. Addressing this requires modular, hierarchical XAI frameworks capable of providing contextualized, multi-level explanations. Innovative methods are also needed to handle emerging paradigms such as multi-agent reinforcement learning and LLM-based agentic AI, which demand capturing complex cross-agent dynamics and delivering natural language explanations. The LLM-driven agentic AI approach in Open RAN exemplifies how autonomous fault identification and mitigation can be enriched with interpretable natural language outputs, enhancing human understanding [5].

Realizing these directions involves concrete implementation strategies, including federated learning for privacy preservation without sacrificing explanation fidelity, multi-agent coordination models for distributed intelligence, and lightweight XAI models optimized for low-latency, resource-constrained edge devices—an imperative in 6G and beyond Open RAN networks [25]. Modular and hierarchical explanation architectures facilitate managing the complexity of heterogeneous network layers and promote interoperability. Open-source efforts and benchmarking platforms focusing on privacy, robustness, and explainability are recommended to accelerate innovation and adoption.

Table 21 synthesizes these future directions by associating key challenges with potential technical solutions, providing a structured roadmap that interlinks privacy, security, scalability, and AI integration themes covered in this survey.

Key research questions driving progress in XAI for telecommunications include: How can federated XAI approaches guarantee stringent privacy protections without undermining explanation quality and user trust? What defense mechanisms most effectively counter adaptive adversarial attacks on explainability in fast-changing, real-time network analytics? How can hierarchical XAI frameworks be standardized to ensure seamless interoperability across diverse telecom infrastructure layers and vendors? Which optimization methods most successfully balance computational efficiency and interpretability in resource-constrained LLM-driven agentic AI deployments? Tackling these requires interdisciplinary collaboration spanning AI, control theory, wireless communications, and cybersecurity domains.

Anticipated disruptive innovations foresee agentic AI systems embedding LLM-driven agents that autonomously perform fault detection and network self-healing [5]. Empirical results demonstrate significant improvements such as a 17% increase in fault detection accuracy and a 40% reduction in network downtime compared to traditional approaches [5]. Seamless AI integration within Open RAN architectures supports adaptive, explainable control loops that bolster network resilience, throughput, and resource efficiency [25]. Collectively, these advancements herald transparent, robust, and efficient AI-driven telecommunications, paving the way for intelligent, self-optimizing networks that dynamically adapt to evolving environmental and operational conditions.

Table 21: Summary of Future Directions, Challenges, and Solutions in Explainable AI for Telecommunications and Control

Future Direction	Challenges	Potential Solutions
Privacy-preserving XAI	Balancing data confidentiality with maintaining explanation fidelity and user trust	Federated explainability, decentralized interpretation frameworks supporting privacy and compliance [18]
Adversarially robust explanations	Vulnerability to attacks undermining network security and explanation integrity	Integration of anomaly detection, adversarial training, and robust model design to secure explanations [24, 25]
Scalable, hierarchical interpretability	Managing complexity and contextualization across multi-layer, distributed, and multi-agent architectures	Modular, hierarchical explanation models that provide multi-level context and scalability [5]
Multi-agent and LLM-driven XAI	Complexity in cross-agent interactions and computational demands on constrained edge devices	Agentic AI frameworks with lightweight LLM optimization and multi-agent coordination mechanisms [5]
Real-time edge deployment	Meeting stringent latency and resource constraints for timely interpretable outputs	Lightweight XAI models optimized for edge devices, supported by hardware acceleration techniques [25]
Interdisciplinary synergy	Bridging disciplinary gaps among AI, control theory, and wireless communications	Modular and layered frameworks integrating theoretical foundations with practical implementation challenges

8.8 Applications in Telecommunications and Networking

This section provides a comprehensive overview of the key applications of artificial intelligence (AI) in telecommunications and networking. We detail the primary objectives, scope, and challenges specific to these domains, systematically examining how AI techniques enhance network performance, reliability, and security. The discussion focuses on four major application areas: network optimization, traffic prediction and management, resource allocation, and fault detection and diagnosis. Each area leverages AI methodologies to address specific problems: network optimization improves routing and load balancing to increase throughput and reduce latency; traffic prediction and management utilize machine learning models to forecast network demand and mitigate congestion; resource allocation applies AI for dynamic and efficient distribution of bandwidth and computing power; and fault detection and diagnosis employ intelligent algorithms to monitor, detect, and promptly resolve network failures, thereby enhancing system robustness.

Compared to traditional methods that often rely on static rules or manual configurations, AI approaches offer adaptive and data-driven capabilities that allow networks to respond dynamically to changing conditions. For example, conventional network optimization techniques might use fixed routing protocols that do not adapt well under varying traffic loads, whereas AI-driven optimization can learn traffic patterns and adjust routes in real time to improve efficiency. Similarly, traditional traffic prediction methods are generally limited by heuristic models, but AI models such as deep learning provide higher accuracy in forecasting network demand, enabling proactive congestion management.

Practical deployments have demonstrated these advantages: telecommunications operators increasingly employ AI-based resource allocation to optimize bandwidth use in 5G networks, achieving better service quality and fairness. In fault detection, AI algorithms have been successfully integrated into network management systems to enable early identification of anomalies, reducing downtime and maintenance costs. These examples underscore the practical impact of AI in advancing telecommunications and networking beyond the capabilities of classical approaches.

8.8.1 Network Optimization. AI techniques have been extensively applied for network optimization by learning from historical data to dynamically adjust configurations in complex and large-scale networking environments. This adaptive approach enhances throughput and reduces latency while efficiently managing the heterogeneous nature of modern networks. Despite these advantages, challenges remain in achieving scalable real-time processing and integrating diverse and often noisy data sources. Compared to traditional rule-based and statistical approaches, AI methods provide superior adaptability and scalability; however, they may introduce additional computational overhead and demand careful design and deployment strategies to balance performance improvements with resource constraints.

8.8.2 Traffic Prediction and Management. Traffic prediction models employ machine learning techniques to forecast network congestion, enabling proactive traffic management strategies that enhance overall service quality and network efficiency. These models must contend with challenges such as highly dynamic and nonlinear traffic patterns, as well as data sparsity and variability across different time scales and network segments. Although AI-based approaches, including deep learning and ensemble methods, typically outperform traditional statistical models by better capturing complex temporal and spatial dependencies, their effectiveness depends on the availability of large volumes of high-quality, representative traffic data. Additionally, model interpretability and computational efficiency remain important considerations for practical deployment. Commonly used evaluation metrics include prediction accuracy, root mean square error (RMSE), and mean absolute error (MAE), which collectively offer a robust assessment of model performance by measuring both overall correctness and magnitude of prediction errors across diverse traffic scenarios.

8.8.3 Resource Allocation. AI-driven resource allocation addresses complex constraint satisfaction challenges in network resource management by optimizing the allocation of limited resources to meet diverse and often conflicting user demands. These approaches aim to improve efficiency and fairness through dynamic balancing of objectives such as maximizing throughput, minimizing latency, and providing equitable service to users. Compared to traditional heuristic and rule-based methods, AI-based techniques can adapt more flexibly to changing network conditions and demonstrate superior performance in resource utilization and fairness metrics. However, key challenges persist, including the need to provide theoretical fairness guarantees, reduce the computational overhead of advanced AI models, and enhance interpretability to facilitate practical deployment. Additionally, future research should investigate the trade-offs among performance gains, model complexity, and transparency, ensuring that AI-driven resource allocation solutions are both effective and deployable in real-world network environments.

8.8.4 Fault Detection and Diagnosis. AI-based fault detection systems are designed to identify and diagnose anomalies within network operations, effectively handling imbalanced datasets and minimizing false alarms. Their ability to detect subtle and rare events

significantly enhances network resilience. However, balancing the detection rate and false positive rate remains a critical challenge, particularly in real-world environments where false alarms can disrupt service continuity. To comprehensively evaluate these systems, metrics such as detection rate, precision, recall, and false positive rate are indispensable.

While AI methods offer substantial benefits over traditional techniques in adaptability and performance, they also introduce challenges related to computational overhead, data requirements, model interpretability, and deployment complexity. Overcoming these limitations is vital to unlock AI's full transformative potential in telecommunications. Future research directions should emphasize scalability, real-time implementation, fairness, and explainability to ensure the development of robust and efficient AI-enabled network systems.

This structured overview highlights the interconnected nature of AI applications, underscoring their collective role in enhancing the robustness and intelligence of modern communication networks.

8.8.5 AI-Driven Adaptive Control Applications. This section provides a detailed overview of the objectives, key AI techniques, applications, and associated challenges in the implementation of AI-driven adaptive control within telecommunications networks. The explicit measurable goals of such AI-driven adaptive control include enhancing Quality of Service (QoS) metrics such as throughput, latency, and reliability; optimizing resource utilization to reduce operational expenditure (OPEX); improving fault resilience through early anomaly detection and remediation; and ensuring privacy and computational efficiency in distributed environments.

The integration of artificial intelligence (AI) techniques—especially machine learning (ML) and deep learning (DL)—has significantly reshaped adaptive control mechanisms in telecommunications. These methods enable critical functions like dynamic resource allocation, congestion control, fault tolerance, and traffic prediction. Specifically, deep learning architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs) have demonstrated outstanding ability to extract complex spatial-temporal features from network data, thereby enhancing predictive accuracy and control responsiveness [2, 7, 24, 35**? ? ? ?** , 36].

This advancement marks a paradigm shift from traditional static, heuristic-based methods toward data-driven, adaptive frameworks capable of learning from large-scale historical and real-time network states. Reinforcement learning (RL), for example, has been effectively applied to dynamic radio resource allocation, achieving optimized trade-offs between throughput and latency in heterogeneous wireless environments [36**?**]. Deep learning models like CNNs and RNNs outperform classical statistical predictors in traffic forecasting by leveraging nonlinear dependencies and temporal correlations within network flows [2**? ?**]. GANs facilitate synthetic data generation and anomaly detection, improving network fault management by detecting rare or previously unseen failure events and addressing challenges posed by sparse failure datasets [7**?**].

However, deploying complex AI models introduces critical challenges. The substantial computational overhead in training and inference of deep learning models affects real-time applicability, particularly in large-scale or resource-constrained edge environments [2**?**]. Federated learning (FL) offers a promising distributed training methodology that preserves data privacy and lowers communication overhead. Robust FL frameworks incorporate gradient sparsification, error feedback mechanisms, and adaptive client selection to mitigate client dropout, limit bandwidth consumption, and accommodate heterogeneous device capabilities [**?**]. Furthermore, joint optimization of model parameter sizing and bandwidth allocation has been shown to reduce FL training time by up to 30% and increase accuracy by 3-5%, according to wireless network simulations with diverse device and channel conditions [2].

Generalization remains a key concern, as AI models must adapt continuously to heterogeneous, evolving network scenarios, necessitating efficient retraining or adaptation strategies [**?**]. Privacy concerns arising from extensive network data collection motivate privacy-preserving frameworks, such as FL, maintaining data locality while collaboratively improving model performance [7, 35]. Despite challenges, AI-driven adaptive control has demonstrated substantial improvements in network self-optimization, reducing OPEX and enhancing QoS across several performance indicators [24**?**].

In summary, AI-driven adaptive control in telecommunications networks seeks to automate and dynamically optimize network operations through advanced ML and DL techniques. The explicit objectives include measurable improvements in throughput, latency, fault detection accuracy, and resource efficiency. While the benefits are considerable, overcoming computational constraints, ensuring privacy, and enabling robust generalization are critical for practical deployment and continued advancement of these adaptive control systems.

8.8.6 Evaluation Metrics and Benchmarking. Comprehensive evaluation of adaptive algorithms in realistic communication and wireless scenarios necessitates metrics that integrate both classical network performance and AI-specific qualities, including robustness and semantic fidelity. To improve clarity, we categorize the key evaluation metrics and benchmarking considerations as follows:

Classical Network Performance Metrics. Traditional benchmarks remain fundamental indicators of network health. These include throughput, latency, packet loss, and bit error rate (BER) [2**?**]. Latency models that encompass transmission, propagation, processing, queueing, and retransmission delays are especially critical in emerging applications such as the Tactile Internet, which demands ultra-low latency and high reliability [7].

Semantic Fidelity Metrics. With the emergence of semantic communications, new metrics quantifying semantic integrity have become essential. Semantic similarity metrics, such as BLEU scores applied to textual or annotated image data, offer a means to assess content fidelity following AI-enhanced compression and error correction [24, 36**?**]. These metrics align closely with frameworks that blend deep learning with knowledge graphs and knowledge bases, enhancing semantic context consistency and error correction capabilities [36]. Incorporating such semantic-level evaluations addresses limitations of purely physical-layer metrics by better reflecting end-user perceived quality.

Table 22: Summary of AI Applications, Challenges, and Evaluation Metrics in Telecommunications and Networking

Application Area	Challenges	Evaluation Metrics
Network Optimization	Scalability, real-time processing, heterogeneous data sources	Throughput, latency, resource utilization
Traffic Prediction and Management	Dynamic traffic patterns, data sparsity	Prediction accuracy, RMSE, MAE
Resource Allocation	Complex constraint satisfaction, fairness among users	Resource efficiency, fairness indices
Fault Detection and Diagnosis	Imbalanced data, anomaly identification, minimizing false alarms	Detection rate, false positive rate, precision, recall

Operational and AI-Specific Metrics. Evaluations also consider computational efficiency, convergence speed, and resilience to adversarial perturbations or network faults [2, 7**?**]. This ensures that adaptive algorithms can perform robustly in dynamic and potentially hostile environments. Metrics balancing accuracy with computational and latency constraints, alongside interpretability and privacy considerations, are crucial for AI-driven network systems [24**?**].

Benchmarking and Standardization. Despite these advances, the field lacks widely adopted standardized benchmarks tailored to semantic communications and AI-driven network control. Existing datasets such as MNIST and CIFAR-10 are commonly used for supervised and federated learning evaluations [2], but they inadequately capture the complexities of semantic fidelity, multimodal data integration, or dynamic, real-time adaptive networking. There are ongoing initiatives and calls within the community to develop novel datasets and repositories that better reflect real-world communication challenges, incorporating multimodal semantics and edge-cloud interactions [24].

Towards Standard Benchmark Protocols. To foster uniformity and comparability across studies, we propose best practices that include clearly defined metrics encompassing semantic, physical-layer, and operational performance. Complementing these metrics, benchmarking protocols should enforce consistent testbed configurations and promote the use of open-source simulation frameworks [24]. Additionally, standardized evaluation frameworks must carefully balance trade-offs among accuracy, latency, computational overhead, interpretability, and privacy in AI-enabled network systems. Establishing such comprehensive benchmarks is pivotal to enable rigorous, reproducible performance comparisons and accelerate practical deployments in wireless and edge-cloud environments.

In summary, advancing evaluation metrics and benchmarking requires structured metric taxonomies, standardized datasets reflective of semantic and AI-driven communication challenges, and community-wide adoption of uniform protocols to promote reproducibility and cross-comparison.

8.8.7 Edge and Cloud Synergistic AI Solutions. The exponential growth of network data and the imperative for ultra-low-latency services have precipitated architectures that synergistically combine edge computing with cloud intelligence. By partitioning AI workloads between decentralized edge nodes and centralized cloud platforms, such hybrid frameworks optimize latency constraints while leveraging substantial computational resources to enhance network intelligence and robustness [6, 20, 24, 38**? ?**].

At the edge, AI models perform real-time inference and handle local data processing, crucial for latency-sensitive applications such as the Tactile Internet and autonomous vehicle control [7]. To accommodate resource limitations inherent to edge devices, lightweight deep learning models or compressed representations are employed, enabling efficient on-device execution without compromising responsiveness [20**?**]. Meanwhile, cloud-based AI systems aggregate global network insights, manage intensive training processes, and distribute updated models back to edge nodes, supporting continual learning and adaptive decision-making [38**?**].

Federated learning exemplifies this edge-cloud synergy by facilitating decentralized model training across heterogeneous devices while preserving data privacy and reducing communication overhead [7]. Addressing device heterogeneity involves adaptive model compression and personalized federated optimization techniques that match model complexity to individual device capabilities. Synchronization challenges caused by diverse update frequencies and intermittent connectivity are mitigated through asynchronous federated learning protocols and hierarchical aggregation schemes. For example, grouping devices based on similarities in data distribution or computational power promotes more stable and efficient training rounds [7**?**].

Additionally, distributed AI systems confront security challenges including adversarial attacks and data poisoning, spurring research into robust model architectures and trustworthy deployment strategies at scale [6, 24]. The dynamic interplay between edge and cloud computing thus represents a critical frontier for delivering intelligent, responsive, and secure network control in next-generation wireless ecosystems, with ongoing developments guided by advancements in AI model design, communication protocols, and network architecture [6, 7, 24**?**].

Importantly, the design of these hybrid AI systems is closely linked to adaptive control principles, enabling real-time adjustments of AI model parameters and resource allocation in response to varying network conditions, traffic loads, and application demands [7, 24]. For instance, adaptive control strategies can dynamically balance the trade-offs between latency, accuracy, and energy consumption across the edge-cloud continuum. Evaluating the performance of such synergistic AI deployments requires comprehensive metrics that capture latency, inference accuracy, communication overhead, model convergence rates, and system robustness [7, 24**?**]. These metrics provide critical feedback loops for both control and optimization processes, fostering continual system adaptation aligned with stringent quality-of-service requirements intrinsic to emerging applications like the Tactile Internet.

To enhance accessibility, the underlying concept involves a two-tiered architecture: the edge tier acts as a near-user processor enabling low-latency, privacy-preserving AI tasks, while the cloud tier functions as a global coordinator and heavy-duty trainer [7**?**]. This separation supports scalable and secure AI services tailored

to next-generation networks' heterogeneity and complexity. The interdependence between adaptive control, evaluation metrics, and edge-cloud AI solutions forms the backbone of future wireless AI ecosystems, ensuring responsiveness, reliability, and intelligence at scale.

8.8.8 Resilient Control of Cyber-Physical Systems. Cyber-physical systems (CPS) underpinning telecommunications infrastructure require resilient control strategies capable of maintaining reliable operation despite actuator faults and sophisticated cyber attacks. A prominent approach involves neural network-based finite-time resilient control methodologies for nonlinear time-delay systems, utilizing radial basis function neural networks (RBFNNs), advanced observer designs, and Lyapunov–Krasovskii functionals to ensure robustness and rapid convergence [3].

This framework models the system's unknown nonlinearities and fault signals through RBFNNs. Specifically, the nonlinear dynamics $f(x(t), x(t-\tau))$ are approximated as

$$f(x(t), x(t-\tau)) \approx W^T \Phi(x(t), x(t-\tau)) + \varepsilon,$$

where W represents unknown weights, Φ denotes the known basis functions, and ε is the approximation error. The unknown weights W are estimated online using adaptive laws. These adaptive laws allow the controller to learn and compensate in real-time for uncertainties in the system, including time delays and external disturbances. Concurrently, a specially designed observer estimates both the system states and fault signals. This observer can handle discrepancies caused by unknown false data injection (FDI) attacks and measurement inaccuracies, which enhances the system's ability to detect and isolate faults dynamically [3].

In simpler terms, imagine a system whose behavior depends not only on its current state but also on past states due to delay, and where sensors and actuators may be compromised by faults or malicious attacks. The neural network approximates the unknown nonlinearities so the controller can predict system trends even when direct knowledge is unavailable. The observer then dynamically figures out whether sensor readings or actuator outputs are faulty or tampered, giving the controller accurate estimates to base decisions on.

The adaptive control laws merge these state and fault estimates to guarantee finite-time convergence of both system states and estimation errors. Unlike traditional control methods that only assure eventual stability, finite-time control ensures that errors disappear within a known, short interval. This characteristic is crucial in hostile or fault-prone environments, where rapid recovery is necessary to maintain safe and reliable operation. The use of Lyapunov–Krasovskii functionals, which incorporate the effects of time delays explicitly, provides mathematical rigor to these stability guarantees despite unknown nonlinearities and disturbances.

To illustrate, consider a nonlinear system experiencing unknown actuator faults and false data injection attacks that cause incorrect sensor data. The proposed control strategy uses RBFNNs to approximate nonlinear terms involving delayed states. The observer tracks both the true system states and any faults or attacks, enabling the controller to adjust inputs in real-time, quickly offsetting the impact of faults. Simulation studies on benchmark nonlinear systems confirm the approach's superior fault tolerance and faster stabilization relative to classical adaptive controllers [3].

In summary, this resilient control framework integrates adaptive neural network control with observer-based fault diagnosis and robust stability theory to achieve real-time mitigation of faults and cyber attacks through dynamic control adjustments. This approach forms a comprehensive resilience paradigm for CPS, addressing both component faults and malicious interventions. Remaining challenges include extending methods to systems with stochastic elements and time-varying delays, supporting multiple actuators and sensors, and validating effectiveness through hardware-in-the-loop experiments reflecting practical deployment scenarios [3].

Thus, this strategy offers a promising pathway to enhance the security and reliability of CPS by enabling real-time learning and estimation mechanisms that not only detect but actively counteract faults and cyber threats.

8.8.9 Open Research Frontiers. Emerging research trajectories in telecommunications underscore the significance of multi-agent systems, stochastic modeling, and hardware-in-the-loop simulation platforms as foundational frontiers advancing adaptive control and AI integration [3]. Multi-agent frameworks enable scalable, decentralized decision-making across heterogeneous network entities, thereby enhancing robustness and adaptability in complex, dynamic environments. However, challenges remain in ensuring coordination among agents under communication constraints and heterogeneous capabilities, which can degrade overall system performance and resilience. Metrics such as convergence speed, fault tolerance rate, and control accuracy serve as concrete milestones for evaluating progress in these systems.

The integration of stochastic models provides a nuanced characterization of wireless channel variability, environmental uncertainties, and human-in-the-loop behaviors, which necessitates adaptive control methodologies capable of balancing performance with risk management effectively [3]. A key challenge here involves managing the trade-off between model complexity and real-time applicability, as overly complex stochastic models may impede timely decision-making. Important metrics include the estimation error bounds, robustness margins against uncertainty, and computational latency.

Hardware-in-the-loop platforms serve as critical testbeds bridging the gap between algorithmic development and real-world hardware constraints such as timing delays, sensor inaccuracies, and communication limitations [3]. These platforms expedite prototyping, validation, and fine-tuning of resilient control schemes under conditions closely aligned with realistic network environments. Risks include insufficient fidelity in simulated hardware responses and scalability issues when extending to large networked systems. Validation benchmarks focusing on real-time responsiveness, fault injection coverage, and hardware-software integration fidelity are important for gauging platform effectiveness.

Complementary research areas focus on developing explainable AI paradigms in network control to ensure transparency and interpretability, as well as integrating reinforcement learning with classical control theory to merge long-term policy optimization with provable system stability guarantees [24]. Challenges in explainability stem from the inherent complexity of AI models and

the potential trade-offs between interpretability and performance. Reinforcement learning integration must address stability guarantees and sample efficiency in dynamic network environments. Key evaluation criteria include explainability scores, convergence rates, and stability certification under varying conditions.

Tackling computational complexity remains a pivotal concern with efforts directed towards model compression techniques, distributed AI frameworks, and energy-efficient algorithms, which are imperative for deploying intelligent control in resource-constrained edge environments [6]. Challenges include maintaining performance under hardware limitations and mitigating latency in distributed settings. Metrics such as energy consumption per inference, model size reduction ratio, and distributed computing overhead quantify advancement in this area.

Collectively, these research directions underscore a transformative potential for next-generation telecommunications systems characterized by intelligence, autonomy, and resilience. However, realizing this potential requires addressing fundamental challenges related to scalability, uncertainty, interpretability, and computational limitations.

To delineate a structured research roadmap, Table 23 summarizes near-term and long-term objectives with associated challenges and metrics.

Near-term objectives include designing and benchmarking adaptive multi-agent control strategies endowed with robust fault tolerance, leveraging hardware-in-the-loop platforms for realistic validation [3]. Concurrently, efforts aim to advance explainable AI methods and fuse reinforcement learning with classical control techniques to achieve interpretable and stable network management solutions [24]. Addressing computational challenges via lightweight distributed AI and energy-aware algorithms establishes a foundation for practical deployment in edge computing scenarios [6]. Milestones such as improved convergence guarantees, demonstrable explainability, and energy reduction targets guide these efforts.

Long-term research goals envision the seamless integration of distributed resilient control mechanisms for complex multi-agent networks operating under dynamic, uncertain wireless conditions. This involves extending stochastic adaptive control techniques to accommodate time-varying delays and enhance resilience against cyber-physical attacks, as demonstrated by frameworks incorporating adaptive neural network-based fault detection and mitigation [3]. Future investigations aspire to realize fully autonomous, intelligent wireless systems by harmonizing adaptive control, explainable AI, and scalable reinforcement learning frameworks while ensuring robustness, scalability, and sustainability within next-generation telecommunications infrastructures. Metrics for success include system-wide autonomy levels, robustness indices against attacks, and scalability efficiency.

9 Cross-Cutting Themes and Integration Considerations

This section synthesizes the key challenges, solutions, and interactions across the various themes discussed in preceding sections, highlighting their intersections through concrete examples and case studies. By examining these cross-cutting issues, we provide a cohesive understanding of how different AI components and methods integrate, addressing their combined limitations and opportunities.

9.1 Scalability and Interpretability Trade-offs

A primary challenge common across multiple themes is the trade-off between scalability and model interpretability. For example, large-scale transformer models achieve state-of-the-art performance on natural language processing tasks but often lack transparency in their decision-making processes. This issue intersects with ethical AI considerations where explainability is critical for user trust and accountability. In healthcare diagnostics, for instance, complex models must provide clear rationales for predictions to comply with regulatory standards and gain clinical acceptance. Addressing these trade-offs requires developing methods that maintain interpretability without sacrificing the ability to scale effectively.

9.2 Integrating Learning Paradigms

Another pervasive theme involves integrating learning paradigms, such as combining supervised learning with reinforcement learning to develop robust agents capable of adapting in dynamic environments. A relevant case study from autonomous driving illustrates this integration: perception modules trained on labeled images are fused with reinforcement learning policies that adapt to real-time driving conditions. This integration surfaces technical challenges including aligning heterogeneous learning objectives and managing uncertainty propagation. Implementation of unified frameworks in practice involves designing modular architectures that facilitate communication and coordination between components trained with different paradigms, ensuring consistency and minimizing cascading errors.

9.3 Data Privacy and Security

Data privacy and security are also essential cross-cutting concerns. Federated learning techniques, originally designed for privacy-preserving model training, have found applications ranging from mobile device personalization to collaborative healthcare research. The main challenge lies in balancing data utility with privacy guarantees and computational efficiency. For example, privacy-preserving federated methods often limit data sharing but can suffer from bias due to uneven client participation, necessitating algorithmic safeguards tailored to these contexts. Practically implementing unified frameworks requires incorporating privacy-preserving mechanisms seamlessly into modular systems, while retaining interpretability and scalability.

9.4 Towards Unified Frameworks: Practical Implementation

Despite increasing recognition of these interconnected challenges, most existing solutions address them in isolation. To foster AI systems capable of holistic performance, we propose a research roadmap focused on three priorities. First, developing unified conceptual frameworks that explicitly model the interactions among scalability, interpretability, privacy, and adaptability. Such frameworks would support joint optimization by capturing trade-offs and

Table 23: Research Roadmap: Open Frontiers in Intelligent Telecommunications Control

Research Frontier	Near-Term Objectives	Long-Term Objectives	Challenges & Metrics
Multi-Agent Systems	Develop adaptive, fault-tolerant control	Integrate distributed resilient control under dynamic, uncertain conditions	Communication constraints; Convergence speed; Fault tolerance rate; Control accuracy
Stochastic Modeling	Enhance adaptive methods balancing performance and risk	Extend to time-varying delays and cyber-physical attack resilience	Model complexity vs. real-time use; Estimation error bounds; Robustness margins; Latency
Hardware-in-the-Loop	Leverage real-world validation platforms	Scale to large network prototypes with hybrid simulation	Fidelity limitations; Scalability; Responsiveness; Fault injection coverage
Explainable AI	Advance interpretable network control	Achieve explainability-stability trade-offs in large networks	Model complexity; Explainability scores; Stability certification; Convergence
Computational Efficiency	Implement lightweight distributed AI and compression	Realize energy-efficient, scalable edge intelligence	Hardware limitations; Latency; Energy consumption; Model size; Overhead

synergies across these dimensions. Practically, this entails formalizing multi-objective optimization problems and designing algorithms and architectures that allow components to share necessary information without compromising privacy or interpretability.

Second, designing modular architectures that enable flexible integration of diverse learning paradigms and privacy-preserving techniques while preserving interpretability standards. These architectures should implement interfaces for communication between heterogeneous components, and embeds uncertainty quantification to control error propagation dynamically.

Third, establishing standardized benchmarks and evaluation protocols that assess multi-dimensional criteria—including transparency, robustness, and privacy—under realistic deployment conditions. This is essential for consistently comparing methods and understanding practical trade-offs.

9.5 Methodological Considerations

From a methodological perspective, integrated AI system design should incorporate explicit multi-objective optimization strategies to balance competing demands. For instance, optimization methods can simultaneously address model accuracy, explainability, privacy preservation, and computational cost. Embedding uncertainty quantification within modular pipelines can help control error propagation when composing heterogeneous components. Moreover, adaptive learning schemes that dynamically respond to environmental feedback can enhance system robustness and foster user trust.

9.6 Summary

In summary, addressing cross-cutting themes in AI requires moving beyond siloed approaches toward comprehensive system-level thinking. Emphasizing joint frameworks, modular design, and multi-criteria evaluation will better align AI capabilities with the complex requirements of practical deployment, fostering transparency, privacy, and adaptability in concert.

9.7 Scalability and Real-Time AI Inference

The deployment of artificial intelligence (AI) within telecommunications demands scalable solutions capable of real-time inference across heterogeneous and dynamically evolving network environments. This requirement is challenging due to the high computational complexity of contemporary AI models, such as deep neural networks and large language models (LLMs), coupled with the variability of network resources and stringent latency constraints [6, 13, 39**?**]. For instance, incorporating AI into Open RAN architectures mandates efficient processing pipelines that adhere to tight timing budgets, enabling rapid control loop adaptations critical for tasks like spectrum management and interference mitigation [18].

Edge-cloud collaborative frameworks offer distinct advantages by distributing AI inference workloads to optimize latency and computational resource utilization; however, they face scalability constraints especially when coordinating multiple edge nodes or base stations [6]. To address these challenges, scalable AI architectures must combine algorithmic compression methods, such as model pruning and quantization, with hardware acceleration and modular, parallel system designs. Synchronization and consistent model updating—particularly in federated or distributed learning paradigms—are critical for maintaining real-time performance while preserving data privacy and handling heterogeneous network conditions [13].

These complexities are especially pronounced in perceptive mobile networks (PMNs), where AI-driven interference management and sensing require dynamic adaptation to fluctuating network loads without compromising communication quality [39]. Advanced AI frameworks in PMNs leverage coordinated beamforming and deep learning-based interference prediction across multi-cell architectures to maximize sensing signal-to-interference-plus-noise ratio (SINR) while preserving communication reliability [18]. Cooperative sensing among multiple base stations enhances robustness under high network loads, demonstrating up to 20% improvement in detection probability and significant interference reduction.

Moreover, real-world AI-enabled routing and traffic management systems dynamically adapt to network conditions, achieving improvements of up to 30% in both throughput and latency [39]. Similarly, AI-powered software-defined networking (SDN) frameworks deploy machine learning models—including Random Forest classifiers and LSTM networks—for real-time traffic classification, anomaly detection, and resource allocation; experiments show up to 92% classification accuracy alongside an 18% reduction in end-to-end latency in emulated 5G scenarios [13]. These results highlight the practical scalability and real-time responsiveness achievable through integrated AI-SDN architectures.

Furthermore, the integration of AI with Reconfigurable Intelligent Surfaces (RIS) optimizes wireless environment control to improve spectral and energy efficiency. AI techniques, such as deep reinforcement learning, enable dynamic RIS configuration and resource allocation by learning mappings from channel states to optimal configurations, thus addressing high-dimensional optimization challenges and latency constraints in scalable deployments [6].

Overall, realizing scalable and real-time AI inference in telecommunications necessitates lightweight yet robust AI models optimized for dynamic network environments, efficient edge-cloud collaboration, and real-time synchronization mechanisms that support

distributed learning while preserving privacy and system responsiveness. These case studies underscore the critical importance of algorithmic compression, distributed inference, adaptive learning, and hardware-aware designs in meeting scalability demands without sacrificing real-time performance, thereby unlocking AI's full potential for next-generation network applications.

9.8 Privacy Preservation Strategies

Preserving privacy is a critical concern in telecommunications due to the sensitive nature of transmitted data and strict regulatory frameworks. Prominent strategies for privacy preservation include federated learning, edge computing, and lightweight distributed AI methods that localize data processing, thereby reducing exposure risks [6, 13, 18, 24]. Federated learning enables collaborative model training across decentralized entities while keeping raw data on-site, effectively mitigating risks inherent in centralized data collection [6]. Despite its advantages, federated learning faces challenges such as communication overhead, the heterogeneity of client devices, and susceptibility to inference attacks. Edge computing complements these approaches by performing AI inference near data sources, which reduces both the privacy attack surface and communication latency [13]. Deploying lightweight AI models at the edge—through techniques like quantization and knowledge distillation—further enhances privacy protection while improving computational efficiency [24].

Moreover, integration of encryption techniques plays a vital role in securing data transmissions and model parameters within telecommunications networks. End-to-end encryption ensures that data remains confidential during communication between distributed entities, effectively preventing unauthorized access and eavesdropping. Homomorphic encryption allows computations to be performed directly on encrypted data without decryption, enabling privacy-preserving model training and inference [24]. Combined with differential privacy mechanisms, which add calibrated noise to shared model updates, these approaches rigorously protect individual data contributions from leakage while maintaining overall model utility. In Software-Defined Networking (SDN)-enabled 5G and beyond frameworks, these encryption and privacy-preserving algorithms are embedded alongside AI models within SDN controllers, facilitating real-time traffic classification and anomaly detection without compromising data confidentiality [13]. However, designing algorithms that balance encryption overhead, privacy guarantees, and model accuracy remains complex, particularly under stringent latency and computational constraints.

Additionally, the rapid evolution of AI within open, multi-vendor ecosystems accentuates the need for standardized frameworks that embed stringent privacy safeguards while maintaining interoperability across diverse network infrastructures [13, 18]. Such frameworks integrate encryption, secure multiparty computation, and federated learning protocols, enabling seamless and privacy-respecting collaboration among heterogeneous stakeholders in telecommunications systems.

9.9 Explainability and Trust

This subsection aims to elucidate the critical role of explainability in fostering trust and transparency within AI-driven telecommunications systems, emphasizing current advancements, challenges, and their operational and regulatory implications.

Establishing trust and transparency in AI-driven telecommunications systems is essential, given that automated decisions directly affect service quality and network reliability [18, 24**?** , 25]. Explainability techniques empower operators and stakeholders to interpret AI decision-making processes, identify erroneous outputs, and align these decisions with domain expertise. For instance, incorporating explainable AI within network management systems elucidates the rationale behind resource allocation or anomaly detection outcomes, thereby fostering confidence and enabling effective human-in-the-loop oversight [**?**]. Empirical studies demonstrate that AI-empowered frameworks can meaningfully improve network performance while maintaining system transparency. For example, the interference management framework in perceptive mobile networks applies AI to dynamically allocate resources based on explainable interference predictions, resulting in a 20% improvement in detection probability and a 30% reduction in sensing interference while preserving communication quality [18]. Such explainable insights assist operators in validating AI decisions against operational constraints, reducing inadvertent network disruptions.

Common explainability approaches in telecommunications include attention visualization, which highlights which input features or network parameters most influence AI predictions, and feature attribution methods that quantify the contribution of individual variables to final decisions [24]. For example, attention maps in deep reinforcement learning-based resource allocation can illustrate the temporal or spatial focus of the AI agent, enabling operators to grasp the learned policies. Counterfactual reasoning techniques further complement this by evaluating how minimal changes in input conditions could alter AI outputs, helping identify critical decision thresholds and increasing interpretability. These workflows not only facilitate debugging and model validation but also provide practical pathways for human experts to audit and adjust AI behaviors to network dynamics.

Despite these benefits, the widespread use of complex deep learning models often creates opaque, black-box systems, revealing a fundamental trade-off between prediction accuracy and interpretability [25]. Current research includes developing inherently interpretable models and post hoc explanation methods such as attention visualization, feature attribution, and counterfactual reasoning. While these methods have shown progress, they remain immature for comprehensive telecommunications applications and require further adaptation to meet domain-specific needs [24]. Moreover, explainability is essential not only for operational transparency but also for regulatory compliance and risk mitigation, as erroneous AI decisions could trigger cascading network failures [18]. The evolving ecosystems of Open RAN and emerging 6G networks introduce additional trust challenges because of their distributed and multi-agent learning frameworks, necessitating transparent AI mechanisms to ensure security, fault tolerance, and interoperability [25].

In summary, advancing explainability in AI systems is a vital challenge underpinning systemic trust, responsible deployment, and regulatory adherence in telecommunications. Continued efforts toward interpretable AI models, tailored explanation techniques, and integration of trust frameworks are imperative for the reliable and transparent operation of next-generation networks.

9.10 Interoperability and Standardization Challenges

The integration of AI into telecommunications networks faces considerable interoperability and standardization challenges arising from diverse multi-vendor equipment, heterogeneous protocol stacks, and disparate technology domains [13, 18, 25]. Fragmented AI models and incompatible data schemas hinder seamless AI-driven control and coordination across Open RAN components, including radio units (RU), distributed units (DU), and centralized units (CU) [18]. The lack of unified AI interfaces and standardized telemetry data formats creates barriers to implementing distributed intelligence and federated learning, constraining scalability and limiting cross-vendor collaboration [13]. Additionally, varying regulatory requirements and privacy policies across jurisdictions and operators compound the fragmentation in AI adoption. Early efforts by standardization bodies to define AI-specific protocols, interfaces, and data representations are ongoing but remain in initial stages, despite their critical role in enabling modular, plug-and-play AI capabilities and ensuring reproducibility and reliability of AI-enhanced network functions [25].

Table 24 summarizes key standards initiatives addressing these challenges. These efforts can be leveraged in real deployments to enhance interoperability and standardization as follows. The O-RAN Alliance's AI/ML Working Group's definition of standardized AI model interfaces and telemetry data formats enables distributed intelligence and federated learning mechanisms to function across heterogeneous Open RAN components, which has been demonstrated to improve throughput, latency, and energy efficiency in real network scenarios [25]. Similarly, 3GPP SA6 integrates AI data models within 5G/6G network management frameworks to provide seamless AI control over multiple network slices, which supports dynamic, real-time resource allocation and service optimization under diverse traffic conditions [13]. The ETSI ENI framework facilitates context-aware AI management and closed-loop automation workflows, translating directly into improved fault detection and network self-healing capabilities observed in testbeds [13]. Furthermore, IEEE's work on interoperability for AI systems aims to standardize data exchange formats and reproducibility, which provides essential foundations for multi-vendor AI component integration and reliable operation in production networks [18]. The ITU-T Focus Group contributes by establishing standardized machine learning workflows and privacy frameworks that are critical for compliance with regional regulations while enabling collaborative AI model training across operators [25].

Real-world applications of these standards help overcome current challenges such as AI model silos, high computational overhead, and privacy concerns by enabling modular, interoperable AI solutions that can adapt to heterogeneous environments. Deployments have reported up to 92% accuracy in traffic classification and significant latency and throughput improvements when leveraging AI-enhanced SDN controllers conforming to 3GPP and O-RAN specifications [13]. Likewise, federated learning compliant with these standards safeguards user data privacy while enabling efficient distributed training, which is vital for scalability [25]. Altogether, these standardized approaches form a foundational ecosystem allowing telecom operators to deploy AI-driven network functions with greater scalability, flexibility, and cross-vendor collaboration, thus accelerating the shift toward fully automated and optimized next-generation networks.

Addressing these interoperability gaps demands interdisciplinary initiatives focused on harmonizing software stacks, unifying data semantics, and developing AI models robust to domain shifts and heterogeneity across multi-technology ecosystems.

9.11 Security and Robustness

As AI technologies permeate critical telecommunications infrastructure, ensuring security and robustness against adversarial threats, data poisoning, and erroneous model outputs is fundamental to operational reliability [5, 18, 24]. AI models are vulnerable to attacks exploiting weaknesses in training data, model parameters, and inference processes, potentially resulting in misclassifications, compromised routing decisions, or degradation of service quality. Such attacks are especially harmful in complex AI-enabled networks where flawed decisions may cascade, causing widespread disruptions across multiple layers and services [24]. Defensive strategies include adversarial training, development of robust model architectures, sophisticated anomaly detection systems, and hierarchical control mechanisms equipped with fallback options to mitigate AI failures [5]. Additionally, integrating explainability aids in the early detection of abnormal AI behaviors, while federated learning frameworks can minimize insider threats by limiting data exposure [18].

Given the resource constraints common at the network edge in 5G and beyond, emerging lightweight security solutions are essential. These include model compression techniques such as pruning and quantization to reduce computational overhead while preserving robustness, enabling deployment of secure AI on edge devices [18, 24]. Lightweight anomaly detection models and efficient encryption schemes tailored for edge environments enhance protection without imposing significant latency or energy costs [24]. Moreover, federated learning and edge computing synergize to protect data privacy and reduce centralized vulnerabilities, balancing the privacy-utility trade-offs crucial under regulatory frameworks [18]. These approaches facilitate real-time inference with acceptable security guarantees in constrained settings, addressing a critical challenge for scalable AI-driven network management [24].

Recent advances demonstrate the promise of agentic AI approaches leveraging large language models (LLMs) embedded within flexible Open Radio Access Network (O-RAN) architectures to enhance resilience. Such LLM-driven agents improve fault detection accuracy and mitigation success by autonomously monitoring network telemetry, interpreting complex faults through natural language understanding, and executing self-healing actions [5]. Experimental evaluations show fault detection accuracy increasing from 78% to 95%, and mitigation success rate rising from 70% to

Table 24: Current Standards Initiatives Addressing AI Interoperability and Standardization in Telecommunications

Standards Body	Initiative/Group	Scope and Focus
O-RAN Alliance	AI/ML Working Group	Defines AI model interfaces, standardized telemetry data formats, and protocols for Open RAN components (RU, DU, CU) to realize distributed intelligence and federated learning [25].
3GPP	SA6 (Enhancements for AI)	Integration of AI data models and interfaces in 5G/6G network management for seamless AI control across heterogeneous network slices [13].
ETSI	Experiential Networked Intelligence (ENI)	Framework for context-aware AI management, standardizing AI-driven closed-loop automation workflows and interoperability between network elements [13].
IEEE Standards Association	P1931.1 (Interoperability for AI Systems)	Developing guidelines for AI component interoperability, data exchange formats, and reproducibility standards applicable to telecommunications infrastructures [18].
ITU-T	Focus Group on Machine Learning for Future Networks	Standardizing machine learning workflows, data representation, and privacy frameworks for multi-vendor network environments [25].

91%, alongside a 40% reduction in downtime and a drop in throughput degradation from 25% to 10%, significantly enhancing network performance. These improvements are summarized in Table 25. However, these benefits must be balanced against challenges including computational overhead, potential erroneous decisions, security risks, and interoperability issues among multi-vendor environments.

Mitigation strategies for these agentic AI limitations involve hierarchical agent designs that delegate tasks across specialized, lightweight AI components to reduce latency and computational load. Continuous validation frameworks employing real-time monitoring, anomaly detection, and rollback mechanisms ensure erroneous decisions are promptly identified and corrected, maintaining operational integrity [5]. Security risks are addressed by integrating multi-layered defensive architectures combining traditional cybersecurity practices with AI-specific safeguards, such as adversarial input detection and secure model update protocols. Interoperability challenges within heterogeneous multi-vendor environments motivate the development of standardized AI interfaces, modular architectures, and open data formats to facilitate secure and scalable collaboration [5]. Such combined approaches ensure that agentic AI systems can reliably support self-healing and adaptive functionalities without compromising network security or performance.

Therefore, designing AI systems with intrinsic robustness, continuous validation processes, and adaptive security measures is imperative for their trustworthy integration into future telecommunication infrastructures. Scalability requires efficient AI architectures combining algorithmic compression, hardware acceleration, and modular design to meet real-time inference demands, while privacy preservation leverages federated learning, edge computing, and lightweight models balancing privacy-utility trade-offs under regulatory constraints [18, 24]. Explainability remains critical for building trust, auditability, and regulatory compliance amidst opaque deep models [24], and interoperability challenges arising from multi-vendor heterogeneity necessitate standardized AI interfaces and data formats supporting scalable collaboration [5]. Security demands robust defenses against adversarial attacks and errors, incorporating fallback mechanisms and continuous validation without sacrificing system performance.

Collectively, these emerging techniques and frameworks underscore that ensuring security and robustness in AI-driven telecommunications requires a multilayered approach combining advanced AI methodologies, network architecture flexibility, and rigorous operational safeguards.

10 Synthesis and Future Directions

The surveyed literature indicates a rapidly evolving landscape at the intersection of artificial intelligence, control theory, and wireless communication technologies. To guide future research and practical implementations, it is essential to synthesize the key methodological advancements and outline potential disruptive paradigm shifts, as well as address challenges encountered when transitioning from theory to real-world applications. In the following, we connect each future research challenge to prior thematic discussions presented in this survey, propose detailed roadmaps, suggest concrete milestones, and highlight actionable research questions that can foster progress in AI-enabled wireless control systems.

Implementing these research directions requires a coordinated effort bridging AI algorithm design, control theory, wireless protocol development, and practical system engineering. For example, scalability challenges connect back to distributed learning methods in Section 3 and resource-aware wireless schemes in Section 4, while robustness aspects rely heavily on control-theoretic guarantees discussed in Section 5. The integration of AI and classical control, emphasized throughout Section 6, remains a core research avenue demanding new theoretical insights and interdisciplinary validation. Low-latency demands specifically leverage advances in AI-aided scheduling from Section 4, and security concerns draw from privacy-preserving and adversarial defense mechanisms in Section 7.

Beyond research, practical implementation pathways must incorporate open-source platforms and standardized benchmarks to enable transparent and systematic evaluation. Existing initiatives such as [placeholder for relevant open-source frameworks] can be expanded to incorporate wireless control scenarios, enabling the community to share datasets, algorithms, and deployment experiences. Field trials and industrial collaborations will be vital to validate theoretical predictions, uncover hidden challenges, and guide algorithmic refinements toward deployable AI-enabled wireless control systems.

This synthesis articulates a comprehensive roadmap that tightly links future opportunities with concrete research questions and milestones, grounded in the thematic structure of this survey. Such clarity and specificity are crucial for advancing the state of the art and accelerating adoption of robust, scalable, and secure AI-driven wireless control in real-world applications.

10.1 Methodological Contributions and Frameworks

The integration of AI techniques with control and wireless systems has led to innovative frameworks that enable adaptive, resilient, and efficient operation in complex environments. To advance this field, future research should prioritize the development of comprehensive and standardized methodological frameworks that unify these interdisciplinary approaches. Such frameworks would support consistent benchmarking, reproducibility, and scalability of results across diverse applications. Essential attributes of these

Table 25: Performance Improvements of LLM-Driven Agentic AI in O-RAN Resilience [5]

Metric	Baseline	Proposed LLM-Driven Agent	Improvement
Fault Detection Accuracy	78%	95%	+17%
Mitigation Success Rate	70%	91%	+21%
Downtime Reduction	-	40%	-
Throughput Degradation	25%	10%	-15%

Table 26: Summary of Future Research Challenges and Opportunities in AI-Enabled Wireless Control Systems

Research Challenge	Opportunity	Actionable Research Questions / Milestones
Scalability and Complexity	Development of scalable AI algorithms for large-scale networks	How can AI models be optimized for computational complexity and energy efficiency in massive deployments? Milestones include prototyping scalable distributed control algorithms tested in large network simulations and physical testbeds that assess industrial or urban environments. Specific projects could involve optimizing energy consumption in edge devices and designing adaptive models that gracefully degrade performance under constrained computational resources. These efforts link to Section 3 on distributed AI frameworks and Section 4 on resource-aware wireless protocols.
Integration of AI and Control Theory	Synergistic methods combining AI learning with classical control	How to design hybrid frameworks that provide formal stability guarantees while enabling AI adaptability? Research directions include proposing architectures that tightly couple online learning with established methods like model predictive control or Lyapunov-based techniques, backed by rigorous stability proofs and simulations (Section 2). Example projects may investigate interpretable AI models facilitating safety-critical applications and validating these frameworks in complex wireless control scenarios.
Low-Latency and Reliability	AI approaches enabling ultra-reliable low-latency communication (URLLC)	What specialized AI driven scheduling, and resource allocation schemes can fulfill stringent URLLC requirements in wireless control loops? Challenges involve predicting channel states and network congestion under dynamic conditions. Milestones include developing latency-aware AI models for dynamic scheduling policies (Section 4), benchmarking the delay-reliability trade-offs in real-time control loops, and integrating adaptive control policies capable of mitigating packet loss effects. Practical projects could implement these models in industrial IoT settings requiring millisecond-level responsiveness.
Data Privacy and Security	AI methods preserving privacy and ensuring network security	How to safeguard data privacy and defend against adversarial attacks in AI-enabled wireless control systems? Research should target secure federated learning frameworks embedded within control architectures, differential privacy mechanisms, and defenses against adversarial inputs (Section 7). Concrete milestones include analyzing system vulnerabilities through simulated and physical attack scenarios, designing robust AI models resilient to manipulations, and disseminating effective mitigation strategies in realistic deployments.
Real-World Deployment	Bridging theory-to-practice gap through prototyping and field trials	What practical barriers—such as hardware limitations, scalability, and unforeseen environmental effects—hinder real-world adoption? Projects must include deploying pilot AI-enabled wireless control systems in industrial and urban environments, collecting detailed empirical data on system performance and failures, and developing open-source toolkits alongside standardized benchmarks to promote reproducibility and community collaboration. These efforts span the thematic breadth of all sections discussed.

frameworks include robustness against system uncertainties, real-time adaptability to dynamic and stochastic conditions, and data-driven optimization strategies that enhance system performance. Moreover, preserving the theoretical guarantees foundational to control theory and wireless communications is critical to ensuring reliability, safety, and interpretability in AI-enabled systems. By establishing these unified frameworks, research can bridge current gaps between disparate methodologies, streamline integration efforts, and accelerate the practical deployment of intelligent control and wireless systems in real-world scenarios.

10.2 Practical Implementation Pathways and Challenges

While theoretical advancements abound, the practical implementation of AI-enabled control and wireless systems confronts several critical challenges. These challenges include computational resource constraints, stringent latency requirements, maintaining robustness amid dynamic and uncertain environments, and seamless integration with existing legacy infrastructure. To address these issues, future research should prioritize the development of scalable and efficient algorithms that are compatible with edge computing paradigms and distributed architectures, thereby reducing latency and resource demands. Furthermore, as AI systems are increasingly deployed in sensitive and mission-critical applications, it is essential to rigorously manage privacy, security, and ethical considerations throughout the design and deployment phases to ensure trustworthy and responsible AI operation.

10.3 Potential Disruptive Innovations and Paradigm Shifts

Emerging trends indicate several potential disruptions in the field, each bearing transformative implications. A major innovation lies in the convergence of AI, control theory, and wireless communication technologies, which is paving the way for self-organizing and self-optimizing networks. These networks promise to dramatically enhance efficiency and responsiveness by autonomously adjusting to environmental and operational changes. For instance, autonomous network management systems leveraging reinforcement learning algorithms can dynamically allocate resources without human intervention, thereby improving latency and throughput under varying network conditions.

Another anticipated paradigm shift arises from the fusion of model-based and data-driven approaches, yielding hybrid methods that harness the complementary advantages of both paradigms. Such approaches integrate the theoretical guarantees and interpretability of model-based designs with the adaptability, scalability, and learning capacity of data-driven techniques. This synergy is exemplified by recent prototypes in channel estimation and signal processing, which demonstrate superior robustness and performance relative to purely model-based or purely data-driven counterparts.

Furthermore, advancing research on cross-layer design that integrates AI across multiple system levels offers significant potential to transform conventional architectures. By enabling joint optimization from the physical layer to the application layer, AI-enabled cross-layer protocols facilitate comprehensive and holistic system adaptation. For example, these protocols have shown to improve quality of service in next-generation networks through simultaneous optimization of routing, scheduling, and power control.

Collectively, these emerging innovations embody a shift toward highly integrated, intelligent network systems capable of autonomously managing complex and dynamic environments. This represents a substantial departure from traditional communication system designs, signaling new frontiers and opportunities in the development of future networks.

10.4 Comparative Summary of Key Approaches

To facilitate a clear understanding of the relative merits, limitations, and practical considerations of the approaches discussed throughout this survey, Table 27 provides a detailed comparative summary. It highlights the principal characteristics, implementation challenges, and key application domains of representative methods, enabling readers to grasp their contextual suitability and trade-offs.

10.5 Interdisciplinary Synergies

The symbiotic relationship between AI, control theory, and wireless communication is poised to deepen significantly, paving the way for innovative system designs that overcome traditional domain-specific boundaries. To foster impactful progress, researchers should actively encourage interdisciplinary collaborations that integrate advances across sensing technologies, computational methodologies, and theoretical frameworks. These collaborations are essential for developing comprehensive and robust solutions that effectively

Table 27: Comparative summary of key AI-based control and wireless methodologies

Approach	Methodological Strengths	Practical Challenges	Application Domains
Model-based control with AI integration	Strong theoretical foundations and interpretability facilitate reliability and safety assurances	High computational complexity and the need for accurate system models may limit adaptability	Autonomous systems, robotics, and safety-critical control
Data-driven machine learning methods	High adaptability and scalability enable handling complex, dynamic environments	Requires large volumes of data and may lack formal performance guarantees	Network optimization, resource management, and traffic prediction
Hybrid model-data fusion frameworks	Combines theoretical rigor with flexibility to improve robustness	Complexity in integration and parameter tuning poses practical hurdles	Smart grids, 5G/6G communication networks, and cyber-physical systems
Reinforcement learning for control	Enables real-time learning and policy optimization in uncertain environments	Balancing exploration and exploitation remains challenging, along with convergence guarantees	Unmanned aerial vehicles (UAVs), Internet of Things (IoT) device management
Edge AI and distributed architectures	Supports low-latency processing and scalable deployment across network edges	Communication overhead and privacy concerns require careful management	Smart cities, industrial automation, and distributed sensing

tackle complex real-world challenges by leveraging complementary strengths from each field.

Moving forward, consolidating diverse methodological advances into unified frameworks that seamlessly integrate AI, control, and wireless communication technologies will be crucial. This integration must explicitly address practical implementation challenges, such as scalability, latency, and reliability, to facilitate the translation of theoretical innovations into deployable systems. Furthermore, anticipating and adapting to emerging paradigm shifts—including hybrid models and adaptive, data-driven control strategies—will promote the creation of more flexible and resilient system architectures. By emphasizing and strengthening these interdisciplinary synergies, the research community can accelerate the development of autonomous, efficient, and dependable systems with broad and meaningful impact across diverse applications.

10.6 Synergies Across AI, Resilient Control, and Wireless Technologies

The fusion of artificial intelligence (AI), resilient control strategies, and wireless technologies has driven substantial progress toward real-time, adaptive, and secure network management within dynamic and uncertain environments. A salient example of this interdisciplinary synergy is the development of adaptive neural network finite-time control methods designed for nonlinear systems with unknown time delays, actuator faults, and false data injection (FDI) attacks. As detailed in [3], such systems are modeled with unknown, possibly time-varying delays and combined faults/attacks affecting actuators and sensors. Radial basis function neural networks approximate the unknown nonlinear function $f(x(t), x(t-\tau))$ by representing it as $W^T\Phi(x(t), x(t-\tau)) + \varepsilon$, where W are adaptive weights estimated online via adaptive laws. An observer-based fault detection mechanism estimates system states and faults despite measurement discrepancies introduced by FDI attacks. The control law combines these elements to ensure finite-time convergence of the state and estimation errors, providing a significant resilience enhancement compared to traditional asymptotic methods by enabling faster responses in the presence of substantial unknown disturbances. Simulation results on benchmark nonlinear systems demonstrate the method's robustness and superiority in stabilizing states and detecting faults under complex cyber-physical threats.

Definition 10.1 (Finite-Time Convergence). Finite-time convergence refers to the property of a system's state or estimation error reaching an equilibrium exactly within a finite time interval, in contrast to asymptotic convergence where the equilibrium is approached gradually over infinite time.

In parallel, AI has invigorated wireless communications through advanced paradigms such as Perceptive Mobile Networks (PMNs). These networks integrate coordinated beamforming with deep learning algorithms to predict and mitigate interference in complex multi-cell environments [**?**]. Specifically, an AI-empowered framework dynamically allocates resources by learning interference patterns through coordinated sensing across multiple base stations, exploiting macro-diversity and array gains. This approach simultaneously optimizes beamforming strategies and resource scheduling to enhance both communication quality and sensing accuracy, improving the sensing signal-to-interference-plus-noise ratio (SINR) in heterogeneous and interference-prone conditions. Such adaptive resource orchestration helps maintain robust wireless communication performance despite challenges like low-latency inference and practical channel acquisition constraints.

Definition 10.2 (Perceptive Mobile Networks (PMNs)). PMNs are wireless networks that combine communication and sensing functionalities by leveraging advanced signal processing and AI to provide both data transmission and environment perception capabilities.

Further extending this ecosystem, AI-driven optimization of reconfigurable intelligent surfaces (RIS) presents a powerful approach for shaping wireless propagation environments. By learning mappings from channel state information to effective RIS configurations, these AI-empowered systems enhance spectral efficiency and robustness under uncertain and time-varying channel conditions [18]. This integration of AI, control theory, and wireless technologies exemplifies how networks can achieve context-aware, adaptive decision-making and resilient operation amid cyber-physical uncertainties, advancing next-generation secure and adaptive wireless systems.

Definition 10.3 (Reconfigurable Intelligent Surfaces (RIS)). RIS are planar surfaces consisting of numerous small elements whose electromagnetic properties can be dynamically manipulated to control the propagation of wireless signals in the environment.

10.7 Critical Enablers

A set of pivotal enablers underpins the convergence of AI, control, and wireless technologies. These enablers not only address fundamental technical challenges but also represent active research areas with diverse approaches and trade-offs, as summarized in Table 28.

Below, we provide a critical discussion of each enabler with examples from current studies and highlight open research gaps, including relevant performance observations and application impacts:

Federated Learning. This approach facilitates collaborative model training across decentralized nodes without raw data exchange, addressing stringent privacy concerns inherent in wireless networks. In Open RAN contexts, federated learning enables dynamic spectrum management and fault detection by aggregating local AI insights, leading to improved spectral efficiency and robustness

Table 28: Summary and Analysis of Critical Enablers in AI-Driven Wireless Networks

Enabler	Function and Benefits	Challenges and Controversies	Illustrative Example
Federated Learning	Enables distributed intelligence while preserving privacy, e.g., collaborative spectrum management	Communication overhead, model convergence issues, heterogeneous data distributions	Open RAN dynamic spectrum allocation [6]
Privacy-Preserving AI	Protects sensitive user and network data, maintains user trust	Balancing privacy with model accuracy and efficiency; trade-offs among differential privacy, encryption, and secure computation	Secure data exchange in multi-operator networks
Edge Intelligence	Decentralizes computation to reduce latency and optimize resource usage at the network edge	Limited edge resources, coordination across nodes, model consistency, heterogeneity of devices	Real-time adaptive interference mitigation
Explainable AI	Provides transparency and interpretability for black-box models, enhancing trust and regulatory compliance	Complexity of explanations, potential trade-off between interpretability and accuracy [24]	Visualizing reinforcement learning decision paths
Scalable Distributed Architectures	Support multi-agent RL and adaptive control across heterogeneous and large-scale network segments	System complexity, scalability bottlenecks, interoperability issues	Large-scale network fault detection systems

[6]. Experimental analyses demonstrate that AI-enabled methods using federated learning outperform traditional heuristics by effectively adapting to dynamic channel conditions. However, challenges such as communication overhead, non-iid data distributions, and convergence difficulties still limit scalability and require further algorithmic enhancements to optimize trade-offs between learning accuracy and resource consumption.

Privacy-Preserving AI. Closely related to federated learning, privacy-preserving mechanisms—including differential privacy, secure multi-party computation, and homomorphic encryption—protect sensitive user and network information while maintaining user trust. These methods must carefully balance privacy guarantees against model accuracy and computational overhead. In wireless network scenarios, this balancing act is crucial to operating within constrained latency and processing budgets. Ongoing research focuses on identifying optimal configurations that minimize the impact on inference performance while securing data, especially in multi-operator environments where data sharing is sensitive.

Edge Intelligence. By decentralizing processing to network edge nodes, edge intelligence significantly reduces latency and bandwidth consumption compared to centralized cloud architectures. Applications such as real-time interference mitigation and localized adaptive control benefit from edge deployment, achieving faster response times and more fine-grained network adaptation. Yet, limited computational resources and the heterogeneity of edge devices pose constraints on model complexity and scalability. Coordinating distributed inference while maintaining model consistency and coping with device diversity remain open challenges, requiring innovative distributed learning algorithms and system designs to harness the edge advantage without sacrificing accuracy.

Explainable AI. Given the widespread deployment of complex neural architectures and reinforcement learning agents in network management [24], explainability enhances system transparency, fosters user trust, and supports compliance with regulatory standards. Techniques that interpret AI decisions and visualize agent behaviors enable operators to identify system weaknesses and biases, thus improving reliability. Performance studies highlight the trade-off that often exists between interpretability and prediction accuracy, necessitating new model designs tailored for wireless applications that maintain both high performance and explainability. This is vital as network decisions impact service quality and security.

Scalable Distributed Architectures. Addressing the scale and complexity of next-generation wireless systems demands distributed frameworks that combine multi-agent reinforcement learning and adaptive control. Such architectures provide resilience and fault tolerance across diverse, heterogeneous network segments. Evaluations of large-scale deployment scenarios reveal that these frameworks can improve fault detection accuracy and system responsiveness. Nonetheless, challenges related to system complexity, scalability bottlenecks, and interoperability among varied network components remain. Research efforts are devoted to developing scalable, interoperable, and efficient distributed frameworks that preserve performance while managing networking and computational overhead.

In summary, these enablers collectively advance AI's feasibility in wireless and cyber-physical systems by tackling privacy, interpretability, latency, and scalability challenges. Performance insights from recent studies demonstrate notable gains in spectral efficiency, network robustness, and management accuracy attributable to these technologies. However, ongoing research is essential to resolve existing trade-offs between performance, complexity, and trustworthiness, ultimately enabling the realization of fully intelligent and adaptive wireless networks.

10.8 Identified Research Needs

This section outlines the key research gaps that must be addressed to advance AI-enhanced resilient control and wireless systems. The primary objectives are to develop computationally efficient, robust, low-latency, interpretable, and interdisciplinary AI frameworks that achieve resilient and scalable performance in dynamic cyber-physical environments. Addressing these challenges will enable reliable operation, real-time responsiveness, and user trust in next-generation cyber-physical infrastructures. To guide research focus, the needs are ranked by urgency and potential impact as shown in Table 29.

Computational Complexity Reduction: This need is the highest priority due to the critical demand for deployable solutions on resource-constrained devices. Current adaptive neural network finite-time resilient control approaches for nonlinear time-delay systems demonstrate robustness to unknown actuator faults and cyber-attacks but require high neural network capacity and exhibit sensitivity to noise and parameter tuning [3]. Concrete strategies include dynamic pruning techniques that remove less significant network connections during training to lower complexity, distributed resilient control frameworks leveraging multi-agent consensus protocols to distribute computation, neural network compression methods such as quantization and low-rank decompositions, and FPGA or ASIC implementations tailored for hardware-efficient inference and control.

Robustness Against Uncertainties: Ranked second given the substantial impact adversarial conditions exert on system reliability. AI models deployed in networked sensing and control remain vulnerable to adversarial perturbations, noisy measurements, imperfect channel state information, and dynamic interference patterns [18]. Robust AI training methodologies including adversarial

Table 29: Summary of Key Research Challenges Ranked by Urgency and Potential Impact, with Specific Illustrative Techniques

Rank	Research Need	Challenges	Potential Solution Paths and Illustrations
1	Computational Complexity Reduction	High neural network capacity requirements, sensitivity to noise and parameter tuning, time-varying delays	Adaptive dynamic pruning (e.g., magnitude-based pruning), distributed resilient control via multi-agent consensus to share computational load, neural network quantization and low-rank factorization, hardware-efficient algorithms leveraging FPGA and ASIC implementations [3]
2	Robustness Against Uncertainties	Vulnerability to adversarial perturbations, noisy telemetry, imperfect channel information, and dynamic interference	Resilient AI with robust training (e.g., adversarial training, domain randomization), cooperative multi-agent reinforcement learning enhancing diversity and coordination, real-time adaptation via online learning, interference mitigation using coordinated beamforming and AI-based channel estimation [18]
3	Latency Minimization	Real-time inference demands, resource constraints, accuracy-latency trade-offs in edge deployments	Lightweight architectures such as MobileNets and EfficientNets, hardware accelerators including GPUs and TPUs, scalable reinforcement and deep learning for dynamic network management optimization, edge-cloud synergy enabling adaptive offloading of intensive tasks [24]
4	Interpretability Enhancement	Lack of transparency in AI decisions, regulatory and operational trust issues	Explainable AI frameworks such as SHAP and LIME providing model-agnostic insights, model introspection techniques like layer-wise relevance propagation, visualizing decision pathways, domain-aware explanations tailored to control and networking contexts
5	Interdisciplinary Collaboration	Integrating diverse expertise across control, communications, and AI domains	Holistic multi-layer frameworks combining algorithmic design, hardware optimization, and network layering; fostering joint research initiatives and collaborative platforms accelerating innovation

training and domain randomization improve resilience. Cooperative multi-agent reinforcement learning leverages diversity and coordination to mitigate interference through coordinated beamforming and AI-enabled channel estimation. Real-time adaptive learning mechanisms empower systems to counteract evolving disturbances and adversarial inputs dynamically.

Latency Minimization: Placed third to reflect stringent real-time operational demands, especially in edge computing environments. Balancing inference speed and accuracy under limited computational resources remains challenging [24]. Lightweight neural network architectures such as MobileNets and EfficientNets facilitate fast inference. Specialized hardware accelerators like GPUs, TPUs, and ASICs significantly reduce computation latency. Scalable reinforcement learning and deep learning algorithms dynamically optimize resource allocation and network management. Synergistic edge-cloud computing configurations enable adaptive offloading of resource-intensive tasks to satisfy low-latency requirements.

Interpretability Enhancement: This research need supports operational trust and compliance with regulatory frameworks, crucial in safety-critical cyber-physical systems. Explainable AI (XAI) techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) provide insights into AI decision rationale. Model introspection methods including layer-wise relevance propagation and decision path visualization facilitate effective debugging and validation. Tailoring explanation modalities to the control and wireless communication domains enhances interpretability and user trust.

Interdisciplinary Collaboration: Although ranked last, this remains essential for comprehensive progress in AI-empowered cyber-physical systems. Integrative efforts spanning control theory, wireless communications, and AI are vital to address complex system challenges holistically. Developing joint frameworks that unify algorithmic innovation, hardware design, and network layering, alongside fostering cross-disciplinary research consortia, accelerates advancements toward resilient, scalable AI infrastructures.

In summary, the prioritized research objectives detailed here target the creation of resilient, scalable, interpretable, and efficient AI solutions tailored for complex cyber-physical environments. By clearly articulating challenges alongside concrete illustrative techniques, this framework aims to concentrate future research efforts and bridge critical gaps, thereby enabling next-generation systems to achieve reliable, real-time operations with enhanced robustness and trustworthiness.

10.9 Anticipated Innovations

This section explicitly highlights the key objectives of forthcoming advancements within AI-empowered wireless networks, focusing on transformative innovations that promise enhanced autonomy, security, and efficiency while addressing current implementation challenges. The synthesis frames these innovations within a conceptual model emphasizing the interplay of technological capabilities, operational constraints, and mitigation pathways, balanced by current feasibility considerations and expected development timelines.

Multi-Agent Collaborative Learning: Distributed learning and decision-making across network nodes promise improved adaptability and fault tolerance in complex environments. Multi-agent reinforcement learning schemes, notably within Open RAN contexts, have demonstrated significant gains in resource allocation, anomaly detection, and network resilience [6, 25]. However, challenges such as coordination complexity, communication overhead, scalability, and adversarial robustness remain substantial. To operationalize mitigation strategies, future research may explore hierarchical agent architectures that reduce coordination burdens by structuring agents into manageable groups and enable local decision-making with limited inter-agent communication. Lightweight consensus mechanisms can minimize latency and computational demands through simplified agreement protocols suited for dynamic network topologies. Federated learning frameworks incorporating privacy-preserving and adaptive convergence protocols can be deployed in practice by integrating secure aggregation and differential privacy techniques, balancing privacy with real-time responsiveness. These advancements are crucial for scalable, decentralized AI-enabled networks, providing practical pathways for deployment by leveraging modular algorithm designs tested in evolving Open RAN testbeds [25].

Hardware Acceleration: Specialized AI accelerators embedded in edge devices can substantially reduce inference latency and energy consumption, enabling real-time adaptive control and interference mitigation tasks essential for next-generation wireless systems [24]. The complexity of hardware design and integration challenges necessitates modular accelerator architectures offering configurable performance-to-energy ratios tailored to resource-constrained environments. In practice, this could be realized by adopting programmable hardware fabrics and cross-layer hardware-software co-design that allow flexible adaptation to varying workloads. Cost-effective fabrication methods and open accelerator standards will be pivotal for widespread adoption and interoperability with heterogeneous network elements. Collaborative ecosystems between hardware designers and network researchers can accelerate development cycles, and deployment-ready frameworks that support plug-and-play accelerator modules will facilitate integration in existing network infrastructures. Progress in this area is anticipated to mature steadily within the upcoming 2–4 years, driven by these practical implementation considerations.

Quantum Computing Integration: Quantum technologies hold promise for breakthroughs in optimization speed and security, enabling accelerated resolution of complex network computation tasks beyond classical capabilities. Despite significant challenges, including immature hardware technology, high quantum error rates, and the need for application-specific quantum algorithms, ongoing research and experimental prototypes present viable mitigation strategies. Hybrid quantum-classical architectures, leveraging early

quantum processors for specialized optimization sub-tasks while maintaining classical control, offer a near-term approach to integration. Operational deployment may take the form of quantum accelerators interfaced with classical network management platforms to solve discrete optimization problems more efficiently. Concurrently, the development of noise-resilient quantum algorithms and error-correction protocols tailored to network optimization and security is a crucial research direction, likely progressing beyond 5 years from now, as hardware and algorithmic maturity improves. Overall, phased quantum adoption strategies that incrementally incorporate quantum advantages into existing systems will enhance practical readiness.

Blockchain Security Mechanisms: Blockchain-based approaches enhance the security of decentralized AI agents by ensuring data integrity and providing transparent audit trails, thereby reinforcing trustworthiness in collaborative learning systems [25]. Main challenges include blockchain-associated latency overhead, scalability limitations, increased computational demands, and privacy/regulatory compliance concerns. To operationalize mitigation strategies, scalable consensus protocols such as variants of proof-of-stake can be integrated into network systems to reduce energy consumption and transaction confirmation times. Off-chain processing techniques, including state channels and sidechains, enable execution of transactions without burdening the main chain, thereby reducing latency impacts critical for real-time operations. Privacy-preserving blockchain implementations combined with compliance-aware smart contracts facilitate data protection regulation adherence by embedding access controls and audit functionalities directly into the protocol. Careful architectural alignment of blockchain solutions with network performance imperatives, such as selective blockchain use in latency-tolerant tasks or hybrid on/off-chain models, is essential to minimize effects on latency-sensitive operations. These solutions are expected to mature within a 3–6 year horizon, integrating seamlessly with evolving AI-controlled network environments.

Collectively, these anticipated innovations envision fully autonomous intelligent networks characterized by self-healing capacities, context-aware adaptations, and proactive cyber-physical threat mitigation [5]. Achieving this vision requires managing trade-offs among computational complexity, scalability, security, and interoperability through integrated frameworks. Techniques such as explainable AI improve transparency and trust, hierarchical agent designs enhance multi-agent coordination, and lightweight model development supports efficient edge deployment. By explicitly linking each innovation's specific challenges to detailed and practicable mitigation strategies, this synthesis offers a coherent narrative that advances understanding of operationalization pathways. This creates a connected flow from identifying challenges to proposing actionable solutions within realistic development and deployment timelines.

The following table summarizes the key innovations, their benefits, associated challenges, and prospective mitigation strategies with improved clarity and readability:

In summary, this nuanced synthesis elucidates the multi-dimensional progress and outstanding challenges at the intersection of AI, resilient control, and wireless technologies. By explicitly linking benefits with challenges and forward-looking mitigation approaches, and situating innovations within practical feasibility and development timelines, it establishes a comprehensive and actionable roadmap. This roadmap guides the evolution of secure, adaptive, and intelligent networked systems capable of effectively addressing future demands and operational uncertainties.

11 Conclusion

This survey aimed to comprehensively review the integration of artificial intelligence (AI) techniques into telecommunication networks, focusing on their roles in adaptive control, wireless networking, routing, software-defined networking (SDN), Open Radio Access Networks (Open RAN), and autonomous fault management. Our objectives included synthesizing state-of-the-art AI methods, evaluating their benefits and challenges, and identifying future research directions to guide the evolution towards fully autonomous, self-optimizing network systems.

For example, AI-driven adaptive control algorithms have demonstrated real-time optimization of network parameters, reducing latency and improving throughput in 5G deployments. In wireless networking, reinforcement learning has been effectively employed for dynamic spectrum allocation, illustrating significant gains in spectral efficiency. Similarly, AI-enabled routing strategies have shown resilience and adaptability in rapidly changing network topologies, while advancements in SDN and Open RAN architectures leverage machine learning to enable programmability and flexibility at unprecedented scales. Autonomous fault management systems utilizing anomaly detection and predictive maintenance approaches have improved network reliability by preempting failures before service degradation occurs.

Despite these promising outcomes, challenges remain. The limitations of current AI models include data scarcity in certain network segments, interpretability obstacles in deep learning applications, and concerns related to security and privacy. Open questions persist regarding the standardization of AI integration across heterogeneous network infrastructures, the establishment of robust metrics to evaluate AI performance in live networks, and the development of scalable solutions viable for future 6G systems.

Looking forward, explicit research agendas should emphasize the design of explainable AI models with quantifiable trustworthiness, integration of cross-layer intelligence to optimize end-to-end network performance, and the creation of adaptive frameworks that balance automation with human oversight. Success metrics must include not only traditional network KPIs such as latency, throughput, and reliability but also AI-specific factors like fairness, robustness against adversarial conditions, and energy efficiency. Progress in these areas will be critical to achieving fully autonomous, self-optimizing telecommunication networks that meet the demands of next-generation applications.

Table 30: Summary of Anticipated Innovations: Benefits, Challenges, and Mitigation Approaches

Innovation	Benefits	Challenges / Limitations	Potential Mitigation Strategies
Multi-Agent Collaborative Learning	Enhanced adaptability, fault tolerance, and optimized resource management [6, 25]	Coordination complexity, scalability, communication overhead, adversarial robustness	Hierarchical agent designs enabling local coordination, lightweight consensus protocols reducing latency, federated privacy-preserving frameworks with adaptive convergence
Hardware Acceleration	Reduced inference latency and energy consumption enabling real-time adaptive control [24]	Design complexity, integration challenges, cost considerations, balancing energy and compute resources	Modular configurable architectures with programmable hardware fabrics, open standards, cost-effective fabrication and hardware-software co-design
Quantum Computing Integration	Potential for accelerated optimization, enhanced security in network operations	Immature hardware, high error rates, need for specialized quantum algorithms	Hybrid quantum-classical systems interfacing quantum accelerators with classical control, noise-resilient quantum algorithms, advanced error correction
Blockchain Security Mechanisms	Enhanced data integrity, auditability, and trust in decentralized AI [25]	Latency overhead, scalability constraints, high computational demand, privacy and regulatory concerns	Scalable proof-of-stake consensus, off-chain computation (state channels, sidechains), privacy-preserving protocols, compliance-aware smart contracts

11.1 Key Contributions and Findings

The integration of AI into various telecommunication domains has markedly enhanced functionalities such as adaptive control, wireless networking, routing, SDN, Open RAN, and autonomous fault management. AI-driven adaptive control strategies utilize advanced predictive models to dynamically optimize network resource allocation, thereby improving the network's responsiveness to variable traffic loads and heterogeneous service demands.

Within wireless networking and routing, machine learning techniques—especially ensemble methods like gradient boosting—have proven highly effective in capturing complex nonlinear patterns and addressing class imbalance issues endemic to network datasets. For instance, as demonstrated in recent work [27], gradient boosting methods such as CatBoost and LightGBM significantly outperform traditional algorithms in telecom customer churn prediction. These methods effectively manage nonlinearities and class imbalances without relying on external resampling techniques like SMOTE, while leveraging ensemble architectures that reduce errors and enhance robustness. The key predictive features identified include customer tenure, service usage, and payment methods. Despite the improved predictive accuracy, these advanced models require more computational resources and entail challenges such as hyperparameter tuning and balancing interpretability. This highlights a trade-off between predictive performance and computational efficiency in practical deployments.

Collectively, these advances underscore AI's critical contribution to enhancing efficiency, resilience, and predictive capabilities in contemporary telecommunication infrastructures.

11.2 Challenges and Future Directions

This section aims to outline the primary challenges currently faced when deploying AI in telecommunication networks and to discuss future directions for research and development in this area. The goal is to provide a balanced view by highlighting not only the obstacles but also ongoing debates and potential solutions emerging from recent advances.

Looking ahead, the evolution of telecommunication networks is trending towards fully autonomous, self-optimizing systems capable of continuous self-monitoring and dynamic adjustment. However, deploying AI solutions at scale introduces significant challenges:

- **Scalability concerns**: The computational demands of complex models such as gradient boosting need to be balanced with the requirement for real-time responsiveness. While powerful, these models often require considerable resources, which may limit their applicability in latency-sensitive network operations. Alternative approaches or model simplifications may offer trade-offs between performance and efficiency.

- **Security and privacy vulnerabilities**: AI-integrated control loops may be susceptible to adversarial attacks, posing risks to network stability and data confidentiality. Ensuring the robustness of AI models in the presence of malicious inputs remains an open research area, with ongoing discussions about secure model training and deployment practices.

- **Interoperability difficulties**: The heterogeneous and multi-vendor nature of telecom environments complicates seamless AI solution integration. Standardization efforts and common frameworks are critical, yet consensus remains challenging due to diverse hardware and software ecosystems.

- **Explainability and transparency**: Particularly with unsupervised learning algorithms, lack of interpretability can hinder trust and accountability. This is a significant concern in mission-critical telecom applications where understanding AI decisions is essential for monitoring and compliance. Promising explainable AI (XAI) frameworks offer pathways to address some of these challenges. For example, the neuralization approach that reformulates clustering models as neural networks reveals feature importances and makes cluster assignments interpretable [**?**], thereby improving transparency and trust in autonomous network operations. However, trade-offs between explainability and model complexity are still debated.

To synthesize these points, Table 31 summarizes the key challenges, associated risks, and potential approaches to mitigation or future study directions.

In conclusion, while significant challenges remain, ongoing research and emerging frameworks hold promise for making AI-driven telecommunications networks more scalable, secure, interoperable, and interpretable. Continued investigation into these open issues is vital to realize fully autonomous next-generation networks.

11.3 Holistic Perspective and Interdisciplinary Importance

Our unique synthesis conceptualizes the interplay between AI methodologies and telecom network layers—control, orchestration, and management—highlighting how AI integration across these strata facilitates unprecedented levels of self-governance and resilience. This holistic view underscores the critical need for interdisciplinary research to address the inherent complexity of integration while carefully balancing scalability, security, and explainability requirements. By bridging diverse fields such as machine learning, network engineering, and cybersecurity, this perspective enables more robust and adaptable telecom infrastructures capable of meeting evolving demands.

11.4 Summary of Key Points

To enhance reader takeaway, we provide a concise summary of key points:

Summary of Key Points:

• This survey reviewed AI's transformative impact on telecommunication functionalities, including adaptive control, routing, and

Table 31: Summary of Key Challenges in Deploying AI for Telecommunication Networks and Potential Directions

Challenge	Discussion / Risks	Potential Directions / Trade-offs
Scalability	High computational demand delays real-time response	Model simplification; hardware acceleration; trade-off between complexity and latency
Security and Privacy	Vulnerability to adversarial attacks; data breaches	Secure model training; adversarial robustness research; privacy-preserving methods
Interoperability	Multi-vendor diversity impedes integration	Development of standards; modular AI frameworks; vendor collaboration
Explainability	Lack of transparency, especially in unsupervised models, limits trust	Utilization of XAI methods such as neuralization [**?**]; balancing interpretability and accuracy

fault management, with a focus on predictive and ensemble learning models.

• Gradient boosting methods, such as CatBoost and LightGBM, demonstrate superior performance over traditional algorithms in customer churn prediction by effectively managing class imbalance and nonlinear relationships within telecom datasets [27]. This ensemble approach reduces the need for external resampling techniques but involves higher computational costs.

• Key deployment challenges remain scalability, security and privacy concerns, interoperability across heterogeneous systems, and the need for explainable models to foster trust and adoption.

• Explainable AI frameworks, including the novel neuralization of clustering models that translate cluster assignments into feature attributions, enhance transparency and trust in autonomous network operations [**?**].

• The envisioned future encompasses fully autonomous, self-optimizing networks requiring interdisciplinary research to overcome integration obstacles and to blend AI advances with telecommunications infrastructure.

This structured and explicitly stated summary consolidates the survey's main contributions while emphasizing critical enablers and research avenues necessary to advance AI-driven telecommunications.

11.5 Concluding Remarks

In summary, the progression towards next-generation telecommunication networks is increasingly reliant on advanced AI methodologies characterized by autonomy, efficiency, security, and interpretability. Achieving this transformative vision requires sustained development of AI-driven frameworks capable of robust, scalable, and transparent operation to meet the demanding performance and reliability criteria of future communication infrastructures.

References

[1] S. Aboagye, M.-S. Alouini, and L. Dai. 2024. Multi-Band Wireless Communication Networks: Fundamentals, Challenges, and Resource Allocation. *IEEE Wireless Communications* 31, 5 (2024), 86–93. https://ieeexplore.ieee.org/document/10438479/

[2] A. Ahmed, T. M. Nguyen, and M. Elsayed. 2023. Deep Learning for Telecom Self-Optimized Networks. *IEEE Transactions on Communications* 71, 4 (2023), 2001–2014. https://ieeexplore.ieee.org/document/10811884

[3] Anonymous. 2025. Deep Learning in Wireless Communication Receiver: A Survey. arXiv preprint arXiv:2501.17184. https://arxiv.org/abs/2501.17184 Accessed: 2024-06-01.

[4] M. W. Baidas. 2016. A Distributed Political Coalition Formation Framework for Multi-Relay Selection in Wireless Networks. *Wireless Communications and Mobile Computing* 16, 4 (2016), 2065–2082. doi:10.1002/wcm.2763

[5] Dimitris Bertsimas. 2023. Global optimization via optimal decision trees. *Journal of Global Optimization* 85, 1 (2023), 1–28. doi:10.1007/s10898-023-01311-x

[6] T. Chen, M. Hong, and Z. Su. 2018. Learn-and-Adapt Stochastic Dual Gradients for Network Optimization. *IEEE Transactions on Control of Network Systems* 5, 4 (2018), 1456–1467. https://ieeexplore.ieee.org/document/8110688

[7] Z. Chen, M. Zhao, and X. Wang. 2024. Robust Federated Learning for Unreliable and Resource-Constrained Wireless Networks. *IEEE Transactions on Wireless Communications* 23, 8 (2024), 9793–9809. https://ieeexplore.ieee.org/document/10444714/

[8] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo. [n. d.]. Deep Learning for Wireless Communications: An Emerging Interdisciplinary Paradigm. Online. https://arxiv.org/abs/2007.05952 Submitted Jul. 2020.

[9] X. Ding, Y. Jin, and J. Liu. 2023. Obstacle-Aware Fuzzy Clustering Protocol for Wireless Sensor Networks in 3D Terrain. *International Journal of Wireless Information Networks* 30, 1 (2023), 30–41. doi:10.1007/s10776-022-00595-8

[10] T. Febrianto, J. Hou, and M. Shikh-Bahaei. 2017. Cooperative Full-Duplex Physical and MAC Layer Design in Asynchronous Cognitive Networks. *Wireless Communications and Mobile Computing* 2017 (2017), 1–14. doi:10.1155/2017/8491920

[11] W. S. Fujo, I. J. Al-Mousa, and S. A. Hamed. 2024. Customer Churn Prediction in Telecommunication Industry Using Deep Learning. *Preprints.org* 2024, 0115 (2024). https://www.preprints.org/manuscript/202403.0585/v1

[12] A. Förster, F. Macabiau, and D. Grouset. 2024. A beginner's guide to infrastructureless networking concepts. *IET Networks* 13, 1 (2024), 14–22. doi:10.1049/ntw2.12094

[13] E. Hanasusanto, D. Kuhn, and K. N. Kallas. 2016. Multistage Robust Mixed-Integer Optimization with Adaptive Partitions. *Operations Research* 64, 4 (2016), 980–998. doi:10.1287/opre.2016.1515

[14] M. Imani. 2024. Comparing Traditional Machine Learning and Advanced Gradient Boosting Techniques in Customer Churn Prediction: A Telecom Industry Case Study. *Preprints.org* 2024, 0213 (2024). https://www.preprints.org/manuscript/202403.0213/v2

[15] K. D. Irianto and R. Chandra. 2020. Partial packet in wireless networks: a review of error recovery and loss mitigation techniques. *IET Communications* 14, 15 (2020), 2396–2409. doi:10.1049/iet-com.2019.0550

[16] D. Kuhn, P. Wiesemann, and T. Georghiou. 2019. Wasserstein Distributionally Robust Optimization: Theory and Applications in Machine Learning. *Operations Research* 67, 3 (2019), 814–831. doi:10.1287/opre.2018.1804

[17] Y. H. Kwon, K. J. Han, and Y. S. Choi. 2015. Efficient network mobility support scheme for proxy mobile IPv6. *EURASIP Journal on Wireless Communications and Networking* 2015, 1 (2015), 1–14. doi:10.1186/s13638-015-0437-8

[18] M. Li, Y. Hong, and B. Chen. 2021. A Unified Analytical Framework for Optimal Control Problems in Network Systems. *IEEE Transactions on Control of Network Systems* 8, 4 (2021), 1645–1656. https://ieeexplore.ieee.org/document/9454297

[19] Y. Li, Z. Zhang, L. Wu, and X. Wang. 2022. Real-World Wireless Network Modeling and Optimization: Recent Advances and Challenges. *Chinese Journal of Electronics* 31, 2 (2022), 263–280. https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/cje.2022.00.191

[20] Y. Liu, X. Liang, and P. Zhang. 2020. Data-Importance Aware Radio Resource Allocation. *IEEE Communications Letters* 24, 9 (2020), 2046–2050. https://ieeexplore.ieee.org/document/9098940

[21] R. F. Lopes. 2013. Performance of the modulation diversity technique for - fading channels in wireless communications. *EURASIP Journal on Wireless Communications and Networking* 2013, 1 (2013), 1–12. doi:10.1186/1687-1499-2013-17

[22] Y. Luo, C. Yang, and S. Yu. 2023. Recent Advances in Optical Wireless Communications for 6G Wireless Networks. *IEEE Wireless Communications* 30, 2 (2023), 58–65. https://ieeexplore.ieee.org/document/10325445/

[23] G. A. Mapunda, R. Ramogomana, L. Marata, B. Basutli, A. S. Khan, and J. M. Chuma. 2020. Indoor Visible Light Communication: A Tutorial and Survey. *Wireless Communications and Mobile Computing* 2020 (2020), 46. doi:10.1155/2020/8881305

[24] S. Nadarajah and A. A. Ciré. 2020. Network-Based Approximate Linear Programming for Discrete Optimization. *Operations Research* 68, 6 (2020), 1767–1786. doi:10.1287/opre.2019.1953

[25] A. Nagurney. 2022. Supply chain networks, wages, and labor productivity: insights from Lagrange analysis and computations. *Journal of Global Optimization* 83, 3 (2022), 615–638. doi:10.1007/s10898-021-01084-x

[26] F. Nisar and B. A. Rehman. 2025. An efficient security framework, vulnerabilities, and defense mechanisms in LoraWAN. *Computer and Telecommunication Engineering* 3, 2 (2025), Article ID 3072. https://aber.apacsci.com/index.php/CTE/article/view/3072

[27] D. Niyato. 2023. Editorial: Fourth Quarter 2023 IEEE Communications Surveys and Tutorials. *IEEE Communications Surveys & Tutorials* 25, 4 (2023), 3456–3463. https://ieeexplore.ieee.org/document/10325334/

[28] Dusit Niyato and et al. 2021. Survey on Wireless Communications. *IEEE Communications Surveys & Tutorials* 23, 1 (2021), 1–40. https://ieeexplore.ieee.org/document/9621329/

[29] S. Pawar, L. Bommisetty, and T. G. Venkatesh. 2022. A High Capacity Preamble Sequence for Random Access in 5G IoT Networks: Design and Analysis. *International Journal of Wireless Information Networks* 30, 1 (2022), 1–15. doi:10.1007/s10776-022-00593-x

[30] Y. Qian, H. Chen, and M. Dohler. 2022. Beyond 5G Wireless Communication Technologies. *IEEE Wireless Communications* 29, 1 (2022), 166–172. https://ieeexplore.ieee.org/document/9749229/

[31] E. Shaaban. 2023. Hyperparameter Optimization and Combined Data Certainty for Customer Churn Prediction in Telecommunication Industry. *Preprints.org* 2023, 1478 (2023). https://www.preprints.org/manuscript/202308.1478/v3

[32] X. Shen, Y. Liu, X. Du, and K. K. R. Choo. 2020. AI-assisted Network-slicing based Next-generation Wireless Networks. *IEEE Transactions on Wireless Communications* 19, 3 (2020), 1558–1571. https://ieeexplore.ieee.org/iel7/8782711/8889399/08954683.pdf

[33] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis. 2016. 5G-Enabled Tactile Internet. *IEEE Journal on Selected Areas in Communications* 34, 3 (2016), 460–473. https://ieeexplore.ieee.org/document/7403840/

[34] S. Thapaliya and P. K. Sharma. 2022. Cyber Forensic Investigation in IoT Using Deep Learning Based Feature Fusion in Big Data. *International Journal of Wireless Information Networks* 30, 1 (2022), 16–29. doi:10.1007/s10776-022-00588-7

[35] D. Wen, B. Zhang, and Y. Chen. 2020. Joint Parameter-and-Bandwidth Allocation for Improving Federated Learning Performance in Wireless Networks. *IEEE Transactions on Wireless Communications* 19, 10 (2020), 6780–6793. https://ieeexplore.ieee.org/document/9194337/

[36] Z. Weng, L. Lu, J. Chen, H. Zhang, and L. Hanzo. 2023. Deep Learning Enabled Semantic Communications With Knowledge Graph and Knowledge Base. *IEEE Journal on Selected Areas in Communications* 41, 9 (2023), 2192–2207. https://ieeexplore.ieee.org/document/10038754

[37] Z. Zhao, E. J. Schiller, E. Kalogeiton, T. Braun, S. Burkhard, and M. T. Garip. 2017. Autonomic Communications in Software-Driven Networks. *IEEE Journal on Selected Areas in Communications* 35, 11 (2017), 2431–2445. https://ieeexplore.ieee.org/document/8063402/

[38] H. Zhou, W. Saad, and D. Niyato. 2024. Large Language Model (LLM) for Telecommunications: A Comprehensive Survey on Principles, Key Techniques, and Opportunities. *IEEE Communications Surveys & Tutorials* 26, 2 (2024), 879–913. https://ieeexplore.ieee.org/document/10685369/

[39] D. D. Čvokić, Y. A. Kochetov, and A. Savić. 2022. A variable neighborhood search algorithm for the (r|p) hub–centroid problem under the price war. *Journal of Global Optimization* 83, 3 (2022), 405–444. doi:10.1007/s10898-021-01051-2