This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version XXX.

Deep Learning-Enabled CSI Estimation and Detection in Modern Multi-Antenna Systems

MUHAMMAD UMER ZIA [1] *(Member, IEEE)*, WEI XIANG [2] *(Senior Member, IEEE)*, TAO HUANG [1,3] *(Senior Member, IEEE)*, JAMEEL AHMAD [4] *(Member, IEEE)*, JAWWAD N. CHATTHA [5] *(Senior Member, IEEE)*, RANA ABBAS [6] *(Member, IEEE)*, AND ASAD MAHMOOD [1] *(Member, IEEE)*

[1]College of Science and Engineering, James Cook University, Smithfield, QLD, Australia
[2]School of Computing, Engineering and Mathematical Sciences, La Trobe University, Melbourne, VIC 3086, Australia
[3]Centre for AI and Data Science Innovation, James Cook University, Smithfield, QLD, Australia
[4]Department of Computer Science, School of Systems and Technology, University of Management and Technology Lahore, Pakistan
[5]Department of Computer Arts and Technology, Norwich University of the Arts, Norwich NR24SN, UK
[6]Transport for NSW, Sydney, NSW, Australia

CORRESPONDING AUTHOR: Wei Xiang (e-mail: w.xiang@latrobe.edu.au).

ABSTRACT The remarkable success of deep learning (DL) techniques has established them as a promising approach for channel estimation and symbol detection in 6th-generation (6G) multi-antenna systems. These systems, including massive multiple-input multiple-output (mMIMO), reconfigurable intelligent surfaces (RIS)-supported mMIMO, and their advanced variants, pose diverse challenges and dependencies, in which estimation inaccuracies can significantly degrade performance. Traditional methods, such as least-squares (LS) and minimum mean-square error (MMSE) estimators, exhibit significant limitations, including the requirement for a channel covariance matrix and matrix inversion. Moreover, conventional techniques often achieve lower accuracy than DL-based methods, which can learn complex channel mappings, an advantage that makes DL-based methods promising for 6G systems.

In the realm of symbol detection, 6G multi-antenna systems face several challenges, including high computational complexity and the escalating difficulty of managing signal interference. These factors can significantly degrade the performance of traditional detection methods, particularly in environments characterized by multiple interfering sources and noise. The data-driven and model-driven DL-aided estimation and detection methods have unlocked a new research paradigm that is now under rigorous investigation by academia and industry. In this paper, we provide a comprehensive review of recent advancements in DL-aided channel estimation and detection for multi-antenna systems, with a specific focus on massive MIMO and RIS-enabled massive MIMO systems. We conclude by discussing the challenges inherent in DL-assisted channel estimation and detection, and propose potential future directions to improve efficiency and accuracy.

Index Terms Deep Learning, Channel Estimation, Detection, Massive MIMO, RIS Massive MIMO.

I. Introduction

The rapid development of 6th-Generation (6G) wireless technology is propelling the use of massive multi-antenna systems to enable high data rates and reliable communications [1]–[3]. 6G technology is expected to operate at higher frequencies, such as mmWave/sub-THz band and utilize a multitude of diverse technologies, such as massive multiple-input multiple-output (mMIMO), extremely large MIMO (XL-MIMO), reconfigurable intelligent surfaces (RIS), and cell-free mMIMO (CF-mMIMO) [4]. There are significant challenges in implementing 6G technology, including channel estimation for massive multi-antenna systems and high-accuracy signal detection. These challenges are associated with the huge number of antennas and their corresponding channels, as well as signal attenuation at higher frequencies [5]. Advances in deep learning (DL) techniques have opened new avenues for improving estimation and signal-detection accuracy in 6G multi-antenna communications systems [6], [7].

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

Table 1: Literature Comparison, portion of the filled circle represents DL-aided design discussion

Massive MIMO Estimation	Massive MIMO Detection	RIS Massive MIMO Estimation	RIS Massive MIMO Detection
[8] ◴	[9] ◴	[10] ◐	[11] ○
[12] ◐	[13] ◴	[7] ◴	[14] ○
[15] ○	[16] ○	[17] ◐	[18] ○
[19] ○	[20] ○	[21] ◴	[22] ○
‘ [23] ◐	[24] ◴	[25] ◴	[26] ○
[27] ◐	[28] ◐	[29] ○	[30] ○
Ours ●	Ours ●	Ours ●	Ours ●

By learning complex relationships among signals, deep learning has partnered with conventional signal processing models to address estimation and detection challenges in complex and noisy environments [31].

In massive MIMO systems, accurate channel estimation is crucial for achieving the benefits of beamforming and spatial multiplexing. However, estimating the channel accurately is a dual challenge because channel estimation for a massive array of antennas leads to high signal acquisition costs/training overheads, and a higher capacity feedback link is required to send this channel state information (CSI) to the base station in frequency division duplex (FDD) mode. Transmission in the FDD scenario also suffers from further distortions due to feedback-link noise, quantized CSI, and delay [32]–[34]. In comparison, the time-division duplex (TDD) mode, which relies on uplink channel information, incurs only increased channel estimation complexity due to the large number of antennas and pilot contamination [35].

RIS-aided massive MIMO systems significantly improve signal quality by providing an alternative path that typically has fewer obstacles, thereby reducing attenuation and scattering. RIS surfaces comprise many small, independently controllable, quasi-passive elements. These elements can reflect incoming signals in a desired direction by adjusting the phase shifts of individual elements, thereby enabling control over the high-attenuation wireless propagation channels. In scenarios where direct base-station-to-user link exhibits lower attenuation, the RIS mMIMO system can benefit from the diversity of RIS-assisted links. However, all these advantages depend upon the availability of high-quality CSI regarding direct and RIS-assisted links to optimize RIS phase shifts at the base station [36], [37]. RIS mMIMO channels exhibit several significant physical features, including spatial, spectral, and temporal correlations, and expert knowledge of channel models and their properties is essential for creating deep-learning training datasets. Appropriately designed DL algorithms and architectures trained on channel datasets can certainly improve CSI accuracy; however, the lack of trained DL models for accurate CSI estimation in RIS mMIMO remains a technical limitation [21]. Acknowledging the promising nature of this research domain, several researchers have shown that DL techniques can outperform conventional channel estimation in RIS mMIMO systems by a fair margin [37]–[41].

Deep learning-based detection in massive MIMO and RIS-aided massive MIMO systems is an emerging research area. Conventional detectors, such as maximum likelihood (ML), are optimal but computationally intensive. Hence, near-optimal detectors that could provide decent performance with low complexity are favored by researchers [9]. Linear detectors, such as zero-forcing (ZF) and linear minimum mean-squared error (LMMSE), are low-complexity and hence preferred in massive MIMO systems. However, determining the coefficients for these detectors typically requires matrix inversion, which complicates the design. DL-aided design can avoid matrix inversion complexity given that an appropriate amount of data is available to train DL models [42].

Although channel estimation and symbol detection are distinct domains in communication systems, they are closely interdependent; the accuracy of channel estimation directly affects detection performance, and both collectively determine the overall performance of multi-antenna systems. Moreover, with the advent of deep learning techniques, these two tasks have become more closely intertwined. One key reason for this integration is that both channel estimation and symbol detection share the goal of accurately recovering the transmitted information in the presence of noise, interference, and distortions. In addition, the complexity of estimation and detection in multi-antenna systems increases due to the high-dimensional signal space, making traditional methods less effective [43]. The objective of this research is to provide a comprehensive and up-to-date survey of advances in massive multi-antenna systems, with a particular focus on deep learning-based approaches for enhancing channel estimation and symbol detection. As highlighted in Table 1, existing surveys typically address estimation and detection in isolation, with many dedicating only a section to DL techniques. Moreover, those that focus exclusively on deep learning lack state-of-the-art DL models, such as transformer and large language model (LLM)-based channel predictors, U-Net-based channel denoisers, and the latest DL-driven detection methods.

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

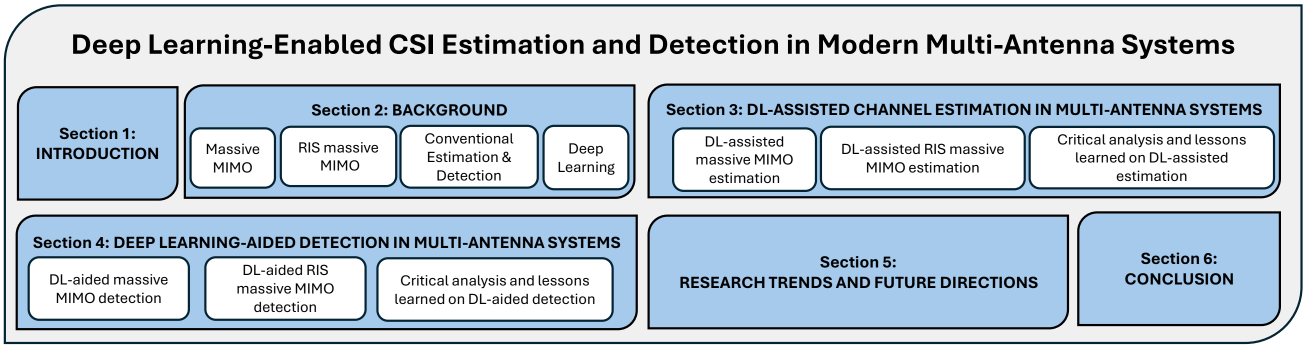

Figure 1: Sections breakdown of the paper

The time-relevance and imperative nature of this survey are due to the growing diversity of DL models applied to various channel types, which has made it increasingly challenging to select the most appropriate DL model for a specific channel while maintaining manageable computational complexity. Unlike previous surveys that neglect a holistic view of this domain, our work provides crucial guidance to the research community on cutting-edge DL techniques that address channel-related challenges, such as high-dimensional signal spaces, interference, and noise in multiantenna systems. In addition, this investigation provides researchers with valuable insights into trends in the selection of data-driven and model-driven DL techniques.

In summary, the investigation aims to bridge these gaps by presenting a unified and cutting-edge perspective on DL-enhanced methodologies for both channel estimation and symbol detection. To provide a clear overview of the scope of our investigation, the paper's structure is outlined in Fig. 1. To this end, the key contributions of this article are highlighted as follows:

1) This survey presents an analytical framework for deep learning-aided channel estimation and symbol detection techniques in modern multi-antenna systems. Unlike previous surveys that primarily discuss conventional estimation and detection approaches with only marginal attention to DL techniques, this survey provides a comprehensive overview of DL constructs, highlighting their dependencies and limitations while emphasizing the need for global solutions. By grouping DL constructs based on their functions, we identify synergies across different DL approaches and analyze the computational complexity and performance trade-offs.
2) A critical analysis of state-of-the-art data-driven and model-driven DL techniques is provided for channel estimation and symbol detection, focusing on the system model, data generation, and assumptions. This analysis identifies design choices that influence the efficiency and scalability of DL models for channel estimation and symbol detection. Furthermore, recent advances in DL constructs, such as federated learning variants, U-Net denoisers, diffusion model-based estimators, and large language model (LLM)-supported approaches, are examined for their potential to address the unique challenges of 6G systems, including high-dimensional signal spaces, interference, and noise. The paper provides an analytical breakdown of the pros and cons of various DL constructs, shedding light on the complexities of their practical implementation.
3) This survey reviews emerging trends and future research directions in DL-aided multi-antenna systems, focusing on innovations in channel estimation and symbol detection. It highlights key challenges in current methods and proposes solutions, emphasizing the role of DL in next-generation systems. The paper explores DL prospects in satellite-based RIS systems, Channel Foundation Models for improved predictions, and Quantum Machine Learning for efficient signal processing. Additionally, it discusses challenges such as adversarial attacks, pixel errors, and the lack of standardization, stressing the need to resolve these issues to enable the real-world deployment of DL-enabled systems.

II. Background

In this section, we present the state-of-the-art in massive MIMO and RIS massive MIMO systems, as well as the relevant constructs and concepts related to deep learning techniques for CSI estimation and symbol detection.

A. Massive MIMO

Massive MIMO technology has become a cornerstone of modern wireless communication systems. This technology, along with its variants, continues to play a pivotal role in enhancing the capacity, reliability, and spectral efficiency of 5G networks, and serves as a foundational technology for future 6G systems [44]. The core principle of massive MIMO is to equip base stations with a large number of antennas to serve multiple users via spatial multiplexing

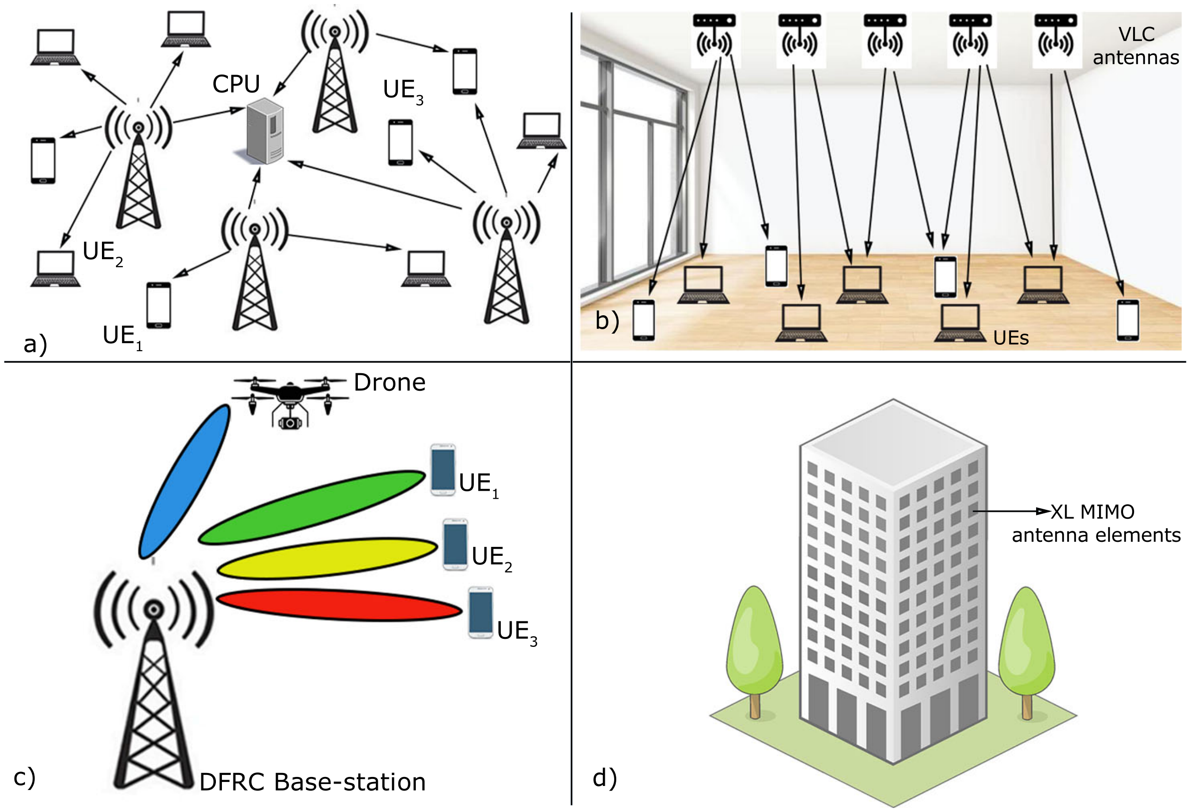

Figure 2: State-of-the-art Massive MIMO Types: a) Cell-free massive MIMO, b) VLC massive MIMO, c) ISAC massive MIMO, and d) XL MIMO.

simultaneously. Conventional massive MIMO often faces challenges, including inter-cell interference and performance degradation at the edges of coverage. To overcome these limitations, innovative variants of massive MIMO have been proposed, including cell-free massive MIMO, which abandons the traditional cell structure in favor of a distributed architecture comprising numerous geographically dispersed access points. These access points collaboratively serve users using coherent joint transmission, offering significant improvements in uniform spectral efficiency and coverage reliability. In addition, massive MIMO has evolved to address a broader set of emerging use cases envisioned for 6G networks. These paradigms include integrated sensing and communication (ISAC), visible light communication (VLC), massive MIMO, and XL-MIMO systems [45]. Each of these extensions introduces new capabilities to the massive MIMO framework. Fig. 2 showcases several state-of-the-art massive MIMO systems, along with their key entities.

VLC massive MIMO leverages the unlicensed visible-light spectrum to provide high-speed data transmission with inherent physical-layer security and reduced RF interference. VLC massive MIMO aims to address the challenge of RF congestion in indoor environments. When combined with MIMO and intelligent beamforming techniques, VLC systems offer compelling solutions for mobility-aware applications and ultra-reliable indoor coverage [46], [47].

The XL-MIMO systems utilize large antenna arrays spanning building facades or infrastructure surfaces, providing enhanced spatial resolution and unprecedented array gains. Recent research has highlighted XL-MIMO's potential to improve signal quality and enable high-throughput services such as holographic communication and immersive augmented reality (AR). However, challenges such as hardware complexity and power consumption remain. Despite these hurdles, XL-MIMO is a promising technology, especially in challenging mmWave and THz bands, offering scalable solutions for high-capacity, dense environments [48], [49].

ISAC massive MIMO systems are developed to overcome the inefficiency of treating communication and sensing as separate systems with dedicated spectrum, hardware, and signal processing frameworks [50]. Its motivation lies in achieving higher spectrum efficiency, lower hardware complexity, reduced cost and energy consumption, and effective support for emerging applications such as autonomous driving, smart cities, and industrial IoT, where real-time sensing and communication are crucial. For instance, in autonomous driving, ISAC enables vehicles to simultaneously communicate with infrastructure to improve real-time situational awareness, an essential feature for autonomous systems and smart infrastructure [51]. In these systems, the large antenna array is used not only for data transmission but also for sensing tasks such as detection, positioning, and tracking, with narrower, more precise beams for both communication and sensing. As a result, the ISAC massive MIMO systems introduce distinct system models and design requirements, including dual-functional beamforming, waveform co-design, interference management between sensing and communication signals, distributed sensing modes, and near-field effects, which are not central concerns in conventional communication-only MIMO systems [50]. In addition, ISAC systems deployment faces significant challenges, including integrating sensing and communication functionalities without causing interference, maintaining high accuracy in environmental sensing, and managing the computational complexity of joint signal processing [52].

Cell-free massive MIMO addresses the need for improved coverage and uniform spectral efficiency in dense urban environments by eliminating the limitations of traditional cell-based architectures. However, deploying cell-free massive MIMO faces challenges related to synchronization, coordination among distributed access points, and increased signal-processing complexity, particularly in channel estimation and resource allocation [53], [54]. Another key area of development involves joint communication and sensing within cell-free massive MIMO frameworks. Scalable architectures have been proposed to enable simultaneous multi-user communication and multi-target detection, highlighting the potential of these distributed systems in supporting ISAC capabilities [51], [55]. Novel implementations also demonstrate how distributed access points in cell-free settings can concurrently execute communication and multi-target tracking, leveraging spherical or adaptive beamforming techniques [55].

To address the escalating requirements for data throughput and spatial resolution, researchers have introduced ultra-massive MIMO (UM-MIMO) and gigantic MIMO (gMIMO) paradigms [56], [57]. These approaches involve deploying antenna arrays with hundreds to thousands of elements, optimized for millimetre-wave and terahertz frequency bands. Such configurations exploit high-dimensional beamforming to harness wide bandwidths and enable dense spatial multiplexing [58].

A transformative concept within the massive MIMO ecosystem is user equipment-centric collaborative MIMO (UE-CoMIMO) [59]. Unlike traditional base-station-driven coordination, UE-CoMIMO enables mobile devices to form virtual antenna arrays through peer-to-peer cooperation dynamically. This distributed model improves spectrum reuse, improves robustness against central point failures, and aligns with the vision of intelligent and decentralized edge networks [60].

Simultaneously, growing concerns about sustainability have led to the emergence of green MIMO technologies [61]. These energy-aware designs incorporate low-power RF circuitry, dynamic resource management, and hardware-level optimizations. Industry efforts, such as Ericsson's 2025 launch of energy-efficient antenna platforms, underscore the commercial viability of such approaches for supporting programmable, sustainable 6G infrastructure [62].

Collectively, these advanced extensions, including cell-free architectures, ultra-dense antenna arrays, ISAC integration, UE-CoMIMO, energy-efficient designs, and VLC-based systems, represent a significant paradigm shift in the evolution of massive MIMO technology. As these innovations continue to mature, they are poised to meet the multifaceted demands of 6G networks, including terabit-level data rates, ultra-reliable low-latency communication (URLLC), centimetre-level positioning, enhanced energy efficiency, and pervasive connectivity in dense, heterogeneous, and highly dynamic environments [63].

B. RIS massive MIMO

Reconfigurable intelligent surfaces have emerged as a pivotal technology for next-generation wireless networks, offering control over the propagation medium through dynamically reconfigurable metasurfaces. These surfaces consist of densely packed, sub-wavelength unit cells whose electromagnetic response can be dynamically tuned to reshape the radio environment. By intelligently manipulating wavefronts, RIS enhances spectral efficiency, energy efficiency, and spatial diversity, making it a foundational enabler for 6G communications [44], [55].

RIS implementations are broadly classified as passive, active, and hybrid. Passive RIS employs low-power components, such as PIN diodes, to control phase shifts without signal amplification, offering a scalable, energy-efficient solution suitable for large-area deployments [44]. In contrast, active RIS integrates amplifiers to compensate for signal attenuation, achieving extended coverage at the cost of higher energy consumption and increased system complexity [64]. Hybrid RIS systems aim to balance these trade-offs by selectively incorporating active elements into an otherwise passive array.

From a structural perspective, RIS can adopt various geometries, including planar, conformal, and spherical configurations. Functionally, they are categorized as reflective, transmissive, or waveguide-fed systems based on their interaction with incident signals [44], [65]. A notable innovation in this domain is the simultaneous transmission and reflection RIS (STAR-RIS), which introduces unit cells that can concurrently reflect and transmit signals. While it splits signals into transmitted and reflected components, it does not amplify the signal. However, a later development known as active STAR-RIS includes reflection-type amplifiers [66]. This dual-mode functionality enables flexible beamforming across multiple coverage zones, significantly improving system performance in scenarios with high user density or physical obstructions [47]. Another advanced RIS architecture is the spherical RIS [55], which has attracted researchers due to its ability to overcome spatial ambiguity and enhance angular resolu-

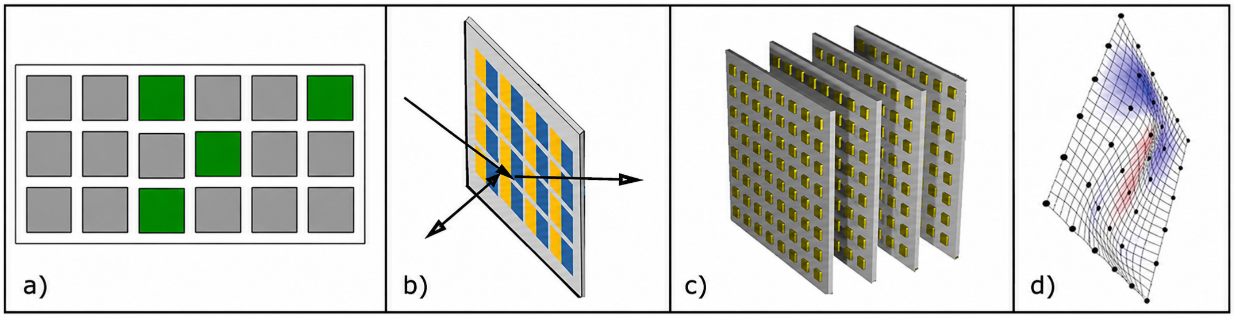

Figure 3: RIS types and architectures: a) Hybrid RIS, where grey and green colors represent passive and active elements, respectively. b) Active STAR RIS, with signal reflection and transmission of signal on the other side, is shown. The active elements are shown in blue colour, capable of amplifying the re-transmitted signal. c) Stacked RIS, where multiple layers of RIS elements are arranged vertically to improve spatial diversity. d) Flexible RIS, designed with a deformable or pliable structure.

tion, two limitations commonly encountered in conventional planar surfaces. These developments collectively underscore the potential of RIS to redefine the wireless landscape by transforming the physical medium into a programmable entity. Notable RIS structural types are shown in Fig. 3.

In the domain of combining advanced RIS systems with massive MIMO, Lin et al. [55] proposed a spherical RIS system for ISAC, employing unit cells arranged on a three-dimensional spherical surface to enable high-precision channel estimation and user self-localization. The spherical geometry, coupled with spherical harmonics and phase-mode excitation, provides full angular coverage, thereby improving the accuracy of extracted channel parameters, such as azimuth, elevation, delay, and gain. This enables unique decoupling of signal paths, which is critical for applications in localization, tracking, and ambient sensing.

Additionally, Stacked RIS, another emerging design, comprises multiple vertically stacked RIS layers, each independently tunable to enhance wave manipulation via cumulative phase control. This architecture enhances beamforming, extends communication range, and improves robustness in dynamic or obstructed environments [67]. Alongside these innovations, research is progressing on graphene-based RIS for THz applications [68], transparent RIS for integration into glass surfaces [69], and AI-controlled RIS for adaptive, real-time optimization [70]. These advancements collectively indicate a shift from static RIS configurations toward intelligent, multi-functional surfaces that are integral to the vision of 6G wireless systems supporting ubiquitous, context-aware connectivity.

C. Conventional Estimation & Detection

This subsection reviews conventional channel estimation and detection techniques together with their underlying mathematical models. Several methods have been proposed for CSI estimation in high-frequency and wideband wireless systems, including ultra-wideband (UWB) [71] and millimeter-wave (mmWave) communications [72]. These methods primarily aim to address the challenges associated with high-dimensional CSI acquisition and the resulting feedback overhead [73]. Similarly, in the area of detector design, numerous conventional schemes have been developed, and several comprehensive surveys have discussed their principles, performance, and limitations [13], [74], [75]. To maintain consistency with the scope of this survey, only the two most commonly used conventional channel estimation and detection methods are briefly reviewed.

We consider uplink transmission, where a mobile user sends pilot and data symbols to a BS equipped with M antennas. Let N_u denote the number of transmitted spatial streams, with the single-antenna mobile station. Let τ_p denote the pilot length. The operators $(\cdot)^H$, $(\cdot)^T$, $\mathrm{vec}(\cdot)$, $\otimes$, and $\|\cdot\|_F$ denote the Hermitian transpose, transpose, vectorization operator, Kronecker product, and Frobenius norm, respectively.

1) Estimation in massive MIMO systems

We consider a massive MIMO system employing uplink pilot transmission in a time-division duplexing framework, consistent with the predominant focus of multi-antenna research on TDD-based transmission. In addition, we assume pilot transmission occurs within coherence time, with fixed RIS phase shifts. In that case, the BS receives the transmitted pilot signal, which can be written as

$$\mathbf{Y}_p = \mathbf{H}\mathbf{X}_p + \mathbf{N}_p, \tag{1}$$

where $\mathbf{Y}_p \in \mathbb{C}^{M\times\tau_p}$ is the received pilot matrix, $\mathbf{H} \in \mathbb{C}^{M\times N_u}$ is the uplink channel matrix, $\mathbf{X}_p \in \mathbb{C}^{N_u\times\tau_p}$ is the known pilot matrix, and $\mathbf{N}_p \in \mathbb{C}^{M\times\tau_p}$ denotes additive white Gaussian noise. Under the usual block-fading assumption, $\mathbf{H}$ remains constant over one coherence interval. The objective of the pilot phase is to estimate $\mathbf{H}$ at the BS, after which the obtained CSI is used for coherent uplink detection.

The least-squares (LS) estimator is the simplest and most widely used benchmark, as it relies only on the known pilot sequence and does not require prior channel statistics. It is obtained as

$$\hat{\mathbf{H}}_{\mathrm{LS}} = \arg\min_{\mathbf{H}} \|\mathbf{Y}_p - \mathbf{H}\mathbf{X}_p\|_F^2 . \tag{2}$$

Assuming that $\mathbf{X}_p\mathbf{X}_p^H$ is nonsingular, the LS solution is [15], [19]

$$\hat{\mathbf{H}}_{\mathrm{LS}} = \mathbf{Y}_p\mathbf{X}_p^H \left(\mathbf{X}_p\mathbf{X}_p^H\right)^{-1} . \tag{3}$$

Although LS is attractive because of its simplicity, its performance can degrade severely in low-SNR conditions.

A more accurate conventional estimator is the linear minimum mean-square error estimator, which exploits second-order channel statistics. By vectorizing (1), the pilot observation can be written as

$$\mathbf{y}_p = \mathrm{vec}(\mathbf{Y}_p) = \left(\mathbf{X}_p^T \otimes \mathbf{I}_M\right)\mathbf{h} + \mathbf{n}_p, \tag{4}$$

where

$$\mathbf{h} = \mathrm{vec}(\mathbf{H}), \quad \mathbf{y}_p = \mathrm{vec}(\mathbf{Y}_p), \quad \mathbf{n}_p = \mathrm{vec}(\mathbf{N}_p). \tag{5}$$

Defining

$$\mathbf{A} = \mathbf{X}_p^T \otimes \mathbf{I}_M, \tag{6}$$

the LMMSE estimate is given by [15], [19]

$$\hat{\mathbf{h}}_{\mathrm{LMMSE}} = \mathbf{R}_h\mathbf{A}^H \left(\mathbf{A}\mathbf{R}_h\mathbf{A}^H + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{y}_p, \tag{7}$$

where $\mathbf{R}_h = \mathbb{E}[\mathbf{h}\mathbf{h}^H]$ denotes the channel covariance matrix and σ_n^2 is the noise variance. The estimate $\hat{\mathbf{H}}_{\mathrm{LMMSE}}$ is obtained by reshaping $\hat{\mathbf{h}}_{\mathrm{LMMSE}}$ to dimension $M \times N_u$. Compared with LS, LMMSE generally provides better estimation accuracy, particularly at low and moderate SNR. This gain, however, comes at the expense of covariance knowledge and increased computational complexity due to matrix inversion.

2) Detection in massive MIMO systems

Following channel estimation, the uplink data model is expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{8}$$

where $\mathbf{y} \in \mathbb{C}^{M\times 1}$ is the received data vector, $\mathbf{s} \in \mathbb{C}^{N_u\times 1}$ is the transmitted symbol vector drawn from the modulation alphabet $\mathcal{S}$, and $\mathbf{n} \in \mathbb{C}^{M\times 1}$ is AWGN. In practice, $\mathbf{H}$ is replaced by its estimate $\hat{\mathbf{H}}$. The task of the detector is then to recover $\mathbf{s}$ from $\mathbf{y}$ using $\hat{\mathbf{H}}$. The zero-forcing (ZF) detector is a standard linear detector that removes inter-stream interference through linear inversion. Its combining matrix is [16], [20]

$$\mathbf{W}_{\mathrm{ZF}} = \left(\hat{\mathbf{H}}^H\hat{\mathbf{H}}\right)^{-1}\hat{\mathbf{H}}^H, \tag{9}$$

and the detected symbol vector is

$$\hat{\mathbf{s}}_{\mathrm{ZF}} = \mathbf{W}_{\mathrm{ZF}}\mathbf{y}. \tag{10}$$

ZF is widely used because of its analytical simplicity, but it may suffer from significant noise enhancement when the channel Gram matrix is ill-conditioned.

A more robust linear detector is the LMMSE detector, which accounts for both interference and noise. Its combining matrix is given by [16], [20]

$$\mathbf{W}_{\mathrm{LMMSE}} = \left(\hat{\mathbf{H}}^H\hat{\mathbf{H}} + \sigma_n^2\mathbf{I}_{N_u}\right)^{-1}\hat{\mathbf{H}}^H, \tag{11}$$

and the corresponding symbol estimate is

$$\hat{\mathbf{s}}_{\mathrm{LMMSE}} = \mathbf{W}_{\mathrm{LMMSE}}\mathbf{y}. \tag{12}$$

Unlike ZF, the LMMSE detector explicitly accounts for the noise variance and therefore offers a better trade-off between interference suppression and noise amplification.

For performance benchmarking, the maximum-likelihood (ML) detector is often adopted as a reference:

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\min_{\mathbf{s}\in\mathcal{S}^{N_u}} \left\|\mathbf{y} - \hat{\mathbf{H}}\mathbf{s}\right\|_2^2 . \tag{13}$$

The ML detector achieves optimal detection performance over the discrete search space $\mathcal{S}^{N_u}$, but its complexity grows exponentially with the modulation order and the number of transmitted streams. For this reason, linear detectors such as ZF and LMMSE remain the more common practical baselines in massive MIMO systems.

3) Estimation in RIS-massive MIMO

For RIS-assisted massive MIMO, the uplink model becomes more complex because the BS receives both the direct signal from the mobile station to the BS and the RIS-reflected signal. Let N denote the number of RIS elements. Let $\mathbf{H}_{\mathrm{d}} \in \mathbb{C}^{M\times N_u}$ denote the direct MS to BS channel, $\mathbf{H}_{\mathrm{BR}} \in \mathbb{C}^{M\times N}$ the RIS-BS channel, and $\mathbf{H}_{\mathrm{UR}} \in \mathbb{C}^{N\times N_u}$ the MS-RIS channel. For the qth RIS configuration, the RIS phase-shift matrix is

$$\mathbf{\Phi}^{(q)} = \mathrm{diag}\left(\beta_1 e^{j\phi_1^{(q)}}, \beta_2 e^{j\phi_2^{(q)}}, \ldots, \beta_N e^{j\phi_N^{(q)}}\right), \tag{14}$$

where β_n and $\phi_n^{(q)}$ denote the amplitude coefficient and phase shift of the nth RIS element, respectively. The received pilot matrix at the BS is then expressed as [25], [29]

$$\mathbf{Y}_p^{(q)} = \left(\mathbf{H}_{\mathrm{d}} + \mathbf{H}_{\mathrm{BR}}\mathbf{\Phi}^{(q)}\mathbf{H}_{\mathrm{UR}}\right)\mathbf{X}_p + \mathbf{N}_p^{(q)}, \tag{15}$$

where $\mathbf{N}_p^{(q)}$ denotes the noise matrix corresponding to the qth RIS configuration. Defining the equivalent channel

$$\mathbf{H}_{\mathrm{eq}}^{(q)} = \mathbf{H}_{\mathrm{d}} + \mathbf{H}_{\mathrm{BR}}\mathbf{\Phi}^{(q)}\mathbf{H}_{\mathrm{UR}}, \tag{16}$$

(15) can be rewritten compactly as

$$\mathbf{Y}_p^{(q)} = \mathbf{H}_{\mathrm{eq}}^{(q)}\mathbf{X}_p + \mathbf{N}_p^{(q)}. \tag{17}$$

Accordingly, LS- and LMMSE-type estimators may still be applied by treating $\mathbf{H}_{\mathrm{eq}}^{(q)}$ as the effective uplink channel, typically after stacking multiple observations corresponding to different RIS states. However, unlike conventional massive MIMO, RIS-assisted estimation must jointly account for the direct and reflected channels, thereby substantially increasing the estimation dimensionality and overhead.

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

4) Detection in RIS- massive MIMO

During uplink data transmission, the received signal at the BS under the qth RIS configuration is given by

$$\mathbf{y}^{(q)} = \mathbf{H}_{\mathrm{eq}}^{(q)}\mathbf{s} + \mathbf{n}^{(q)}, \tag{18}$$

where $\mathbf{n}^{(q)}$ is AWGN. Once $\hat{\mathbf{H}}_{\mathrm{eq}}^{(q)}$ is obtained, conventional detectors such as ZF, LMMSE, and ML can be directly employed by replacing $\hat{\mathbf{H}}$ in (9)–(13) with $\hat{\mathbf{H}}_{\mathrm{eq}}^{(q)}$. Although the detector structures remain formally unchanged, RIS-assisted systems are considerably more challenging because of the cascaded channel structure, RIS-state dependence, and increased problem dimensionality. In summary, these methods remain the primary analytical baselines for evaluating modern learning-based transceivers.

D. Deep Learning

In recent years, deep learning techniques have attracted considerable attention due to their transformative impact across diverse domains such as computer vision [76], medical imaging [77], connected and autonomous vehicles (CAVs) and unmanned aerial vehicles (UAVs) [78], generative AI, and natural language processing (NLP) [79], while also showing strong potential to advance signal processing in large antenna arrays [80]. The results of deep learning-aided designs have shown increases in spectral efficiency, capacity, and robustness against communication impairments in multi-antenna systems. Deep learning can be integrated into communication systems either as a standalone processing module or in conjunction with conventional signal processing blocks, particularly for estimation and detection. The latter paradigm, commonly referred to as model-driven deep learning, embeds domain knowledge and analytical models into the learning architecture, enabling a principled fusion of model-based interpretability and data-driven adaptability. Fig. 4 illustrates a generic model-driven deep learning framework, where the conventional estimation & detection can either be utilized as input to DL frameworks for further refining or can be combined at the output of DL constructs for improved results.

Based on our extensive literature evaluation on massive MIMO and deep learning [8], [12], [23], [27], the majority of existing works rely on analytically generated channel models, where the channel is expressed in closed-form and treated as perfect CSI, serving as the ground truth. In these approaches, the received pilot signals are typically considered as imperfect observations, or conventional estimators such as LS or MMSE are used to obtain an initial estimate. Subsequently, deep learning models are trained to learn a mapping from these imperfect estimates to the assumed ground-truth CSI. While this framework has shown promising performance, it inherently depends on the accuracy of the assumed channel models and may not fully capture real-world impairments. In contrast, only a limited number of studies incorporate real-world channel measurements obtained from hardware testbeds, which are crucial for capturing practical propagation effects and hardware non-idealities [81], [82]. In these cases, a noisier version is treated as an imperfect observation.

From a practical deployment perspective of DL constructs in different environments, the nature of devices, their processing, and computational capabilities must be considered. For example, the DL constructs for IoT systems must operate under strict constraints on memory, computation, latency, and energy consumption. The typical IoT devices have limited processing capability, making direct deployment of computationally intensive deep learning models impractical [83]. Accordingly, hardware-aware design is essential, where lightweight inference is executed at the edge or gateway, while more complex processing is selectively offloaded to the cloud [84], [85]. This edge-cloud partitioning improves feasibility and supports low-latency response for time-sensitive applications. In addition, compression, feature reduction, quantization, and edge-oriented frameworks should be considered to ensure that inference remains within the computational capacity of the hardware [84]. Thus, deployment-oriented evaluation should extend beyond predictive accuracy to include latency, memory footprint, and runtime resource usage [85]. The same considerations hold for portable or legacy devices with limited processing

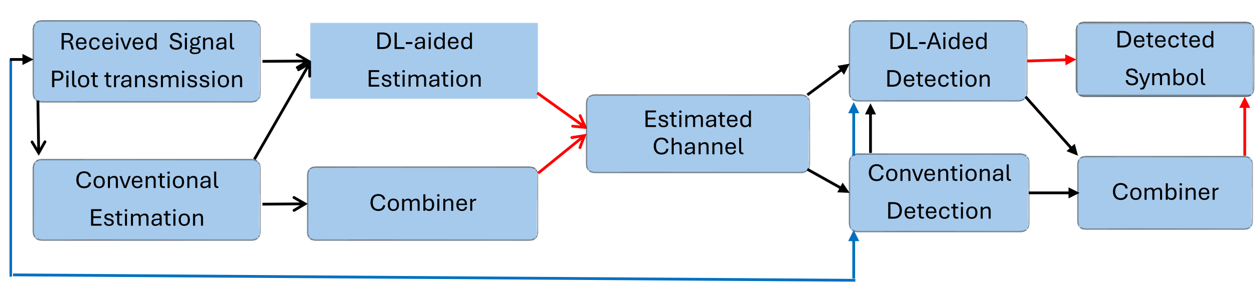

Figure 4: Model-driven deep learning: Red arrows show two ways through which deep learning can be integrated with a conventional estimation and detection pipeline.

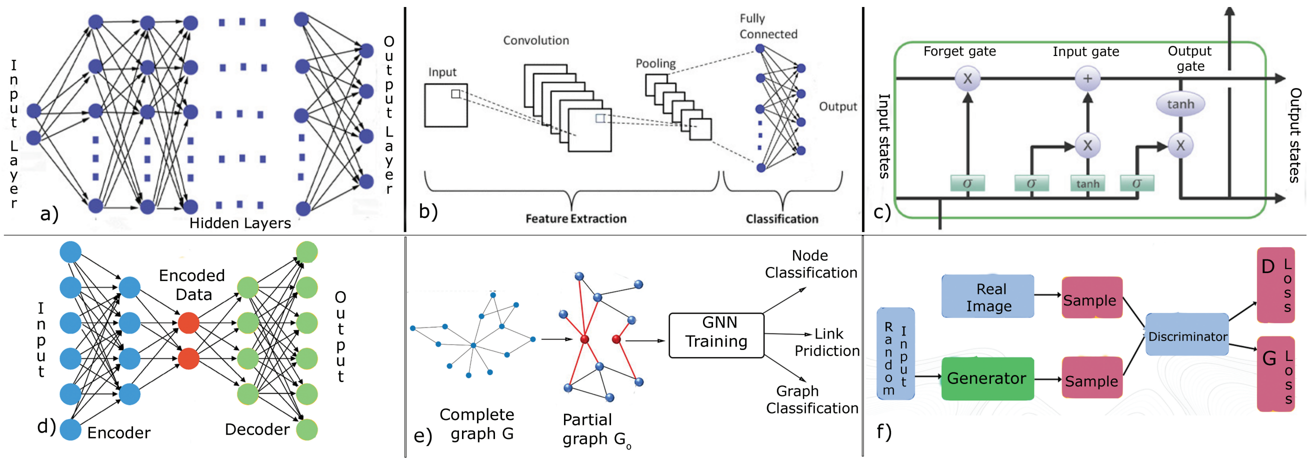

Figure 5: Deep Learning Models: a) Multilayer perceptron, dots represent the variable number of layers and neurons b) Convolutional neural network c) Long short-term memory, where $+$ and $\times$ represent element-wise summation and multiplication, σ and $tanh$ represent Sigmoid and Tangent Hyperbolic activation functions d) Autoencoder e) Graph Neural Network f) Generative Adversarial Network, where $\mathrm{D}_{\mathrm{Loss}}$ and $\mathrm{G}_{\mathrm{Loss}}$ indicate Discriminator loss and Generator loss.

power. Their practical deployment must balance predictive performance against computational cost, memory access, and energy efficiency [86]. These aspects are particularly important when real-time response, privacy preservation, and reduced communication overhead are required, since portable intelligent devices cannot depend continuously on remote cloud execution [86].

For URLLC scenarios, learning-based methods must additionally satisfy stringent end-to-end latency and reliability requirements, since processing and queueing delays can themselves become critical bottlenecks. In 6G-oriented URLLC, end-to-end delay is expected to remain below 1 ms, with packet loss probabilities on the order of 10^{-5} or lower [87]. In this regard, lightweight deep learning and deep reinforcement learning are promising for real-time resource allocation and scheduling, provided they are supported by computationally efficient architectures, low-overhead inference pipelines, and edge-level intelligence that keep processing delay below the transmission interval. Overall, hardware-aware and latency-conscious learning designs are essential for translating deep learning into practical real-time 6G and IoT operation [86], [87].

In the following, a concise, high-level overview of state-of-the-art deep learning constructs for multi-antenna systems is presented.

1) Multilayer Perceptrons

Multilayer perceptrons (MLPs), often referred to as fully connected neural networks, are among the simplest deep learning models. Fig. 5(a) depicts a fully connected neural network, that can be used in both CSI estimation and signal detection tasks [6]. In MLPs, each neuron is considered to be part of several layers and is fully connected with neurons from the previous or subsequent layers. They are instrumental when the associated problem is either low-dimensional or when a lightweight model with relatively low computational demand is needed [88].

Despite their simplicity, an MLP can effectively estimate channels and detect symbols when properly trained, i.e., when the training database accurately captures signal statistics. They can be used as signal detectors in systems where CSI is assumed known, and the task is to directly learn the mapping from noisy received signals to the transmitted symbols [89]. Furthermore, MLPs are particularly well-suited to environments with relatively static, less complex channel characteristics, where their simplicity and lower computational requirements make them a preferred choice over more complicated neural network architectures [90].

MLPs may not learn spatial or temporal dependencies as well as convolutional neural networks (CNNs) or long short-term memory (LSTMs); however, they are still well-suited to hybrid models that combine other deep architectures. They can also be deployed as lightweight components to accelerate signal processing by leveraging their pretraining. [89]. Fully connected DNNs have also shown success in addressing channel feedback impairments, yielding higher-accuracy channel information and improved beamforming [32]. Compared with signal processing, MLPs are more useful for approximating complex mappings between received signals and transmitted symbols because they can handle nonlinearities in the system.

2) Convolutional Neural Networks

Convolutional neural networks can also be used in DL-based CSI estimation because they can process structured data, such as channel information represented as 2D matrices [91]. The structure of a convolutional neural network is shown in Fig. 5(b). In modern multi-antenna systems, the

CSI matrices can be visualized as images. Therefore, the CNN model's capabilities can be used to extract spatial features from noisy images [91], [92]. By employing CNNs, localized patterns and spatial correlations among antennas and subcarriers can be learned, significantly improving the accuracy and reliability of the CSI estimation. Moreover, the inherent ability of CNNs to generalize across channel variations, such as those induced by mobility or interference, enables them to perform robustly in dynamic and challenging wireless environments.

Denoising capability is a significant strength of CNNs, as it is particularly advantageous for noisy or incomplete CSI [93]. In orthogonal frequency-division multiplexing (OFDM) MIMO systems, where pilot signals may be insufficient, CNNs reconstruct missing or corrupted channel data by learning the underlying channel structure. This quality significantly improves CSI estimation accuracy and enables better performance in subsequent tasks, including improved signal detection and resource allocation. In addition, convolutional neural networks reduce pilot overhead by requiring fewer measurements to estimate channels, thereby supporting resource optimization in multi-antenna systems.

3) Long Short-Term Memory

LSTM neural networks can be employed for CSI estimation in dynamic environments due to their ability to capture temporal dependencies. Fig. 5(c) presents the LSTM construct along with its building blocks. LSTMs learn sequential patterns in time-varying channels, which is beneficial for CSI prediction, where the previous states of the channels influence future states [94], [95]. LSTMs introduce memory cells and gating mechanisms that mitigate the vanishing gradient problem, thereby enabling the model to retain information for extended periods. This property is beneficial in wireless communication systems, where CSI changes dynamically due to user mobility, interference, and other environmental factors.

Accurate prediction of CSI using LSTM-based models is a crucial capability for enabling dynamic resource allocation and adaptive system design in 6G multi-antenna systems, thereby ensuring their robustness and reliability [96]. LSTMs can also model the stochastic nature of wireless environments, enabling systems to adapt to rapid changes, optimize spectral efficiency, and enhance overall QoS.

4) Autoencoders

The application of autoencoders for dimensionality reduction and feature extraction in CSI feedback has been extensively explored in the literature [97]–[99]. A generic autoencoder design is shown in Fig. 5(d). Denoising autoencoders, in particular, are effective in CSI estimation and its feedback, making them suitable for achieving accuracy in the presence of interference and nonlinear effects in multi-antenna systems [100], [101]. Several studies have employed autoencoders to compress CSI feedback, thereby reducing overhead in massive MIMO systems. The integration of autoencoders enhances the efficiency of CSI transmission by reducing the feedback required for accurate CSI estimation, particularly in large-scale systems.

Advanced designs, such as variational autoencoders, extend the basic autoencoder framework by introducing probabilistic latent representations, enhancing noise robustness, and facilitating better generalization across diverse channel conditions [102], [103].

5) Graph Neural Networks (GNN)

Fig. 5(e) illustrates a basic structure of a graph neural network. Researchers have highlighted the potential of GNNs for CSI tracking and prediction, particularly when channels exhibit spatial and temporal correlations [104], [105]. In GNN-based models, nodes typically represent entities such as users or antennas, while edges capture the relationships between them, such as interference or spatial correlations. This graph-based representation enables GNNs to effectively learn and exploit underlying dependencies, thereby improving the performance of both CSI estimation and signal detection tasks.

Furthermore, the message-passing mechanism of GNNs enables efficient aggregation of global and local information, making them robust to dynamic channel variations and environmental changes. Given GNN capabilities, the authors of [106] leveraged graph topology to improve generalization and performance under model mismatches. The investigators in [107] utilized a GNN-based estimator to optimize the process and reduce computational complexity while maintaining accuracy.

6) Generative Adversarial Networks

The generative adversarial network (GAN), depicted in Fig. 5(f), is a powerful construct that can help address several challenges related to CSI in multi-antenna systems, including CSI estimation, incomplete CSI prediction, and CSI feedback improvement [108], [109]. The GAN comprises a generator and a discriminator that operate adversarially to generate realistic data distributions. GANs can reconstruct high-quality channel information from limited or noisy CSI estimates, where the generator produces plausible CSI matrices, and the discriminator ensures their accuracy. GANs can also be used for feedback denoising and enhancement, particularly when users experience a low signal-to-noise ratio (SNR) [110]. Conditional generative adversarial networks (CGANS) are also used for channel estimation in RIS-aided ISAC systems, overcoming the limitations of conventional deep learning approaches [111].

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

7) U-Net

U-Net is a deep learning architecture initially developed for biomedical image segmentation. The adoption of its name stems from its unique 'U'-shaped structure, which consists of two main parts, the encoder (contracting path) and the decoder (expansive path).

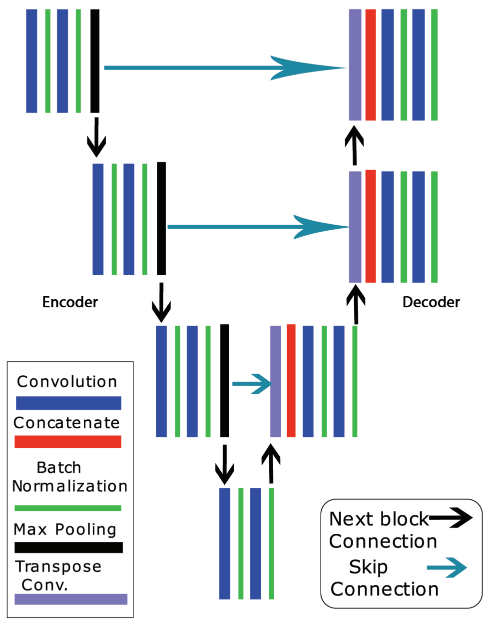

Figure 6: U-Net Architecture

The encoder comprises several convolutional layers that reduce the input's spatial resolution while capturing high-level features. The decoder, in contrast, uses upsampling layers to progressively increase the spatial resolution back to the original input size. This is typically achieved using transposed convolutions. A key feature of the U-Net architecture is the skip connections between corresponding encoder and decoder layers, which enable the network to retain fine-grained spatial information while learning abstract features. A generic U-Net design with 3 encoders/decoders, and a base layer is presented in Fig. 6. To address the complexity of channel estimation in multi-antenna systems, researchers have utilized the U-Net architecture both independently and in combination with other deep learning models [112]–[114].

8) Transformer

Recent investigations have demonstrated the immense potential of transformers for addressing challenges in multi-antenna systems, particularly for CSI estimation and prediction [115]–[117]. The block diagram of the transformer encoder and decoder is presented in Fig. 7. The self-attention mechanism of transformers enables them to model long-range dependencies and complex relationships within multi-dimensional data, such as the spatial and temporal features in CSI. By dynamically weighting the importance of different

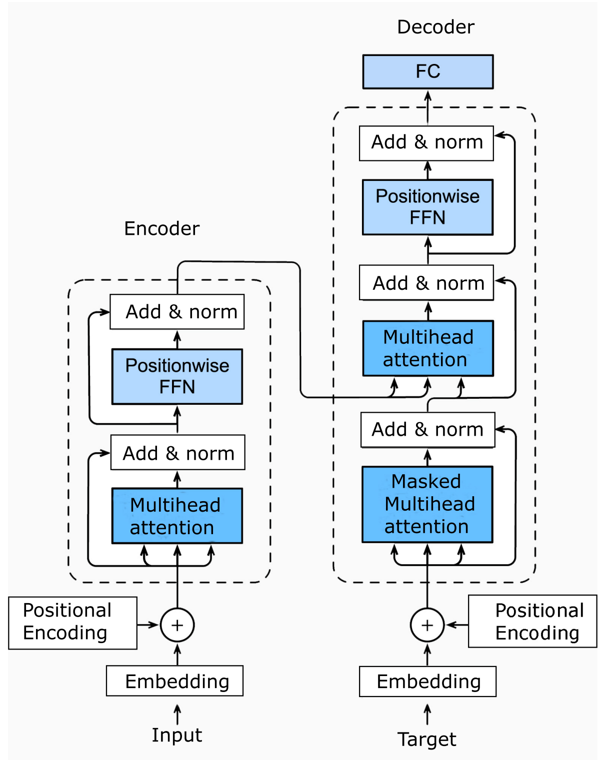

Figure 7: Transformer Architecture

elements in the input, transformers can effectively capture correlations and local patterns.

Concerning CSI estimation, transformers have demonstrated exemplary performance in reconstructing the channel matrix under dynamic and noisy conditions [118], [119]. High-dimensional input processing enables transformers to learn relationships between antennas and subcarriers, thereby accurately recovering missing or corrupted CSI. Moreover, transformers can capture temporal dependencies through incorporated positional encodings, making them effective for time-varying channels. Furthermore, transformers can reduce pilot overhead by estimating CSI with fewer measurements, optimizing resource utilization in multi-antenna systems.

9) Comparative Analysis

Overall, DL architectures employed in wireless communication systems differ primarily in their ability to exploit structural properties of the data, such as spatial, temporal, and relational dependencies. Multilayer perceptrons represent the simplest form, offering powerful nonlinear mapping capabilities but lacking the ability to capture inherent data structure, making them suitable mainly for low-dimensional or static scenarios [88]. In contrast, convolutional neural networks are specifically designed to extract spatial features and have demonstrated strong performance in channel estimation tasks by treating CSI as structured matrices [91], [92]. Recurrent neural networks, particularly LSTM networks, extend this capability to temporal domains by modeling sequential dependencies, enabling effective prediction in time-varying channels [94]. Autoencoders further contribute by learning compact latent representations, which are particularly useful for CSI compression and feedback reduction, while maintaining robustness against noise through denoising variants [100].

More advanced architectures focus on capturing complex dependencies and enhancing reconstruction quality. Graph neural networks model interactions among distributed entities, such as antennas or users, by leveraging graph-based representations, making them suitable for capturing relational structures in large-scale systems [104]. Generative adversarial networks, through their adversarial training paradigm, are capable of learning underlying data distributions and reconstructing high-quality CSI from limited or corrupted observations, although they suffer from training instability [109]. U-Net architectures enhance reconstruction and denoising performance by combining encoder–decoder structures with skip connections, preserving both global and fine-grained spatial features [113]. Recently, transformer-based models have gained prominence due to their self-attention mechanism, which enables the modeling of long-range spatial and temporal dependencies in high-dimensional data [116]. While these models offer superior performance in complex and dynamic environments, they typically require higher computational resources, highlighting an inherent trade-off between model complexity and practical deployment in real-time communication systems.

III. DL-assisted Channel Estimation in Multi-Antenna Systems

Massive MIMO and RIS-assisted massive MIMO have emerged as pivotal technologies for next-generation wireless networks, delivering substantial gains in spectral efficiency, energy savings, and link reliability. However, the high dimensionality introduced by large antenna arrays and densely deployed RIS elements poses significant challenges for accurate and efficient channel estimation.

The availability of CSI plays a key role in 6G systems, for example, in JCAS systems, CSI information helps in localization and sensing capabilities. Specifically, the estimated CSI captures the spatial-temporal propagation characteristics of the wireless channel, including angle-of-arrival/departure, delay (or time-of-arrival), Doppler, and path gain. These parameters are inherently related to the physical geometry of the propagation environment and can therefore be exploited for user positioning, target detection, tracking, and environmental perception [120]. For localization, CSI can be transformed into geometric features for model-based positioning or directly utilized as a high-dimensional fingerprint in data-driven approaches. For example, CSI-based deep learning models, such as CNNs, have demonstrated sub-wavelength positioning accuracy by learning spatial patterns embedded in massive MIMO CSI without requiring additional signaling overhead [121]. In sensing, reflections and motion-induced variations in CSI provide information about surrounding objects, thereby enabling the estimation of target range, direction, and velocity [122]. Furthermore, in JCAS systems, CSI and sensing are tightly coupled. On the one hand, CSI estimation provides the foundation for extracting sensing parameters such as AoA, delay, and Doppler. On the other hand, sensing information, particularly AoA, can be leveraged as prior knowledge to refine CSI estimation, thereby improving communication reliability [120]. This bidirectional relationship highlights CSI as a shared front-end for both communication and sensing tasks.

This section provides a comprehensive overview of recent deep learning–based channel estimation techniques tailored for massive MIMO and RIS mMIMO systems. Furthermore, it offers a critical evaluation of their performance and distills key insights to guide future research and practical deployment.

A. DL-assisted massive-MIMO estimation

The initial investigation in this domain was carried out by Neumann *et al.* [123]. This research effort was focused on learning a low-complexity MMSE channel estimator from a single noisy snapshot. Using the structure of a hierarchical channel model, MMSE and low-complexity approximate MMSE estimators tailored for a specific channel model with a single hyperparameter were derived. However, this approach is limited by the reliance on a single noisy snapshot, which may not fully capture the variability and dynamics of real-world communication channels. Continuing the efforts to reduce the estimation complexity, Dong *et al.* [91] employed a deep CNN and applied wideband channel estimation to mmWave mMIMO systems. The investigators used the tentatively estimated channel matrices across multiple adjacent subcarriers as input to the CNN and leveraged spatial and frequency correlations to achieve performance close to that of the ideal MMSE estimator. Ma *et al.* [124] explored a data-driven deep neural network (DNN)-enabled joint pilot and estimator design for wideband massive MIMO systems. Their proposed design employed an end-to-end DNN architecture comprising a dimensionality-reduction network and a reconstruction network to emulate pilot signals and channel estimators using data-driven deep learning. The research claimed computation complexity and performance gains over simultaneous orthogonal matching pursuit (SOMP)-based estimation. The complexity reduction of hybrid processing-based millimetre wave (mmWave) massive MIMO systems was investigated by Dong *et al.* [125]. In their proposed design, dubbed "spatial frequency-temporal CNN (SFT-CNN)" spatial and temporal correlations were used to improve accuracy. Further, a spatial pilot-reduced CNN (SPR-CNN) was used to reduce pilot overhead, leading to reduced complexity and a performance close to the ideal MMSE estimator. Belgiovine *et al.* [126] argued that conven-

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

tional channel estimation requires excessive computations for a massive MIMO system and explored the estimation of all pairwise channels simultaneously via an MLP-based DNN. Their proposed design accounted for computational complexity and performed slightly better than the conventional LMMSE estimator. Similarly, Yarotsky *et al.* [127] introduced a recurrent neural network (RNN) for beamspace channel estimation in massive MIMO receivers, reporting improved performance with comparable computational cost when compared to other RNN-based approaches.

Another significant area of research is the performance enhancement of DL-aided massive MIMO systems. In this domain, Gao *et al.* [128] investigated an attention mechanism that leveraged the channel distribution to enhance the estimation accuracy of highly separable channels. By exploiting the narrow angular spread, their method achieved significant performance gains with only a modest increase in complexity. The fast time-varying environments and associated non-stationary channel estimation in OFDM mMIMO systems were addressed by Liao *et al.* [129]. Their suggested DNN consisted of a two-dimensional CNN and a bidirectional long short-term memory (BiLSTM) module. This DL framework was trained on a large dataset of fast, time-varying channels, and the results demonstrated improved performance and channel estimation accuracy compared with conventional algorithms. Wei *et al.* [130] argued that the classical iterative algorithms can not accurately perform beamspace channel estimation using a sparse signal recovery framework. Consequently, they proposed a prior-aided Gaussian Mixture LAMP (GM-LAMP) approach. Initially, investigators derived a shrinkage function to refine the approximate message passing (AMP) algorithm, and subsequently, by substituting the shrinkage function in the LAMP network with the Gaussian mixture shrinkage function, a prior-aided GM-LAMP network was developed. Zhang *et al.* [131] proposed the integration of a “fully convolutional denoising network (FCDNet)” with a “learned approximate message passing (LAMP)” network. Their design, dubbed a “fully convolutional denoising approximate message passing (FCDAMP)” network, was capable of learning noise characteristics and denoising the signal, resulting in better performance, especially in low-SNR regions. In [132], He *et al.* presented a learned denoising-based approximate message passing (LDAMP) network to estimate channels in Beamspace mmWave Massive MIMO systems. This design achieved good accuracy with fewer RF chains at the receiver. The investigators also provided an analytical framework for the asymptotic performance of their designed channel estimator. The analysis and simulation results demonstrated superior performance compared with the latest compressed sensing-based algorithms. In [27], Melgar *et al.* highlighted the limitations, such as inadequate performance and covariance matrix unavailability in LS and MMSE estimators, respectively. Consequently, the authors proposed different DNN models that can work in conjunction with conventional channel estimation for improved performance. The performance results were close to those of the MMSE estimator without requiring any prior channel statistics. In an investigation by Ge *et al.*, [133], the issue of inaccuracy due to interpolation in the channel response of data subcarriers in an OFDM massive MIMO system was raised. To address this problem, researchers proposed a feed-forward neural network-based channel prediction framework that could replace interpolation. Simulation results indicated improved performance and the removal of dependencies on the channel autocorrelation matrix, prior statistics of the noise variance, and complex matrix inversion. In [134], Gao *et al.* treated beamspace massive MIMO channel estimation as a 2D image denoising problem. They proposed integrating AMP and FDDNet and demonstrated better performance with fewer RF chains and lower noise levels than conventional estimation techniques. Li *et al.* [135] also dealt with mMIMO channel estimation as the image denoising problem and suggested utilizing a denoising CNN (DnCNN) in the knowledge-driven machine learning (KDML)-based channel estimator. The performance of this design approached that of MMSE estimation without any prior statistical knowledge. Yuan *et al.* [136] presented a machine learning-based TDD design and associated channel estimator under time-varying channel conditions. This design incorporated a pattern extraction module and a CSI predictor. Concerning the CSI predictor, two combination choices were presented, i.e., CNN with AR predictor and CNN-RNN prediction approach. Numerical and simulation results demonstrated a significant gain in prediction quality and the possibility of a trade-off between prediction quality and estimation overhead. Yarotsky *et al.* [127] proposed a design which integrates a recurrent neural network and a nonlinear denoiser for iterative channel tap search and amplitude estimation, respectively. The authors reported better results with this design than with the sparse minimum mean-square error estimator. Balevi *et al.* [137] proposed DNN-based denoising to combat pilot contamination before a conventional LS estimation. Their proposed design approached the performance of the MMSE estimator as the number of antennas, subcarriers, and OFDM symbols increased.

Abboud and Jaber [138] suggested resolving the well-known massive MIMO pilot contamination problem using a two-stage framework. In this design, the first stage employed classical channel estimation to build a memory map of quantized channel state information (QCSI), and the second stage used DL to leverage the learned CSI map to predict the channel of the next user terminal. The authors claimed improvements in the sum rate of the mMIMO system. Balevi *et al.* [137] introduced a technique for estimating channels with low complexity in conjunction with effectively handling pilot contamination. They first denoised the received signal using a DNN, then applied a simple LS-type operation. By assuming a large product of the number of antennas, subcarriers, and coherence time, the authors demonstrated that the deep channel estimator achieves performance close to that of the MMSE estimator. Furthermore, the authors claimed complete elimination of interference. Hirose *et al.* [139] suggested two deep-learning techniques for channel estimation to mitigate pilot contamination effects. The first approach used a fully connected neural network to extract spatial information from contaminated signals, while the second used a CNN. Although the former method provided faster training, the latter approach produced a more accurate estimate of the channel. The proposed estimation methods were capable of handling imperfect timing synchronization among user terminals in neighbouring cells and the effects of channel ageing.

Sabeti *et al.* [140] emphasized the resolution of the excessive training overhead and proposed a deep learning-assisted technique for blind channel estimation in OFDM massive MIMO systems. In their design, the investigators proposed extracting transmitted data symbols without knowledge of the channel impulse response (CIR) using a mathematical formulation, and subsequently using the detected symbols as virtual pilots to estimate the user's CIR. A denoising convolutional neural network was employed to counter channel noise and enhance channel estimation accuracy. To reduce pilot overhead, Chun *et al.* [141] proposed a two-stage deep learning architecture. The first stage of their framework employed a two-layer neural network and a deep neural network to design the pilots and the channel estimator jointly. In the subsequent second stage, an additional DNN was employed on the data to improve channel estimation accuracy further. The knowledge-driven machine learning (KDML) model was presented by Li *et al.* [135], focusing on the design of a channel estimator that does not rely on any statistical properties of the wireless channel. The learning tasks were reconstructed using domain knowledge to simplify the network structure and minimize training overhead. Yuan *et al.* [136] addressed the issue of channel estimation overhead by leveraging the autocorrelation present in the CSI series. Instead of using conventional channel estimation methods, they employed machine-learning-based predictors to obtain CSI. The proposed predictors, namely "CNN combined with AR predictor (CNN-AR)" and "autoregressive network with exogenous inputs recurrent neural network (NARX-RNN)", were utilized to enhance CSI prediction. Numerical results demonstrated the superiority of the CNN-AR compared to other designs in terms of prediction accuracy.

In the domain of parametric channel and Direction of Arrival (DOA) estimation, Huang *et al.* [142] explored a DL design that could learn the channel features and spatial structures efficiently, leading to a better performance than conventional techniques. Zia *et al.* [32] presented a novel parametric channel model for massive MIMO and a corresponding parametric estimation algorithm. Their investigation employed an adaptive deep neural network architecture, termed "Adaptive DNN Compensator (ADN)," comprising an Rx-DNN pool and a Tx-DNN pool to refine channel parameters and mitigate channel feedback errors, respectively.

Understanding and mitigating nonlinear hardware impairments and differences between high and low resolution ADC, along with their impacts on CSI accuracy, are critical for enhancing the performance of a multi-antenna system. The prominent investigations in this domain include a study by Huang *et al.* [143], to calibrate the channel between the uplink and downlink to counter nonlinear hardware impairments. Their suggested DNN was trained with uplink and downlink channel measurements. The performance of their deep learning construct was on par with that of the LS-based designs. The investigation by Gao *et al.* [144] utilized a design dubbed "direct-input deep neural network (DI-DNN)" to estimate the channels from the received signals and further employed "selective-input prediction DNN (SIP-DNN)", which makes use of signals received by the high-resolution analogue-to-digital converters (ADC) antennas to predict other channels. In [145], Balevi *et al.* developed a two-stage deep learning pipeline architecture to estimate the uplink massive MIMO channel using one-bit ADCs. The proposed design was comprised of two generative deep learning models; first for compensating quantization loss and second for denoising. The results indicated that the proposed deep learning-based channel estimator surpasses other state-of-the-art estimators for one-bit quantized massive MIMO systems. Elbir *et al.* [146] focused on three DL frameworks utilizing CNNs to provide channel estimates and hybrid beamformers for OFDM mMIMO systems. The suggested frameworks offered higher spectral efficiency, improved resilience to corrupted channel data, and outperformed the spatial frequency CNN (SF-CNN) presented in [125].

Several investigators employed DL techniques to obtain sparse channel representations, achieving high estimation accuracy with lower computational cost. Jin *et al.* [92] exploited the sparsity of the mmWave channel matrix and represented it as a natural image. They introduced a flexible and fast denoising convolutional neural network (FFDNet) for channel estimation. The investigators also showed that FFDNet outperforms other methods in terms of normalized mean-squared error across varying noise levels. Ma *et al.* [147] proposed DL-based compressed sensing channel estimator. In their proposed framework, the correlation between the received signal vectors and the measurement matrix was fed into a DNN to predict the beamspace channel amplitude. Subsequently, the indices of the leading beamspace channel entries were used for channel reconstruction.

Gao *et al.* [148] proposed a DL-based framework to enhance uplink channel estimation in hybrid analogue-digital (HAD) massive MIMO systems using a novel angular space segmentation method. The method divides the angular space into smaller regions and trains a distinct neural network for each region to exploit channel sparsity. Investigators claimed improvements concerning measurement efficiency and channel estimation capability. In another investigation, Gao *et al.*

[149] presented a DL-based channel estimation approach that translated the sparse Bayesian learning (SBL) algorithm into a DNN. Each SBL layer in the DNN updates the Gaussian variance parameters of the sparse angular domain channel, effectively capturing complex sparsity structures. The authors claimed that the proposed framework could be easily extended to the multi-block case by leveraging temporal channel correlation to adaptively predict the measurement matrix. In [150], Ma *et al.* proposed a model-driven deep learning (MDDL)-based channel estimation for mmWave massive hybrid MIMO systems. In the suggested design, channel sparsity in the angle-delay domain was employed to reduce overhead. The uplink pilot training overhead was addressed through joint training of the phase-shift network and the channel estimator as an autoencoder. The investigators also considered the downlink channel estimation and feedback for the FDD case. They suggested that the pilots at the BS and the user channel estimation can be jointly trained as an encoder and a decoder, respectively. In [151], Wen *et al.* presented CsiNet, a novel DL design for CSI sensing and recovery, which was capable of learning the transformations between CSI and their optimal representations. The authors claimed that the recovered CSI has superior quality compared to current compressive sensing-based approaches.

Ma *et al.* [152] proposed a deep learning compressed sensing (DLCS) scheme for channel estimation, encompassing beamspace channel amplitude estimation, and reconstruction. Moreover, they proposed a hybrid precoding method called deep learning quantized phase (DLQP), which incorporated a training hybrid precoding neural network (THPNN) to obtain analog precoding vectors per user and a deployment hybrid precoding neural network (DHPNN) to obtain an analogue precoding matrix for all users.

Guo *et al.* [34] introduced a deep learning-based framework that utilizes uplink CSI for downlink channel acquisition, aiming to minimize training overhead. Their approach involved an adaptive pilot design that utilized magnitude correlation in bidirectional channels within the angular domain to enhance channel estimation as well as bit allocation for feedback. In [153], Sun *et al.* presented an approach based on a new low-rank model for the double-directional massive MIMO channel estimation. The authors modeled a downlink channel with only a few paths as a low-rank channel and efficiently recovered the channel at the base station using a small number of scalar measurements. Instead of relying on angle parameterization, investigators leveraged the inherent low-rank structure of the double-directional channel model. The authors claimed better performance than sparse estimation designs.

In the domain of state-of-the-art LLM-based designs, Liu *et al.* [154] introduced LLM-empowered channel prediction (LLM4CP), a novel channel prediction method that employs pre-trained large language models to predict future downlink channel state information sequences from historical uplink CSI sequences. The authors reported good channel-prediction performance with low training and inference costs, outperforming existing approaches across full-sample, few-shot, and generalization tests.

Continuing these efforts, Fan *et al.* [155], introduced CSI-LLM. This novel LLM-powered downlink channel prediction technique models variable-step historical sequences to improve CSI prediction. The results demonstrated that CSI-LLM outperforms existing methods, achieving stable, optimal performance for continuous multi-step prediction. Catak *et al.* [156] developed BERT4MIMO, a transformer-based architecture designed to enhance the reconstruction of CSI in massive MIMO systems, leveraging the BERT model's capabilities with feature embeddings and attention mechanisms. The results showed that BERT4MIMO effectively reconstructs CSI under varying mobility scenarios and channel conditions, with strong performance in stationary and high-speed mobility scenarios. Ju *et al.* [116] focused on the development of T-PaCE, a Transformer-based predictive channel estimation technique for mmWave massive MIMO systems, which efficiently predicted downlink channel parameters by learning the spatio-temporal correlation of multipath components. The results showed that T-PaCE significantly outperforms conventional channel acquisition techniques, achieving lower normalized mean square error.

Zheng *et al.* [157] introduced SLRTNet, a deep learning model that combines super-resolution techniques with a transformer architecture for enhanced channel estimation in OFDM systems. The results indicated that SLRTNet significantly outperforms existing models, achieving a 5.5 dB gain in high SNR regimes and demonstrating greater robustness across extended SNR and Doppler frequency shift ranges. Zeng *et al.* [118] also contributed to the LLM-based estimation by investigating the CSIGPT approach, which integrated a Swin Transformer-based CSI acquisition network with federated-tuning and a variational auto-encoder to address CSI acquisition challenges in massive MIMO systems. Authors claimed that SWTCAN outperforms state-of-the-art methods in CSI acquisition, while the federated-tuning method significantly reduces communication overhead, improving efficiency in practical deployment.

U-Net-based estimation and denoising techniques have also garnered significant attention in recent research, with notable contributions such as the work by Zhao *et al.* [112], which employed a ResNet-UNet hybrid model for channel estimation in millimeter-wave Massive MIMO systems, addressing challenges in low-SNR regimes. Continuing these efforts, Imran *et al.* [113] introduced an attention-residual U-Net architecture for channel estimation in massive-MIMO FSO systems, combining attention and residual connections to improve performance. The model significantly outperforms traditional methods and other deep learning models, particularly in terms of MSE under atmospheric turbulence and noise.

The performance comparison of different estimators, along with the computational complexity, is given in Table 2.

Please note that the SNR range is selected based on the drop in bit error rate (BER) and normalized mean squared error (NMSE) to provide a clear performance comparison. The variables associated with computational complexity are presented in the footnote of Table 2.

The studies referenced above clearly underscore the significant impact of CSI imperfections on the performance of massive MIMO systems. However, despite these advances, there remains a critical need to evaluate the comparative advantages of model-driven versus data-driven deep learning approaches to determine which methodology delivers optimal performance.

In conclusion, across the reviewed DL-assisted massive-MIMO estimators, a consistent pattern is the reliance on structural priors, such as sparsity, frequency correlation, or parametric representations, combined with DL-based refinement. These denoising-based studies include CNN- and SFT-CNN-based designs [91], [125] and hybrid AMP–DL approaches such as GM-LAMP and LDAMP [130], [132]. Model-driven architectures (e.g., MDDL [150] and ADN [32]) embed analytical structure and achieve predictable complexity, whereas data-driven estimators (e.g., FCDNet [131], DnCNN-based KDML [135], and FFDNet [92]) depend heavily on large training datasets and channel stationarity. While some of these works assume block-fading or quasi-static channels, others explicitly target fast-varying environments using CNN-BiLSTM predictors [129] or CNN–AR/LSTM-based time-series models [136], resulting in differing claims regarding temporal robustness. Furthermore, studies differ in whether DL inference cost is counted, as some works have treated it as negligible [27], [137], while others provide explicit Big-O analysis [32].

B. DL-assisted RIS massive MIMO estimation

The growing demand for higher data in wireless communications has driven the deployment of massive MIMO systems at higher frequency bands, where severe propagation losses motivate the integration of reconfigurable intelligent surfaces. RIS-enabled massive MIMO has emerged as a promising paradigm for mitigating high-frequency attenuation while enabling partial control of the wireless propagation environment. However, unlike conventional massive MIMO systems with simplistic BS–UE channel, RIS-assisted systems exhibit a fundamentally different propagation structure in which the effective channel is cascaded, comprising of BS–RIS and RIS–UE links, together with an additional direct BS–UE component [10], [17]. This cascaded channel model substantially increases the complexity of channel estimation, which is a critical prerequisite for precoding design and RIS phase-shift optimization [163]. The estimation challenge is prominent in passive RIS configurations, where no signal processing or direct channel observation can be conducted at the RIS, forcing the base station to estimate the cascaded BS–RIS–UE links while separating the direct BS–UE path, often through specialized training protocols [17]. Moreover, the use of large RIS arrays with hundreds of reflecting elements significantly increases the channel dimensionality, resulting in higher estimation complexity, longer training overhead, and reduced spectral efficiency [7], [17]. To address these challenges, DL-based channel estimation methods have attracted considerable attention owing to their ability to learn complex channel relationships and deliver accurate estimation performance with manageable computational complexity. A comparison of notable channel-estimation studies and their respective computational complexities is presented in Table 2.

Among the key CNN-based RIS channel estimators, the key investigations include the research conducted by Elbir *et al.* [38]. Their research employed a deep convolutional neural network trained on received pilot signals to learn about direct and cascaded channels. In this work, the performance of the proposed ChannelNet framework is shown to outperform other DL-based approaches, including a multi-layer perceptron and the SF-CNN. Also, the proposed ChannelNet-based approach does not need to be retrained when user locations change by up to 4 degrees.

Continuing the efforts in this domain, Kundu *et al.* [25] carried out a minimum mean square error channel estimation via a data-driven deep learning approach. The authors first used the LMMSE estimator along with majorization-minimization (MM) for RIS phase shift optimization. Subsequently, deep CNN architectures were employed to cast the channel estimation problem as an image denoising problem. To address the computational complexity of the CNN-based MMSE estimator, the authors of [164] proposed a super-resolution denoising convolutional neural network (SRDnNet), which is a concatenation of earlier works, e.g., the super-resolution convolutional neural network (SRCNN) and the DnCNN. The channel was modeled as a super-resolution image to estimate the channel gains at pilot positions for the RIS-aided mMIMO OFDM system. The results indicated significantly reduced prediction time compared to competitors such as LS, DFT-LS, LMMSE, and ChannelNet, along with more than 10dB NMSE performance gains. Based on the Bayesian MMSE criterion, Liu *et al.* contributed in this domain [158] by proposing a data-driven deep residual (DReL) network to address the channel estimation problem in the RIS-added TDD multiuser communication system. The channel estimation was modeled as a denoising problem, in which the residual noise was learned to estimate the channel coefficients from noisy pilot-based observations. As a realization of the developed DReL framework, a CNN-based DRN (CDRN) was then proposed. In this framework, a CNN denoising block equipped with an element-wise subtraction structure was specifically designed to exploit both the spatial features of the noisy channel matrices and the additive nature of the noise simultaneously. The authors claimed its performance was close to optimal MMSE.

In another investigation by Liu *et al.* [165], the challenge of high-dimensional cascaded channels in RIS-enabled

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

Table 2: Performance comparison of different techniques for DL-assisted channel estimation

Category	Research	SNR (dB)	NMSE / BER drop approx.	DNN Model / Type	Computational Complexity
Massive MIMO	Jiang *et al.* [93]	0 to 10	10^{-1} to 10^{-2} (NMSE)	Dual CNN	$O(8 \times (108 \times N_t K + N_t K \log K + K N_t \log N_t))$
	Dong *et al.* [125]	-5 to 20	10^{-1} to 10^{-3} (NMSE)	Spatial-Frequency-Temporal CNN	$O(\sum_{l=1}^{L_c} M_{1,l} M_{2,l} F_l^2 N_{l-1} N_l)$
	Zia *et al.* [32]	8 to 14	-10^{-1} to -10^{-3}(BER)	Adaptive Neural-Net Compensator	$O(N_t N_r \log(N_t N_r) + P)$
	Ma *et al.* [150]	-10 to 10	8×10^{-1} to 6×10^{-2} (NMSE)	Model-Driven DL	$O(MN_tK + TMGK)$
	Balevi *et al.* [137]	4 to 20	10^{-3} to 10^{-4} (NMSE)	Deep Estimator (DNN followed by LS-estimator)	$O(N_t)$
	Liao *et al.* [129]	18 to 30	10^{-1} to 10^{-3}(BER)	CNN-BiLSTM Hybrid	$O(N_r N_t N^2)$
	Elbir *et al.* [146]	-15 to 6	10^{-2} to 10^{-3}(NMSE)	Multi-Carrier Hybrid Beamforming Network	$O(\sum_{l=1}^{L_c} D_x^{(l)} D_y^{(l)} b_x^{(l)} b_y^{(l)} c_{CL}^{(l-1)} c_{CL}^{(l)})$
RIS-assisted Massive MIMO	Liu *et al.* [158]	0 to 25	10^{-1} to 10^{-3}(NMSE)	CNN based Deep Residual Network	$O(DN_t(N_i+1)\sum_{l=1}^{N_{tr}} n_{l-1}.s_l^2.n_l)$
	Abdallah *et al.* [159]	-20 to 0	2 to -19 (NMSE)	Frequency-Selective Cascaded Channel Estimation	$O(UKPN_t + (K+2K_p) + UQN_iL_GL_u)$
	Xu *et al.* [95]	-5 to 30	10^{-1} to 10^{-5}(NMSE)	Sparse-Connected Long Short-Term Memory,	$O(\frac{2N_{ris}^3+(N_tL+1)N_{ris}^2+N_tLN_{ris}I_{\max}+(N_{ris}+N_t)^3+(N_t+N_{ris})^2U}{\tau}$ $+360KN^2 + (4KM + 4M + 42K)N)$
	Jiang *et al.* [160]	-5 to 8	10^{-1} to 10^{-6}(NMSE)	End-to-End Deep Neural Network,	$O(kN_t + N_t^2 + N_r^2 + N_tN_r + NN_t + NN_r)$
	Ginige *et al.* [161]	0 to 30	10^{-2} to 10^{-6}(NMSE)	Tucker2-BLAS and CNN-AR,	$O(iQ(N_i^2T(U+N_t) + 2N_tUTN_i) + 28514UMQ + 64(Q+1) + 8192)$
	Xiang *et al.* [162]	0 to 16	10^{-1} to 10^{-4} (BER)	PreRec-Correlator-Net	$c_1(LNlog(LN)) + c_2(I_tQ)$

The variables in the computational complexity column are as follows: N_t and N_r represent the number of antennas at the transmitter and receiver, respectively, K represents the number of subcarriers, L_c is the number of convolutional layers, $M_{1,l}$ and $M_{2,l}$ denote the number of rows and columns of each feature map output by the l-th layer, F_l is the side length of the filters used by the l-th layer, N_{l-1} and N_l denote the numbers of input and output feature maps of the l-th layer, P is the number of rays, $M = QN_{RF}$ where Q is the number of timeslots and N_{RF} is the number of RF chains, T is the number of layers in the channel reconstruction network, G represents columns of the redundant dictionary matrix, N is the size of the OFDM symbol, $D_x^{(l)}$ and $D_y^{(l)}$ are kernel dimensions, $b_x^{(l)}$ and $b_y^{(l)}$ are dimensions of the l-th convolutional layer output, and c_{CL}^{L-1} and c_{CL}^{L} are the number of filters in the corresponding layers.

D is the number of denoising blocks, N_i is the number of IRS elements, N_{tr} is the number of training examples, n_l is the number of neural network channels in layer l and s_l is spatial size of filter in layer l, U is the number of users, P is the number of pilot training frames, L_G is the number of paths between BS and RIS and L_u is the number of paths between RIS and user u, K is the frequency-domain subchannels, N_{ris} is the size of RIS, L is the number of layers, k is the length of the bit stream, i is the number of iteration before convergence is achieved, Q is the number of coherence intervals in the pilot-based training phase, T is the minimum pilot length required. All SNR ranges and NMSE/BER values are taken from the respective original papers and were not reproduced under a unified experimental setting; therefore, the comparison should be interpreted as indicative rather than a strictly controlled head-to-head benchmark.

millimetre-wave massive MIMO is tackled by utilizing a deep denoising neural network-assisted compressive channel estimation. The proposed design incorporated a hybrid passive/active RIS framework, where a few receive chains were employed to estimate the uplink user-RIS channels. In the channel estimation stage, a limited number of RIS elements participated in the estimation. Later, the complete channel matrix can be reconstructed from the partial measurements using compressive sensing, and subsequent refinement can be achieved using a complex-valued denoising convolutional neural network (CV-DnCNN).

Xu *et al.* [166] suggested a channel extrapolation method where a subset of RIS elements was turned on, and the cascaded channel was first estimated, which was then extrapolated to estimate the whole channel. In this work, the standard CNN architecture was modified by introducing ordinary differential equations (ODEs) into the neural network to model latent relations among different data layers. Two numerical differentiation methods, namely Leapfrog and Runge-Kutta, were introduced into CNN. Performance results, evaluated using NMSE and the loss function, demonstrated the method's efficacy compared with the standard CNN. The double deep learning-based method was proposed in [167] by Li *et al.*, which addressed cascaded channel estimation, joint optimal RIS phase shifts, and beamformer design to maximize the channel capacity between the source and destination. Two neural networks were introduced in this construct, avoiding complex iterative optimizations that are part of traditional supervised and reinforcement learning.

The approach proposed by Shen *et al.* [168] was based on federated learning (FL), which enabled users to estimate their local channels and then collaboratively update the estimation model via the RIS. The RIS coordinates the FL process and aggregates the locally estimated channel parameters from the users to improve the overall channel estimation accuracy. The results showed that the proposed approach significantly enhances channel estimation accuracy compared to traditional methods, particularly in scenarios with a large number of users. The investigation by Elbir *et al.* [169] further discussed the application of federated learning for channel estimation in RIS-assisted massive MIMO systems. The FL-supported design proposed in this investigation enabled each user to estimate its own channel locally and then share this information with the base station, rather than transmitting raw data. The authors claimed that this FL approach can improve the system's overall performance by enabling information sharing between users and the RIS.

GANs have shown promise for generating synthetic datasets to train machine learning models, ensuring robustness and generalization in real-world scenarios. For example, the investigators in [170] addressed the challenges of commercial standard Wi-Fi devices sensing at mmWave frequencies by using a GAN to generate synthetic CSI samples. The method successfully augmented existing datasets with 30,000 additional CSI samples, thereby improving data consistency.

Additionally, the augmented samples enhanced the generalization of a pose classification model, demonstrating the potential of GANs for advancing joint communication and sensing applications. In [111], the authors proposed a CGAN framework to train two deep learning networks adversarially, enabling the generator to learn a mapping from observed data to real channel conditions and to refine its output based on feedback from the discriminator, thereby optimizing both the training process and estimation accuracy.

Rahman *et al.* [41] proposed a BiLSTM construct to improve the channel estimation and signal detection in RIS-assisted multi-user multiple-input single-output (MU-MISO) OFDM systems. The proposed framework employed a technique to estimate the CSI of direct and cascaded links, along with a deep neural network-based detector to improve signal-detection accuracy. The proposed model was tested with various cyclic prefix (CP) lengths, as well as with a couple of pilot sequences and optimizers. The authors claimed a better performance of the proposed Bi-LSTM network concerning channel estimation and signal detection compared to LSTM, CNN, gated recurrent unit (GRU), CNN-GRU, MMSE, and LS-based designs.

Xu *et al.* [95] focused on a deep learning-based approach for time-varying channel prediction in RIS-assisted MU-MISO networks. The proposed framework obtains CSI via channel prediction and effectively decomposes the cascaded channels into separate BS-RIS and RIS-UE channels. This construct exploits the two-timescale property, in which the BS-RIS channel is quasi-static, and the RIS-user channel is fast time-varying. A deep neural network architecture dubbed "Spatial-Channel LSTM (SCLSTM)" was employed to predict the future channel state. In the SCLSTM-based algorithm, the current CSI and the RIS reflection coefficients were used as inputs to predict the future CSI. Lu *et al.* [171] introduced an innovative hybrid precoding architecture that combined geometric mean decomposition (GMD) with an LSTM autoencoder. This joint design approach employed GMD to reduce computational complexity and improve hybrid precoding performance, whereas the encoder-decoder design served as a dimensionality-reduction strategy. The proposed framework introduced a novel approach to maximize energy efficiency, enabling the RIS to select the optimal hybrid precoding matrix without relying on CSI.

To address the challenge of reduction in complexity while maintaining performance, Kim *et al.* [172], presented a novel channel estimation method for RIS-assisted mmWave mMIMO systems, employing a pruned autoencoder. By utilizing a denoising autoencoder, the system learns to reconstruct channel information from a reduced set of pilot signals, while the Average Percentage of Zeros (APoZ) metric is used to identify and prune less significant neurons, optimizing pilot pattern in terms of both number and location. Experimental results on the DeepMIMO dataset demonstrate that the proposed method achieves nearly identical or superior NMSE performance across various SNR levels and pilot configurations. Working on similar lines, Wu *et al.* [173] proposed a deep learning-based rate-splitting multiple access (RSMA) scheme for RIS-aided tera-hertz (THz) massive MIMO systems. The proposed RSMA scheme leverages the unique features of THz communication, such as wide bandwidth and high directional beamforming capabilities, to enable high-speed data transmission. The suggested RSMA was a two-stage design. In the first stage, the base station employed the RIS to divide the users into two groups based on their channel quality. The subsequent second stage used a DNN-based RSMA scheme to allocate power and data rates to the users within each group. Simulation results showed that the proposed RSMA scheme achieves a higher sum rate and energy efficiency than conventional multiple-access schemes. Further investigations in this domain include the work by Abdallah *et al.* [159], which introduced novel data-driven and compressive-sensing-based techniques for estimating cascaded channels in RIS-assisted multi-user millimetre-wave massive MIMO systems. Their method leveraged the shared sparsity across subcarriers and the double-structured sparsity of angular-cascaded channel matrices. Denoising neural networks were employed in the proposed data-driven channel estimation approaches to improve accuracy.

The computational complexity challenge associated with massive antenna arrays was addressed by Jin *et al.* [39] by proposing a channel estimation algorithm based on deep residual networks (ResNets). By reshaping the channel matrix into a two-dimensional natural image, the channel estimation problem for millimetre-wave RIS systems can be treated as an image super-resolution (SR) problem. The similarity between sparse channel estimation and the image SR problem motivated the authors to employ enhanced deep residual networks to estimate the channels accurately. The authors proposed two variants of the ResNet architecture for channel estimation in RIS, i.e., EDSR and MDSR. The EDSR variant was a single-scale ResNet that employed residual connections and skip connections to improve the learning of channel features and significantly reduce estimation complexity. The MDSR variant was a multi-scale ResNet that utilized multiple feature map scales to capture fine-grained and coarse-grained information. The authors claimed substantial performance improvements in NMSE relative to SNR compared with existing baseline estimators.

To improve performance while reducing complexity, a skip-connection attention (SC-attention)- type DL network was proposed by Liu *et al.* in [40]. It was designed to improve channel estimation performance in double-RIS-aided massive MIMO systems. The suggested SC-attention network utilized self-attention layers and skip-connection structures to process the noisy pilot-based observations and extract more accurate CSI. Self-attention layers enabled the network to selectively attend to different parts of the input, focusing on the most relevant features for the estimation task. Skip-connections aid in the mitigation of the vanishing gradient problem, enabling the network to learn a more complex mapping between the input and output. The suggested design was evaluated through simulations and showed better performance as compared to existing channel estimation methods in terms of accuracy and efficiency. In the investigation by Vu *et al.* [174], a system that combines cognitive radio with RIS to support NOMA was proposed. The system considered the presence of imperfect CSI and imperfect successive interference cancellation (SIC). Specifically, the study focused on a scenario where a source node communicated with two secondary users while adhering to a multi-primary receiver constraint. Moreover, insights into the system's performance were offered by deriving a closed-form expression for the users' outage probabilities and system throughput in line-of-sight (LoS) and non-LoS scenarios. These expressions were validated through Monte Carlo simulations. The investigation by Ren *et al.* [175] focused on utilizing long-term CSI to optimize the deployment and configuration of RIS in multiuser MISO systems. Their objective was to maximize the system's performance by jointly optimizing the RIS phase shifts and transmit-beamforming vectors. To achieve this, the authors proposed the deep deterministic policy gradient (DDPG) algorithm, which enables the RIS to dynamically adapt its phase shifts, taking into account both the current and past CSI. The results demonstrated the effectiveness of the DRL-based design in optimizing the RIS configuration for multiuser MISO systems. Compared to traditional approaches that rely on instantaneous CSI, the long-term CSI-based design focuses on DRL to achieve improved performance in terms of achievable rates and spectral efficiency. Another application of DDPG algorithm is also presented in [176] for joint optimization of transmit precoding and RIS phase shifts in multi-RIS aided multiuser downlink systems, while explicitly accounting for practical amplitude-phase coupling effects.

The U-Net-based designs in the RIS domain include the investigation by Xie *et al.* [114] proposed a deep compressed sensing-based channel estimation scheme using a U-Net to recover high-dimensional cascaded channel matrices with limited pilot overhead. The architecture, which incorporates stacked residual units, enhances feature extraction and outperforms conventional methods in terms of accuracy, generalization, and robustness. Similarly, Tong *et al.* [177] proposed a diffusion model-based channel estimation method for RIS-aided communication systems, incorporating a U-Net for efficient channel recovery and a phase noise reduction mechanism. Their key contributions include employing progressive distillation for reduced sampling steps and enhancing robustness against RIS phase noise. Simulation results demonstrate that their method outperforms baselines by over 3.2 dB in NMSE and achieves a 3.74 dB gain under phase noise conditions, without retraining for different SNR settings. Xiao *et al.* [178] proposed a U-shaped multi-layer perceptron (U-MLP) network for hybrid-field channel estimation in XL-RIS-assisted mmWave MIMO systems, incorporating a permutator module to model non-stationary channels. Their method used a split attention mechanism and a U-shaped backbone to address the rank deficiency of cascaded channels. Simulation results show that U-MLP achieves better channel estimation accuracy with less pilot overhead compared to state-of-the-art methods.

The field of Transformer-based channel estimation and prediction has garnered significant attention in recent years. The notable investigations in this domain included Guo *et al.* [179], which proposed a novel Transformer-based channel estimation framework called EPformer, which addresses the challenges of RIS-aided communication by formulating the problem as a parallel super-resolution task. The method uses multi-head attention to parallelize the estimation process, enhancing computational efficiency, and incorporates a super-resolution block for denoising. The simulations show that EPformer outperforms traditional methods (LS, MMSE) and a couple of deep learning-based solutions [38], [180] in terms of estimation accuracy and computational time, achieving better NMSE performance across various SNR levels. Qi *et al.* [181] proposed a deep learning-based channel estimation scheme for RIS-assisted time-varying MIMO systems, incorporating a RIS element selection model (RESM) to reduce system overhead and a channel estimation model (FCEM) that uses attention mechanisms to accurately extrapolate complete channel information. They introduced a multi-head attention mechanism in FCEM to enhance accuracy in time-varying environments. Simulation results demonstrate that the proposed scheme outperforms state-of-the-art methods, such as the deep probabilistic subsampling [182] and ordinary differential equations-based scheme [183], in terms of NMSE performance at various SNR values.

Mohsin *et al.* [184] proposed a novel transformer-based distributed machine learning approach for downlink channel estimation in RIS mMIMO systems, incorporating a custom multi-domain attention mechanism to capture spatial, temporal, and channel dependencies. They introduced parallel attention branches to model complex wireless channel characteristics, enhancing computational efficiency and enabling accurate estimation. Simulation results show that the transformer-based distributed machine learning method outperforms CNN and ResNet-based models.

Singh *et al.* [185] proposed a novel graph transformer-based approach for RIS channel estimation, dividing RIS elements into groups and using spatial correlation to reduce pilot overhead. The approach integrates a graph neural network for embedding correlations and a transformer model with an attention mechanism for channel prediction. This model outperforms such as existing methods, such as deep residual learning [158], in terms of normalized mean square error, requiring fewer pilots for similar or better estimation accuracy. Xia *et al.* [186] proposed a two-timescale channel prediction framework with a transformer-based parallel prediction scheme, addressing sequential error propagation in RIS mMIMO systems. Their method utilized the sparse connected multi-causal convolution attention transformer to predict future channels in parallel, improving accuracy in dynamic environments. The proposed transformer incorporates an improved attention mechanism that extracts more local context features, mitigating sequential prediction errors and improving channel prediction reliability. Simulation results show that the proposed method outperforms traditional sequential prediction methods (LSTM, GRU, RNN), demonstrating superior performance in terms of NMSE.

Xiang *et al.* [162] presented a deep learning framework termed PreRec-Correlator-Net to address channel estimation, symbol detection, and phase shift optimization for RIS-enabled mMIMO systems in industrial IoT environments. The framework combines the Pre-Net, Rec-Net, and Correlator-Net modules for enhanced estimation accuracy, detection, and adaptive pilot retransmission. The simulations demonstrated that PreRec-Correlator-Net significantly outperforms traditional models, achieving substantial improvements in BER and MSE at various SNR levels. The overall framework reduced the error floor, offering a reliable solution for industrial IoT communications.

In summary, several investigations concerning DL-assisted RIS channel estimation treat the cascaded BS–RIS–UE channel as a 2D/3D image and employ CNN, U-Net, or super-resolution denoisers for reconstruction, e.g., ChannelNet [38], SRDnNet [164], CDRN [158], and CV-DnCNN-assisted methods [165]. More recent designs incorporate sequential or two-timescale models such as SCLSTM [95] or ODE-enhanced CNNs [166] to address temporal channel variation. Some investigations assume ideal conditions, while others explicitly include interference, hardware impairments, or partial RIS activation [41], [167]. Most studies also vary in validation methodologies, with some studies using NMSE values plotted against certain SNR values, while other studies focus on bit error rates or DL train/test evaluations, complicating direct comparison. Overall, RIS channel estimation research is split between high-complexity architectures yielding strong NMSE performance and lightweight/federated approaches [168], [169] prioritizing scalability and practical deployment. To facilitate the identification of DL architectures suitable for specific channel conditions and deployment scenarios, Table 3 summarizes the relevant studies reported in the literature.

C. Critical analysis and lessons learned on DL-assisted estimation

Deep learning-assisted channel estimation for massive MIMO systems has emerged as a breakthrough solution to address the shortcomings of traditional methods, such as LS, MMSE, recursive least squares (RLS), and Kalman filtering [187]. Conventional methods often rely on assumptions of linearity, time invariance, or specific statistical models, which may not hold in practical deployments, resulting in suboptimal performance. DL techniques, however, exhibit the ability to learn complex, non-linear relationships between received signals and channel estimates without requiring prior assumptions. This adaptability makes DL well-suited to model the intricate dynamics of real-world wireless channels, offering substantial improvements in CSI estimation accuracy, especially in dynamic environments [188].

A significant issue in evaluating DL-assisted channel estimation methods is the potential bias in fair comparison, as the channel models and corresponding assumptions often differ across studies. This disparity can lead to results that may not generalize well to other channel conditions due to deep learning models' inherent data dependency. Key assumptions that may differ in the context of mMIMO and RIS-enabled mMIMO are [12], [23], [24], [27], [28], [189]:

- Block fading vs. time-varying channels
- Employing statistical CSI instead of complete CSI
- Taking measured channels instead of assuming channel model
- Structured sparsity assumption
- Spatial or temporal channel correlations
- Cascaded channels in RIS with or without direct BS to MS links
- Including or neglecting inter-user interference (IUI)
- Self-interference in hybrid RIS
- Deployment constraints, such as IoT devices, that prohibit computationally heavy DL-model deployment.

These challenges highlight the importance of carefully considering the channel model and relevant assumptions in the investigations. Similar care should be taken with DL constructs plugged into modern multi-antenna systems, as well as with the datasets used for training and testing. Researchers should exercise caution when making direct performance comparisons across studies, as differing assumptions may lead to discrepancies in results and interpretations. In the subsequent discussion, we analyze the issues pertinent to DL-assisted channel estimation in massive MIMO and RIS massive MIMO systems.

1) Massive MIMO Systems

In addition to the functional division of deep learning-assisted estimation discussed in section A, which includes low-complexity estimation [91], [125]–[127], estimation for higher accuracy [128]–[131], pilot decontamination-based designs [137]–[139], parametric estimation [32], [142], and sparse channel estimation [147], [148], a further categorization can be made based on the level of DL involvement in conventional designs. These categories offer valuable insights into how DL is integrated into the system model.

1) Hybrid models with equal involvement of DL and signal processing aim to attain maximum performance by utilizing the strengths of signal processing and DL techniques. In such models, conventional algorithms are often used to preprocess received signals or reduce the dimensionality of the input data, while DL is

Table 3: DL techniques in practical deployment scenarios. The cited studies are representative examples reported in the literature and are not reproduced under a unified experimental setting.

Deployment scenario	Key constraints	Representative DL techniques	Rationale	Relevant Studies
High mobility (vehicular / high Doppler)	Short coherence time; rapidly time-varying CSI; need prediction/tracking	CNN–BiLSTM predictors, temporal CNN–RNN estimators, transformer-based forward prediction	Sequence models and transformers capture fast temporal variations; multi-head attention helps track Doppler-induced non-stationarity	[129], [136], [116]
Low-latency edge systems (URLLC / edge receiver)	Strict inference latency; limited compute/memory; predictable runtime	Model-driven / unfolded DL; lightweight denoisers; compact MLP refiners	Unfolding yields fixed-depth inference; lightweight CNNs reduce compute while preserving accuracy	[150], [125], [93]
RIS-dense environments (large N_{RIS}, cascaded channels)	High-dimensional cascaded CSI; pilot overhead; scalability with RIS size	CNN/ResNet denoisers; U-Net super-resolution; compressive sensing + DL refinement	Cascaded channels behave like structured images; CNNs/U-Nets reconstruct dense CSI from sparse pilots	[158], [165]
Hardware impairment (e.g., low-resolution ADC, non-linear RF chains)	Quantization distortion; RF nonlinearity; mismatch between uplink/downlink channels	DNN-based calibration; CNN denoisers; generative models for quantization recovery	DL can learn non-linear hardware distortions and interpolate missing information that SP-based estimators cannot recover	[143], [144], [145]

employed to further refine channel estimates. Notable designs in this category include an investigation by Wei *et al.* [130], in which they proposed GM-LAMP that integrates traditional iterative methods with DL-based denoising modules to achieve optimal results, particularly in low-SNR regimes. Similarly, the investigation by Zia *et al.* [32] employed signal processing (a higher-dimensional Fourier transform followed by the EM algorithm) to extract channel parameters, and then used an MLP to refine the parameter estimates. To conclude, these hybrid approaches balance computational efficiency and adaptability while keeping interpretability to a moderate level.

2) Models that have a dominant signal processing part (also known as model-driven deep learning) use signal processing as the primary mechanism, with DL playing a supplementary role in enhancing performance. These models typically employ enhanced conventional methods to process the received signal and extract features, which are then fed into simpler DL models for refined output. Prominent research work in this domain includes the investigation by Ma *et al.* [150], who proposed using a deep neural network for joint pilot and estimator design in wideband massive MIMO systems. Another investigation in this context is that of Elbir *et al.* [146], which explores deep learning-based designs for channel estimation and hybrid beamforming in frequency-selective, wideband mm-Wave systems. Signal processing dominant approaches are particularly useful in systems where computational resources are not limited (suitable for real-time signal processing) or where the emphasis is on improving specific aspects of the system.
3) Models with limited or no signal processing (also known as data-driven deep learning) eliminate the reliance on traditional signal processing by directly feeding the received signal into DL architectures. These models aim to learn end-to-end mappings from raw received signals to channel estimates, allowing them to capture the non-linearities and complexities of real-world channels adaptively. Key research efforts in this area include Melgar *et al.* [27], which focused on the limitations of LS and MMSE estimators, noting inadequate performance and the unavailability of the covariance matrix in LS and MMSE estimators, respectively. Another investigation conducted by Zhang *et al.* [131] focused on FCDNet to extract channel information from noisy inputs, particularly in cases when SNR levels are significantly low. DL-dominant models excel in scenarios where domain knowledge is limited, or channels are moderately dynamic, and their statistics can be captured in a database.

Each category highlights a unique approach to integrating DL into massive MIMO systems, offering trade-offs among complexity, interpretability, and performance. Lessons from these classifications suggest that the choice of approach should align with the application's specific requirements, including resource availability, channel conditions, and system scalability. The data analysis that can suggest the level of DL involvement to optimize performance across diverse scenarios presents a key research direction.

However, it should be noted that the performance gains offered by deep learning-based estimators are not achieved without trade-offs. Unlike LS and MMSE estimators, which rely on channel covariance information and matrix inversion operations, DL-based estimators typically require access to a sufficiently large and representative training dataset prior to deployment [25], [189]. In other words, deep learning shifts part of the estimation burden from explicit model knowledge and matrix-based processing to offline data collection and training [189]. This trade-off is particularly attractive in massive MIMO and RIS-assisted systems, where accurate covariance acquisition, reliable channel modeling, and repeated matrix inversion become increasingly challenging. In such settings, a DL-based solution may offer a more practical alternative, although its effectiveness remains strongly dependent on the quality, diversity, and scale of the training data [27].

Moreover, DL-based estimators commonly employ the mean square error loss function and learn a nonlinear mapping from input observations to desired outputs by iteratively adjusting weights and biases through nonlinear activation functions. From this perspective, a DL estimator can be viewed as a nonlinear data-driven mechanism for minimizing mean square error over the available training set, with the expectation that the learned mapping will generalize well to unseen future data [27], [189]. This is fundamentally different from conventional linear estimators, since the optimization is not restricted to a closed-form model but is instead carried out implicitly through training, allowing the network to capture complex relationships among noisy received signals and transmitted symbols.

A critical requirement for the practical deployment of DL-assisted channel estimation is robustness across diverse environments. Many existing DL models demonstrate high performance in specific scenarios but may fail to generalize when the channel conditions deviate from those used during training. Robustness can be achieved through [190]–[192] :

- Training on Diverse Datasets: Models trained on datasets representing a wide range of channel conditions, user distributions, and hardware imperfections are more likely to generalize.
- Domain Adaptation Techniques: These techniques enable models to adapt to new or unseen domains by bridging the gap between the source domain (training environment) and the target domain (deployment environment). Domain adaptation improves model performance in the target domain without requiring extensive retraining.
- Data Augmentation and Transfer Learning: Techniques such as transfer learning and domain-specific fine-tuning allow pre-trained models to adapt to new environments with minimal additional training.
- Incorporating Robust Features: Hybrid models often enhance robustness by combining domain knowledge with DL.

Lessons from these studies suggest that while DL-aided channel estimation holds significant potential, several challenges remain in its way to becoming a key enabler of 6G. Many DL methods require extensive training datasets, and their performance heavily depends on the representativeness of training data. Moreover, scalability and interpretability are still critical concerns, particularly in systems with high-dimensional channel matrices. Future research must aim to strike a balance between complexity and performance, exploring hybrid architectures that combine domain knowledge with data-driven insights.

2) RIS-Aided Massive MIMO Systems.

The potential of RIS-aided MIMO systems hinges on addressing the critical challenge of accurate and efficient channel estimation [193]. This section examines several key issues related to DL-aided channel estimation in RIS-enabled mMIMO systems. In RIS-aided mMIMO systems, the optimization of phase shifts requires precise estimates of both the BS-RIS and RIS-user channels. These estimates are crucial for feeding optimization algorithms that dynamically adjust the RIS phase shifts to enhance system performance [17]. Traditional channel estimation methods often rely on extensive signal processing techniques to extract the necessary CSI at the base station, which can be computationally expensive and time-consuming. In contrast, deep learning techniques offer distinct advantages in scenarios where conventional signal processing methods struggle. By leveraging the power of neural networks, DL-based approaches can efficiently estimate and separate complex, high-dimensional channels from a single observation, thereby significantly reducing computational overhead [17]. A fully passive RIS, which reflects incident signals through adjustable phase shifts, offers a cost-effective and energy-efficient solution for enhancing coverage and mitigating signal blockage in 6G networks. While passive RIS eliminates the need for RF chains and complex signal processing, the deployment of DL-based channel estimation in RIS-enabled mMIMO systems necessitates the use of extensive datasets and advanced training strategies to effectively model the high-dimensional, cascaded links (BS-RIS-UE).

The high-dimensional nature of the cascaded channels, due to sequential channel coefficients and scaling with the large number of BS antennas, RIS elements, users, and carriers, further complicates the estimation process. Traditional methods, in this context, become computationally prohibitive. Consequently, deep learning techniques are increasingly proposed to address this complexity by reducing dimensionality prior to or during estimation. Key strategies include employing sparsity assumptions in compressive sensing approaches [193], utilizing autoencoders for compression and joint estimation [193], applying denoising through CNN variants [159], employing tensor decomposition to reduce rank [193], and implementing federated learning to minimize training efforts per user by utilizing shared global knowledge [169]. Furthermore, estimating the high-dimensional cascaded channels often requires schemes that involve switching RIS elements on/off or cycling through different phase configurations [193], which increases the number of pilot signals necessary during training. Convolutional neural networks offer a promising solution by treating the sparse pilot responses as low-resolution images and applying super-resolution techniques to reconstruct the full channel matrix [159]. CNNs are particularly well-suited for this task due to their ability to extract local patterns, process high-dimensional spatial data, and perform effective denoising as pilots traverse the cascaded channels. This capability aligns seamlessly with the structure of RIS channels, effectively addressing the inherent challenges in channel estimation. Many deep learning-based channel estimation techniques treat the channel matrix as a 2D/3D image, neglecting its inherent graph structure and permutation invariance. Recent studies have applied graph models, such as graph attention networks and transformers, to map pilot signals to a graph, improving robustness and estimation accuracy. These methods are effective in capturing long-range dependencies across RIS elements, users, and time slots. However, they remain underexplored, as they require significant computational resources for training [194], [195].

Active RIS-aided MIMO systems are particularly effective in highly disruptive and time-varying environments. In such conditions, traditional channel estimation techniques, or even static neural networks, quickly become outdated. Retraining or online adaptation is costly [169], and adaptive model design via federated learning is much needed for dynamic channels. However, most existing deep learning models assume static channels or rely on slow-fading blocks, a limitation that is rarely addressed in contemporary research [183]. With the integration of active RIS elements in hybrid and STAR RIS configurations, it becomes possible to directly measure the BS-RIS and RIS-UE channels independently. Here, the problem becomes that of power consumption and cost rather than complexity. Active RIS is useful for long-range or high SNR use cases where passive RIS cannot provide sufficient link budget [196].

It is important to note that due to the emergence of state-of-the-art RIS surfaces such as stacked RIS, Flexible RIS, or Cell-free RIS, this topic is potentially in its infancy. These innovative configurations, which are more aligned with the 6G vision, remain under-researched in the literature but hold significant potential to meet the evolving demands of next-generation multi-antenna systems [197].

IV. Deep Learning-aided Detection in Multi-Antenna Systems

The large-scale deployment of 5G and 6G multi-antenna systems has introduced significant challenges in the signal detection domain [9], [16]. The fundamental objective of signal detection is to accurately estimate the transmitted signal vector $\mathbf{x}$ from the received signal vector $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ represents the channel matrix, which is obtained from the channel estimation module, and $\mathbf{n}$ denotes additive white Gaussian noise (AWGN). While traditional mMIMO detection algorithms [16], [198] demonstrate strong performance for small-scale systems, their computational complexity becomes prohibitive when applied to massive MIMO configurations where N_r and N_t may number in the hundreds. Thus, there is a need for low-complexity signal detection schemes that can perform well and scale to large system dimensions.

Recent advancements have shown that DL-enabled detection techniques can significantly improve performance in several scenarios involving RIS, STAR RIS, and hybrid RIS configurations. These DL-driven methods enhance the detection accuracy, robustness, and energy efficiency in environments with complex propagation conditions and large antenna arrays [41]. By leveraging the potential of RIS and hybrid RIS, these techniques offer a promising solution to address the scalability and complexity issues inherent in massive MIMO systems, providing a practical approach to next-generation communication networks [162], [199].

A. DL-aided massive-MIMO detection

Deep learning has emerged as a promising solution to the massive MIMO detection problem, offering two principal advantages: firstly, the ability to learn complex channel environments through data-driven and model-driven approaches [200], and secondly, the ability to develop existing detection algorithms into trainable neural architectures [201]. However, significant challenges remain in designing architectures that strike a balance between performance, complexity, and generalizability. Employing deep learning-aided detectors can be prohibitively complex for MIMO systems, for example, a transmitter with four antennas supporting the 64-QAM alphabet requires a 16.7 x 10^6 candidate comparison for ML detection [9].

Samuel *et al.* [200] introduced a fully connected DL-based detector dubbed "DetNet" characterized by a simple architecture that was highly efficient on fixed channels, where the channel matrix $\mathbf{H}$ is deterministic, and Rayleigh fading channels. It also achieved near-optimal performance with low modulation orders while maintaining low computational complexity. However, the design was data-driven, which relied on a significantly larger set of parameters, typically ranging from 1 to 10 million. This large parameter space necessitates extensive offline training, often involving long training times, to ensure the model's accuracy. In a bid to achieve better results, Datta *et al.*

proposed a collaborative learning-based detection technique for mMIMO systems, employing multiple fully connected DNNs in conjunction with an iterative meta-predictor to minimize detection errors. Simulation results show that the proposed technique significantly outperforms conventional methods such as Stair Matrix Detection (SMD) [202], in terms of BER performance. In the study conducted by Tan *et al.* [203], DNN was utilized to enhance belief propagation (BP) detection algorithm for mMIMO systems under both i.i.d. Rayleigh and correlated fading channels, with varying antenna configurations. The research explored several DNN-based variants of BP detection algorithms, including the Deep Neural Network-Simplified Message Passing Detector (DNN-sMPD), Deep Neural Network-Damping Belief Propagation (DNN-dBP), and Deep Neural Network-Max-Sum (DNN-MS). These novel implementations significantly improved detection performance by utilizing the powerful feature-extraction capabilities of DNNs, while maintaining robustness in diverse channel conditions. In the study by Liao *et al.* [42], a model-driven deep learning method was proposed for massive MIMO detection, utilizing a DNN to unfold an iterative algorithm for multiuser interference cancellation. The model incorporated two auxiliary parameters at each layer to better manage residual errors and adjust layer relationships, effectively enhancing the detection process. Simulation results indicated that the proposed method outperformed conventional detectors like LMMSE and deep learning aided detectors DetNet [200], achieving superior BER performance with lower complexity. Kumar *et al.* [204] proposed a deep learning-based method for signal detection in mMIMO systems utilizing optical non-orthogonal multiple access (O-NOMA). They employed a DNN to address interference issues caused by overlapping signals in O-NOMA systems. The proposed DNN outperformed traditional detection techniques, achieving a BER of 10^{-3} at an SNR of 4 dB, which was significantly better compared to the 7 dB performance of the conventional MMSE detector.

Ahmad *et al.* [205] proposed a DL-supported channel estimation and detection method for wavelet-based massive MIMO-NOMA systems. The proposed DNN was designed to estimate and detect symbols with or without pilot signals, mitigating pilot contamination and reducing transmission overhead. The results showed that the DNN-based method improved the successive interference cancellation-based detection, providing a 4 dB gain over traditional fast Fourier transform (FFT)-based NOMA systems in terms of symbol error rate.

Yu *et al.* [206] proposed a data-driven DNN dubbed “AMIC-Net”, for uplink massive MIMO detection. The network utilized a sparsely connected DNN structure and introduced an extrapolation factor as a learnable parameter to enhance the convergence speed and robustness of the iterative multiuser interference cancellation algorithm. The results demonstrated that AMIC-Net achieved nearly the same detection accuracy as OAMP-Net [207], but with at least 10 times lower computational complexity, especially in high-order QAM modulation.

Keawin *et al.* [208] proposed a CNN-based signal detection algorithm for ultra-massive MIMO systems, integrating multiple machine learning algorithms such as extreme learning machine (ELM), regularized extreme learning machine (RELM), and outlier-robust extreme learning machine (ORELM) for enhanced signal detection accuracy. Their proposed method demonstrated improved performance in terms of MSE and BER, and channel capacity compared to traditional methods like ZF and MMSE.

Yao *et al.* [209] proposed a model-driven DL detector network, termed “Block Gauss-Seidel Network (BGS-Net)”, to address the high computational complexity associated with traditional Gauss-Seidel methods in mMIMO systems. The BGS-Net combined DL with a Gauss-Seidel iterative method, reducing large matrix inversions into smaller ones for efficient detection. Simulation results showed that BGS-Net outperformed Gauss-Seidel and MMSE in terms of SER while maintaining lower complexity.

Zhang *et al.* [210] proposed an efficient DNN-supported AMP detection algorithm for mMIMO systems. The AMP-DNN detector significantly reduced BER without requiring prior signal information, especially in spatially correlated channels, while maintaining low computational complexity. The computational complexity of the proposed AMP-DNN was claimed to be lower as compared to other detectors, such as DetNet [200].

Kumar *et al.* [211] introduced the successive interference cancellation with reinforcement learning (SIC-RL) detector for mMIMO–NOMA systems, significantly improving signal detection under Rayleigh fading channels. The proposed SIC-RL outperforms traditional detectors such as MMSE and ML, achieving SNR gains of 11.2 dB for 512-QAM and 6.6 dB for 256-QAM at a BER of 10^{-3}, while minimizing spectral leakage by 35%.

Brennsteiner *et al.* [212] proposed LAMANet, an approximate message passing-based detector enhanced by deep learning to significantly improve detection performance while minimizing computational complexity. By integrating learnable matrices within the AMP algorithm, they mitigated performance degradation in untrained detectors, thereby eliminating the need for extensive online training as required by previous approaches. Empirical results demonstrate that LAMANet achieves performance comparable to MMNet [213], while using the simpler AMP algorithm that effectively reduces preprocessing complexity. Li *et al.* [214] introduced a learning-based symbol detection algorithm for mMIMO-OFDM systems using Multi-Mode StructNet, which integrates multi-mode reservoir computing with real-time learning via the alternating recursive least squares algorithm. The algorithm adapts to dynamic environments and compensates for nonlinear distortions caused by transceiver

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

circuits, outperforming traditional methods like LMMSE in terms of BER.

Ye *et al.* [215] introduced "ChannelNet", a purely data-driven learning-based mMIMO detector, which utilizes off-the-shelf neural network components to overcome scalability bottlenecks in previous data-driven methods. By incorporating antenna-wise shared feature processors and channel-embedded layers, ChannelNet achieves permutation invariance and equivariance, significantly reducing computational complexity. Empirical results demonstrate that ChannelNet consistently outperforms or matches the performance of state-of-the-art detectors, including AMP and MMSE, across various antenna configurations and modulation.

In the domain of sparsely connected networks, Gao *et al.* [216] addressed the issue of the exponential number of parameters needed by DetNet [200]. Their proposed design was a more simplified version of the DetNet architecture dubbed "Sparsely connected neural Network (ScNet)" which significantly reduced the number of inputs and network layers by using sparse connectivity instead of full connectivity, while maintaining comparable BER performance. Another fast-convergence, sparsely connected detection network (FS-Net) was proposed by Nguyen *et al.* [217]. FS-Net was obtained by optimizing the DetNet [200] and ScNet [216] designs. The FS-Net simplified network connections to reduce complexity. The loss function considered the correlation between each layer's output and the desired solution to achieve faster convergence and better performance than conventional detectors. Jin *et al.* [218] proposed a DL-based interference cancellation-aided signal detection method, using the enhanced sparsely connected network (EScNet) for mMIMO systems. They introduced two novel detection algorithms, EScNet-PIC and EScNet-SIC, which reduce the number of training parameters and significantly improve detection performance. The complexity of the proposed EScNet was lower than that of the fully connected DetNet model [200]. The experimental results demonstrated that EScNet-SIC outperforms DetNet [200] by achieving a performance gain of 2.4 dB at a BER of 10^{-3}.

In the work by Khani *et al.* [213], MMNet, a deep learning-based architecture, was proposed for signal detection in mMIMO systems, utilizing an iterative soft-thresholding approach combined with a nonlinear denoising step. This design incorporates an online training algorithm that used the spectral and temporal locality of channel matrices, resulting in substantial reduction of computational complexity. MMNet achieved a performance within 2dB of the optimal ML detector while requiring 10 times less computational effort. Table 4 compares the detection performance of notable works in this domain.

In essence, massive-MIMO DL-based detectors broadly follow two patterns, i.e., unfolding model-based algorithms into trainable architectures, DetNet [200], DNN-BP variants [203], FS-Net [217], BGS-Net [209], MMNet [213], and data-driven architectures such as AMIC-Net [206], AMP-DNN [210], and StructNet [214]. While unfolded architectures provide stability and interpretable complexity, data-driven detectors report stronger BER gains under non-linearities and imperfect CSI. However, discrepancies arise due to heterogeneous assumptions: some studies consider fixed or quasi-static channels [200], while others evaluate correlated, dynamic, or interference-limited scenarios [42], [205]. Additionally, some works consider DL inference cost when reporting complexity [210], whereas others focus solely on iterative components [217], making cross-paper efficiency comparisons inconsistent. These divergences underscore the need for standardized experimental environments when assessing DL-enabled massive-MIMO detectors.

B. DL-aided RIS massive-MIMO detection

Recent advancements in deep learning techniques have significantly contributed to the development of efficient detection methods for RIS-assisted massive MIMO systems. In particular, the research concerning DL-aided detection in RIS mMIMO includes a pioneering study by Rahman *et al.* [41], which introduced a novel Bi-LSTM-based model for channel estimation and signal detection in a RIS multiuser MISO OFDM system. The researchers trained their proposed DNN using two distinct optimization algorithms, highlighting the effectiveness of this approach compared to existing DNN models. A comparative analysis of their method with four alternative DNN models demonstrated its superior performance in detecting signals, as well as its robustness against residual estimation errors.

Wang *et al.* [222] explored the joint design of multi-antenna receivers and RIS phase shifts, specifically targeting the uplink of RIS-assisted mMIMO systems in the absence of complete CSI information. Their DL-based receiver, referred to as "DeepSIC RX", employed multiple DNNs, where the configurations of RIS reflection patterns, which significantly influence the propagation channel, were treated as hyper-parameters. Their simulation results, conducted under non-linear propagation environment and the presence of AWGN, demonstrated the robustness of their approach.

In addition to these efforts, Singh *et al.* [224] proposed an RIS-assisted generalized spatial modulation (RIS-GSM) scheme aimed at improving the signal reception and spectral efficiency. Their approach utilized a block-based DNN design to develop a low-complexity symbol detector for RIS-GSM MIMO systems. Authors compared their proposed design with conventional signal detectors, such as the maximum likelihood, block zero-forcing, and block minimum mean square error detectors, and a better performance was shown as evidence of this design's effectiveness. Another work in the context of RIS-assisted GSM is presented by Liu *et al.* [219], which presented a novel complex-valued DNN and complex-valued CNN for signal detection. Unlike traditional real-valued neural networks that process real and imaginary components separately, their approach directly handles complex-valued signals, preserving the inherent

Table 4: The Computational Complexity vs Performance of DL-Based Detectors

Category	Research	Architecture	SNR (dB)	BER	Computational Complexity
Massive MIMO	Zhang *et al.* [210]	AMP-DNN	2 to 17	10^{-2} to 10^{-5}	O(LN_tN_r)
	Gao *et al.* [216]	ScNet	4 to 12	10^{-2} to 10^{-4}	O($3N_t$ L)
	Nguyen *et al.* [217]	FS-Net	4 to 12	10^{-2} to 10^{-5}	O($8N_t^2$ + $8N_t$) L
	Jin *et al.* [218]	EScNet-SIC	4 to 16	10^{-1} to 10^{-3}	O($(L+2N_t-1)N_r^2+(3L+2N_t-1)N_r$)
	Li *et al.* [214]	StructNet	20 to 28	10^{-2} to 10^{-4}	O($p\kappa+\kappa o$)
	Ye *et al.* [215]	ChannelNet	11 to 15	10^{-1} to 10^{-3}	O(N_tN_r)
	Khani *et al.* [213]	MMNet	9 to 14	10^{-2} to 10^{-4}	O(bN_t^2L)
	Tan *et al.* [203]	DNN-dBP	4 to 14	10^{-1} to 10^{-4}	O(MN_tN_rL)
RIS-assisted Massive MIMO	Liu *et al.* [219]	CVNN	-8 to 4	10^{-1} to 10^{-4}	$O(NN_u((2N_rN_u+2N_r-1)2\delta_1+\sum_{k=1}^{E-1}\delta_{k+1}(2\delta_k-1)))$
	Sagir *et al.* [220]	DNN-CRIS	-48 to 32	10^{-1} to 10^{-4}	$O\left(pn_1+on_\kappa+\sum_{k=1}^{\kappa-1}n_kn_{k+1}\right)$
	Azizi *et al.* [221]	DFSM-Det	10 to 20	10^{-1} to 10^{-3}	($O(N^2+6N)$ additions $O(N^2+5N)$ multiplications)/layer
	Wang *et al.* [222]	DNN	-12 to 2	10^{-1} to 10^{-5}	–
	Jin *et al.* [223]	DNN-JSDc	-20 to -10	10^{-1} to 10^{-4}	$O(2^{B_i}(4N_r\delta_2+\sum_{q=2}^{Q-2}\delta_q\delta_{q+1})+\delta_{Q-1}M+2N_rN+N_r)$

The variables in the computational complexity column are as follows: N_t represent the number of antennas at the transmitter, N_r represent the number of antennas at the receiver, N_u is the number of transmit antennas activated in a timeslot, L is the number of layers in neural network, F is the number of feature maps, R is the filter size, M is the order of the modulation scheme, p is the number of inputs and κ is the number of hidden layers of the multilayer neural network, and the k^{th} hidden layer has n_k neurons with o outputs neurons or number of outputs. $B_i = \lfloor \log_2 L \rfloor$ (opposite phase reflection pattern index bits, $L = \binom{N_g}{K}$), δ_q are neurons in layer q (total layers Q), M is modulation order, $N = N_g \times N_p$ is total RIS elements, and N_g is RIS groups. All SNR ranges, and BER values are taken from the respective original papers and were not reproduced under a unified experimental setting; therefore, the comparison should be interpreted as indicative rather than a strictly controlled head-to-head benchmark.

structure of communication data. The proposed detectors demonstrated near-optimal BER performance, approaching ML detection, while significantly reducing computational complexity. The proposed detectors exhibited strong robustness under Nakagami-m fading channels and imperfect CSI, outperforming conventional MMSE detectors by 7 dB and real-valued neural network alternatives by 2 dB.

Xiang *et al.* [162] proposed a novel deep learning-based symbol detection technique for RIS-assisted massive MIMO systems. The approach, termed PreRec-Correlator-Net, used deep learning to enhance detection reliability in the presence of imperfect CSI and interference. This method significantly reduced the BER, improving system performance compared to conventional techniques that neglect multi-user interference. The simulations showed that the deep learning-aided detection mechanism outperforms traditional methods by achieving a lower error floor, mitigating the effect of multi-user interference.

In the investigation by Ioannou *et al.* [225], authors proposed an advanced deep learning-based methodology for RIS beam management, focusing on the optimization of beamforming and beam steering through the use of feedback DNN. Simulation results revealed that the proposed method achieved superior performance, particularly in high-mobility scenarios, with a notable reduction in pilot overhead. In the study by Abid *et al.* [199] a deep learning-based signal detection method for RIS-aided OTFS systems was proposed. Investigators suggested a DNN-based detector to improve signal detection in doubly selective, high-speed communication channels. This novel approach effectively mitigated performance degradation caused by NLOS conditions, enhancing BER performance. Simulation results demonstrated that the DNN-based detector outperforms conventional detectors like ZF and MMSE achieving a lower BER in challenging high-mobility environments.

Similarly, Liu *et al.* [226] extended the application of RIS-aided spatial modulation and space shift keying by incorporating DL techniques for detection. By employing the concept of deep unfolding, they examined model-driven deep learning detection methods tailored specifically for RIS-supported systems. Liu *et al.* also compared the performance of their proposed methodology with data-driven DL detectors. This study, alongside the aforementioned contributions, emphasized the substantial potential of DL techniques for enhancing the detection capabilities of RIS-aided MIMO systems, particularly in scenarios where the received signal is corrupted by residual estimation errors, multi-user interference, and Gaussian noise.

In [227], Du *et al.* proposed a semi-blind joint channel estimation and symbol detection scheme for RIS-empowered multiuser mmWave systems, based on the PARATUCK2 tensor model. This scheme effectively minimized the symbol detection and channel estimation errors, showing superior performance compared to existing methods. The simulation results further validated that the proposed semi-blind approach offers improved detection accuracy in RIS-aided multiuser mmWave systems, providing a compelling solution for enhancing detection capabilities in complex communication environments.

Two DNN-assisted cooperative RIS models, namely, $\mathrm{DNN_R}$ and $\mathrm{DNN_{R,D}}$, were proposed for IoT-based wireless networks by Sagir *et al.* in [220]. The key novelty of this investigation lied in the integration of DNN to optimize RIS phase shifts and replace ML detection with DNN-based symbol estimation, reducing complexity while maintaining performance. The authors demonstrated that their DNN-optimized RIS relays achieve BER close to ML detector, even in high-noise environments, with a 7dB SNR gain for 32 reflecting elements. Notably, the $\mathrm{DNN_{R,D}}$ cooperative RIS model reduced computational overhead significantly compared to traditional ML/MMSE detectors. Xiong *et al.* [228] proposed a novel framework for separate channel estimation and data detection in semi-passive RIS-aided millimeter wave massive MIMO systems. Their key contribution was the development of a double-layer bilinear generalized approximate message passing (DL-BiGAMP) algorithm, which utilized partial active elements at the RIS to first obtain coarse estimates of the user-to-RIS and RIS-to-base station channels, and then refines these estimates while simultaneously recovering data symbols. Numerical results demonstrated that the proposed DL-BiGAMP method significantly outperforms existing techniques by improving BER, especially at higher SNRs.

In [223], Jin *et al.* proposed RIS-based opposite phase index modulation (RIS-OPIM) scheme. Their proposed construct enhanced spectral efficiency and reliability by dividing RIS elements into groups that reflect signals either in-phase or with opposite-phase shifts, maximizing the Euclidean distance between reflection patterns while maintaining full aperture gain. The authors derived a theoretical upper bound for BER under ML detection, validated by simulations showing superior performance over existing RIS-based modulation methods. A machine learning-based approach for joint symbol detection and RIS reflection optimization was proposed by Wang *et al.* [222], considering a passive RIS environment. Their contribution was focused on treating RIS phase configurations as hyperparameters and optimizing them alongside a DNN-based receiver dubbed "DeepSIC" using Bayesian optimization. Their proposed method alternates between training the DeepSIC receiver for multi-user detection and tuning RIS reflections via Bayesian optimization, utilizing pilot transmissions to iteratively improve system performance.

In [221], Azizi *et al.* introduced DFSM-Det, a novel DL-based symbol detector for RIS-assisted mMIMO systems, employing first- and second-momentum mechanisms to enhance detection performance. Unlike existing methods, DFSM-Det incorporated outputs from three preceding layers to progressively refine symbol estimates, achieving faster convergence and improved robustness in high-load scenarios where the number of transmit antennas approaches receive antennas. Their proposed method transforms complex-valued signals into real-valued models and employs a trainable activation function tailored for M-QAM modulation. Simulation results demonstrate that DFSM-Det outperforms OAMP-Net [207] in terms of BER performance, all while maintaining low computational complexity, especially under high system loading conditions.

Adding to these efforts, Khan *et al.* [229] introduced a novel deep learning-based detection approach for RIS-enabled massive MIMO dubbed "DeepRIS". This construct serves a dual role, symbol detection and channel estimation, utilizing RIS-reflected signals. The key contribution of their work was focused on the development of a deep learning-based detection method that eliminates the need for pilot signals, thereby significantly reducing transmission overhead. Simulation results demonstrated that the DeepRIS framework has a significantly improved performance over traditional detection techniques, achieving low BER performance, even in the presence of imperfect CSI. These results highlighted the potential of DL constructs in overcoming the limitations of conventional methods, particularly in handling residual estimation errors more efficiently. Table III presents the complexity vs performance comparison of several key DL-enabled detectors.

Overall, DL-enabled RIS detection approaches generally augment classical detection with neural refiners to counter cascaded-channel errors, residual interference, and hardware distortion, as seen in BiLSTM-based detectors [41], DeepSIC-RX [222], RIS-GSM DNN detectors [224], and complex-valued CNN/DNN constructions [219]. Architectures diverge significantly, while some prioritize temporal modeling or phase-pattern learning, such as PreRec-Correlator-Net [162] and feedback-based beam management [225], while others integrate joint channel estimation–detection pipelines [227], [229]. Moreover, lightweight detectors (e.g., DNN-CRIS [220]) emphasize reduced complexity, while more expressive models, e.g., graph-transformer schemes or OPIM-based detection [223] target maximal BER performance. Varied assumptions on channel availability, RIS activation patterns, and pilot overhead contribute to heterogeneous benchmarking, highlighting the need for unified evaluation frameworks.

C. Critical analysis and lessons learned on DL-aided detection

Signal detection in multi-antenna systems encounters similar challenges for mMIMO and RIS-aided mMIMO due to the analogous operational processes at the transmitter/receiver of both variants. In light of this, this subsection will focus on the analysis of deep learning-enabled detection techniques for both configurations.

The key benefit of using deep learning for massive MIMO detection lies in its ability to handle high-dimensional data and complex channels that traditional methods cannot efficiently process. While traditional detection schemes, such as ZF, MMSE and ML detection, achieve optimal accuracy, their high computational complexity renders them impractical for multi-antenna systems [28]. DL models can effectively model the non-linearities of the mMIMO channels and the inter-user interference in multi-user scenarios [162]. These models learn the optimal detection strategy directly from data, eliminating the need for explicit mathematical modeling of the channel or noise, which simplifies the implementation.

DL-supported signal detection in multi-antenna systems introduces several key challenges, particularly in adapting deep learning-enabled signal processing to real-world scenarios. A primary concern is the substantial amount of training data required (that is not always readily available), necessary to build effective DL models [9]. Training deep networks on large-scale mMIMO datasets demands significant computational resources, which may be impractical for real-time applications, especially as the number of antennas increases. Furthermore, while DL-based methods can improve performance in many cases, they remain highly dependent on the quality of the training data, making them less robust in dynamic, unpredictable environments where channel conditions frequently change. This cost-benefit trade-off has driven innovation in DL techniques, although these innovations come with their own unique set of challenges.

DetNet [200], a pioneering deep learning-based detector, has shown significant promise in enhancing detection performance. However, despite its advantages, it introduces substantial computational complexity, particularly during the offline training phase, which can limit its practical applicability in large-scale systems. One major issue regarding DetNet [200] and FS-Net [217] is the convergence of DNN models for realistic channel conditions, such as spatially correlated channels. This underscores the need for further optimization and customization of DNN-based models to handle real-world complexities [230] effectively.

Several researchers have highlighted the power of data-driven deep learning approaches for solving complex signal detection problems in mMIMO systems [206], [214], [230]. These methods offer significant performance improvements, especially in environments with imperfect CSI or interference. For example, to deal with imperfect CSI, user interference, and channel noise, Ahmad *et al.* [205] adopted a wavelet-based solution combined with DL for advanced channel estimation and symbol detection in mMIMO-NOMA systems.

Model-driven deep learning is beneficial for DL-aided detection because it combines the strengths of conventional signal processing techniques with the adaptability and performance enhancements offered by deep learning. Model-driven techniques, such as [162], strike a balance by using fewer trainable parameters, enabling rapid training with smaller datasets. Another example in this regard is the design presented by Yao *et al.* [209] to address shortcomings of the Gauss-Seidel iterative algorithm. Their proposed BGS-Net reduced the computational burden, simplified complex matrix inversions, and improved symbol error rate without sacrificing detection performance. Similarly, the innovative hybrid approach "DFSM-Det", proposed by Azizi *et al.* [221], combined elements of both model-driven and data-driven DL for intelligent symbol detection in RIS mMIMO systems. This design combines conventional model-based optimization techniques, such as momentum, within a deep learning framework, thereby leveraging the strengths of both paradigms to improve performance, reduce computational complexity, and enhance robustness, particularly in highly loaded scenarios.

Li *et al.* [214] highlighted the challenge of real-time learning in massive MIMO systems, emphasizing the importance of adapting to rapidly changing channels and hardware impairments. The lesson learned here is the importance of domain-specific knowledge in improving both training efficiency and detection accuracy in large-scale systems. The ability of deep learning models to generalize across different modulation schemes, antenna numbers, and channel conditions is crucial for practical deployment in evolving communication networks. This is demonstrated by the flexibility of the AMIC-Net and DetNet models in adapting to various settings [206], [210].

Despite the substantial performance gains provided by deep learning, managing computational complexity remains a critical challenge. Sparse neural networks, such as AMIC-Net [206], and the learnable extrapolation factor in AMP-DNN [210] are essential in reducing the complexity of these models while ensuring high accuracy.

Lessons that can be learned from these studies are multifold. The need for low-complexity yet scalable solutions is evident. Sparse architectures and hybrid methods offer promising avenues to address this trade-off, but further optimization and real-world validation are critical. The robustness of DL models under real-world channel conditions remains a key challenge. Future work should prioritize models that adapt dynamically to varying interference patterns, channel statistics, and hardware constraints. Employing domain-specific features, such as temporal and frequency correlations, is critical for enhancing detection performance in practical environments. Techniques that combine channel estimation and signal detection highlight the benefits of integrating traditionally distinct processes. Such frameworks can help mitigate errors introduced at different stages of the communication pipeline.

Overall, while significant strides have been made in mMIMO detection, the field must address real-world challenges, including scalability, robustness, and practical feasibility, to fully realize the potential of these advanced detection techniques.

V. Research Trends and Future Directions

The integration of DL into multi-antenna systems, particularly in the estimation and detection domain, has shown promising results in improving system performance. In the

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

following, we will discuss the state-of-the-art trends and future directions.

- Academia and industry should focus on developing deep learning models that are robust to environmental changes and adaptable to diverse RIS configurations. Real-world wireless channels are highly dynamic and unpredictable, and RIS-assisted massive MIMO systems often operate in non-stationary environments. Research should aim to develop models that can address these challenges and maintain good performance under diverse conditions [231].
- Deep learning has also helped in the context of mobile and satellite-aided RIS, offering a substantial extension in coverage. By using deep learning models, researchers are developing advanced algorithms capable of optimizing RIS configurations in real-time, improving signal coverage area with reliability and efficiency [232].
- Channel Foundation Models (CFMs) have gained significant attention in wireless communication research, particularly for their potential to optimize channel-related tasks. These models utilize large-scale, pre-trained architectures, including generative, discriminative, and hybrid approaches, to efficiently capture complex channel dynamics. Their ability to generalize across diverse channel conditions and operational environments makes them highly adaptable for 6G systems. The development of CFMs has opened avenues for improving tasks such as channel estimation, signal detection, and beamforming optimization, thereby addressing challenges posed by dynamic environments, such as urban areas and high-speed mobile scenarios. Moving forward, research will focus on enhancing their scalability, data efficiency, and fine-tuning mechanisms to support the growing complexity of next-generation wireless systems [233].
- Quantum machine learning (QML) is gaining momentum in the field of RIS, especially for channel estimation and tracking in complex systems such as STAR-RIS. Recent advances propose modular QML frameworks that use quantum RNNs to handle the high dimensionality of channel data and track user mobility. These frameworks, benefiting from quantum parallelism and noise-reduction techniques, have been shown to outperform classical methods, particularly in terms of convergence speed and accuracy in dynamic environments. Future directions emphasize further refining QML models to mitigate imperfections in quantum computing and exploring their integration in cutting-edge multi-antenna systems [234].
- The integration of DL for channel estimation in next-generation wireless networks, especially 6G, holds significant promise, but faces security challenges, particularly adversarial attacks on pilot input signals. Recent studies focus on mitigating these threats through innovative techniques, such as distillation, which enhance the robustness of DL models to perturbations. Research indicates that employing advanced architectures can improve the reliability of channel estimation under adversarial conditions. Future directions will likely include refining distillation methods and exploring novel defensive strategies to ensure the secure deployment of AI-based channel estimation models [235].
- Researchers are increasingly exploring scalable online learning approaches for channel estimation, wherein deep learning models are leveraged to continuously refine and enhance their estimates based on real-time channel measurements. A significant area of focus is the integration of feedback mechanisms that enable the system to incorporate the accuracy of previous estimates, thereby enabling more precise and adaptive tracking of channel variations. This approach holds promise for improving the robustness and performance of communication systems, particularly in dynamic and fluctuating environments [236].
- Recent innovations in federated learning for channel estimation and detection include clustered FL frameworks with blockchain for reduced communication overhead in RIS-enabled systems [237], hierarchical FL networks for improved performance in nonuniform user distributions [238], model-driven FL approaches for better accuracy in mmWave MIMO systems [239], decentralized FL with graph neural networks for collaborative learning across base stations [240], and federated deep reinforcement learning for ISAC systems [241]. These advancements focus on enhancing estimation accuracy, minimizing overhead, and enabling efficient decentralized learning in complex multi-antenna systems.
- At present, no official and universally adopted standard exists within major telecommunication standardization bodies such as 3GPP and ETSI for specific deep learning models used in channel estimation. Instead, current research is largely directed toward developing techniques and frameworks that may later be incorporated into future standards or deployed within the flexible layers of existing communication systems. At the same time, the absence of standardized datasets, reproducible benchmarks, and open-source implementations remains a major obstacle to fair and reliable comparison of DL-assisted systems, as inconsistencies in data and methodologies hinder the ability to draw meaningful, unbiased, and reproducible conclusions. Therefore, the standardization of datasets, benchmarks, and evaluation practices is becoming increasingly important for consistent performance assessment across models and approaches, providing a unified and comparable foundation for advancing deep-learning-based estimation and detection in multi-antenna systems and facilitating their practical adoption in real-world scenarios.
- Recent advancements in RIS highlighted that pixel errors and other hardware impairments significantly

degrade system performance. In response, deep learning techniques have emerged as a promising solution for mitigating these hardware limitations. [242]. Researchers are increasingly focusing on developing deep learning algorithms capable of detecting and correcting pixel-level distortions. Looking ahead, future research will likely focus on creating more resilient models that can dynamically adapt to and counteract hardware impairments.
- In scenarios where labeled data is scarce or difficult to obtain, researchers are increasingly turning to semi-supervised and unsupervised learning techniques for channel estimation and detection. These methods can utilize both labeled and unlabeled data, making them highly effective in improving the accuracy of channel estimation in environments where pilot signals are limited. Techniques such as hard inference learning (HIL) and Gaussian inference learning (GIL) have shown promising results in mMIMO systems [243].
- Energy-efficient and lightweight deep learning models are essential for multi-antenna systems involving small, battery-powered IoT edge devices. Future research should focus on developing deep learning architectures optimized for energy efficiency, specifically tailored for resource-constrained IoT devices [244].
- Hybrid estimation and detection approaches, which integrate traditional signal processing techniques with deep learning or employ multiple deep learning modules for tasks such as estimation or detection, such as twin CNN (TCNN) [245], have shown significant promise in achieving more accurate and robust results. These approaches are particularly beneficial in the context of 6G networks, where high performance is critical to pushing the boundaries of current technological capabilities. Future research should focus on identifying and optimizing such hybrid designs to meet the demanding requirements of 6G, thereby advancing network capacity, reliability, and real-time processing.

In summary, the above-presented challenges & future research directions in DL-enabled multi-antenna systems are diverse and multidimensional. Addressing these challenges requires a comprehensive, well-structured strategy that aligns with current needs while being forward-looking and adaptable to future advancements. To guide the research community about actionable directions that necessitate immediate attention as well as those requiring long-term continuous research efforts, we prioritized the research directions across near, mid, and long-term horizons.

In the near term, the most pressing needs include: (i) the establishment of standardized datasets and reproducible benchmarks to mitigate inconsistencies in reported results caused by varying training datasets and model assumptions, (ii) the development of low-complexity, hardware-aware DL models that meet stringent latency and memory constraints, particularly for IoT devices, and (iii) enhancing the robustness of DL models to rapidly time-varying channels, high interference, and low-SNR regions, through innovative approaches such as online learning, federated learning, and attention-based learning techniques. In the mid-term, research efforts should focus on: (i) the creation of a unified DL framework that integrates joint sensing, localization, and communication in ISAC systems [50], (ii) scalable distributed or federated learning strategies for RIS-dense deployments [169], (iii) developing DL techniques that are resilient to hardware impairments, particularly RIS pixel errors [242], and (iv) establishing standardized frameworks for countering adversarial attacks.

In the long term, the research landscape is expected to shift toward i) the development of Channel Foundation Models [233] trained on large-scale, cross-environment datasets, ii) exploration of quantum machine learning for high-dimensional signal processing [234], and iii) the development of autonomous networks that leverage reinforcement learning and meta-learning to self-optimize pilot overhead, resource allocation, and propagation environments without explicit human intervention.

VI. Conclusion

The rapid evolution of deep learning has significantly reshaped the landscape of massive multi-antenna systems, particularly in CSI estimation and symbol detection. This paper has explored various deep learning-based approaches, emphasizing their ability to address the inherent challenges of mMIMO and RIS-mMIMO. As wireless communication systems evolve towards 6G, the integration of DL frameworks offers a promising avenue to improve system performance, particularly in complex and highly disruptive environments where traditional methods struggle to deliver accurate results.

The rapid growth of RIS technology, characterized by an increasing number of deployed reflective panels (ranging from single to multi-surfaces) and a growing multiplicity of reflected propagation paths, has given rise to a highly dense and dynamically reconfigurable radio environment. This emerging paradigm calls for an in-depth investigation into scalable, efficient design methodologies for dense RIS deployments, particularly those that leverage multi-hop reflections through joint optimization of cooperative active and passive beamforming and intelligent data routing strategies.

Despite the massive scale of the data and associated training challenges, DL-enabled solutions show substantial promise, particularly for optimizing the estimation and detection capabilities required for large-scale RIS deployments. Advances in DL models, ranging from MLPs and CNNs to more advanced architectures such as U-Net and Transformers, have demonstrated their ability to outperform conventional techniques by leveraging the ability to learn from large datasets, yielding robust performance even in non-stationary environments. Moreover, coupling CSI estimation with symbol detection via DL-based methods has highlighted their interdependence, underscoring the need for integrated approaches that simultaneously balance both aspects.

Looking ahead, this review has identified several promising future directions, including the potential of federated learning, online learning, and transfer learning for scalable and adaptive solutions, as well as the application of large language models for dynamic channel prediction. The adoption of these advanced learning paradigms will be critical to overcoming the challenges posed by massive MIMO and RIS systems, particularly in real-time, distributed environments. Additionally, the continued development of hybrid models that integrate DL with conventional signal processing holds substantial promise for enhancing the robustness and efficiency of 6G multi-antenna systems.

In essence, this paper provides a comprehensive overview of the current state of the art in DL-assisted multi-antenna systems, while also highlighting research gaps and potential innovations that will drive the next wave of advancements in wireless communication technology. We hope this paper will serve as a valuable and inspiring resource for future research on DL-enabled multi-antenna systems to unlock the full potential of future-generation wireless communications.

References

[1] F. A. Pereira de Figueiredo, "An Overview of Massive MIMO for 5G and 6G," *IEEE Latin America Transactions*, vol. 20, no. 6, pp. 931–940, 2022.

[2] O. Elijah, S. K. Abdul Rahim, W. K. New, C. Y. Leow, K. Cumanan, and T. Kim Geok, "Intelligent Massive MIMO Systems for Beyond 5G Networks: An Overview and Future Trends," *IEEE Access*, vol. 10, pp. 102 532–102 563, 2022.

[3] A. Almomani and F. Al-Turjman, "Challenges and Opportunities in Integrated 6G and IoT Paradigms: An Overview," in *2022 International Conference on Artificial Intelligence in Everything (AIE)*, 2022, pp. 140–145.

[4] Z. Wang, J. Zhang, H. Du, E. Wei, B. Ai, D. Niyato, and M. Debbah, "Extremely large-scale MIMO: Fundamentals, challenges, solutions, and future directions," *IEEE Wireless Communications*, 2023.

[5] Z. Liu, L. Zhang, and Z. Ding, "Overcoming the Channel Estimation Barrier in Massive MIMO Communication via Deep Learning," *IEEE Wireless Communications*, vol. 27, no. 5, pp. 104–111, 2020.

[6] P. K. Gkonis, "A Survey on Machine Learning Techniques for Massive MIMO Configurations: Application Areas, Performance Limitations and Future Challenges," *IEEE Access*, vol. 11, pp. 67–88, 2023.

[7] S. F. Drampalou, N. I. Miridakis, H. C. Leligou, and P. A. Karkazis, "A survey on optimal channel estimation methods for RIS-aided communication systems," *Signals*, vol. 4, no. 1, pp. 208–234, 2023.

[8] G. N. Reddy, C. V. Ravikumar, and A. Rajesh, "Literature review and research direction towards channel estimation and hybrid pre-coding in mmwave massive MIMO communication systems," *J. Reliab. Intell. Environ.*, Apr. 2022.

[9] M. A. Albreem, M. Juntti, and S. Shahabuddin, "Massive MIMO detection techniques: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3109–3132, 2019.

[10] Y. Yu, J. Wang, X. Zhou, C. Wang, Z. Bai, and Z. Ye, "Review on channel estimation for reconfigurable intelligent surface assisted wireless communication system," *Mathematics*, vol. 11, no. 14, p. 3235, 2023.

[11] M. Tariq, S. Ahmad, M. Ahmad Jan, and H. Song, "Deep Learning Aided Intelligent Reflective Surfaces for 6G: A Survey," *ACM Computing Surveys*, vol. 57, no. 3, p. 63, 2025.

[12] H. Ma, J. Du, H. Wu, L. Xing, and R. Zheng, "A review of channel estimation research methods for massive MIMO systems," *Physical Communication*, vol. 71, p. 102632, 2025.

[13] M. A. Albreem, W. Salah, A. Kumar, M. H. Alsharif, A. H. Rambe, M. Jusoh, and A. N. Uwaechia, "Low Complexity Linear Detectors for Massive MIMO: A Comparative Study," *IEEE Access*, vol. 9, pp. 45 740–45 753, 2021.

[14] J. Wang, W. Tang, Y. Han, S. Jin, X. Li, C.-K. Wen, Q. Cheng, and T. J. Cui, "Interplay Between RIS and AI in Wireless Communications: Fundamentals, Architectures, Applications, and Open Research Problems," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2271–2288, 2021.

[15] H. Xie, F. Gao, and S. Jin, "An Overview of Low-Rank Channel Estimation for Massive MIMO Systems," *IEEE Access*, vol. 4, pp. 7313–7321, 2016.

[16] S. Yang and L. Hanzo, "Fifty Years of MIMO Detection: The Road to Large-Scale MIMOs," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 1941–1988, 2015.

[17] B. Zheng, C. You, W. Mei, and R. Zhang, "A Survey on Channel Estimation and Practical Passive Beamforming Design for Intelligent Reflecting Surface Aided Wireless Communications," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 1035–1071, 2022.

[18] K. M. Faisal and W. Choi, "Machine Learning Approaches for Reconfigurable Intelligent Surfaces: A Survey," *IEEE Access*, vol. 10, pp. 27 343–27 367, 2022.

[19] P. Varshney, R. P. Singh, and R. Jain, "A Systematic Review of Challenges in Channel Estimation for 5G Massive MIMO," in *2024 1st International Conference on Sustainable Computing and Integrated Communication in Changing Landscape of AI (ICSCAI)*, 2024, pp. 1–5.

[20] Y. Komati and S. Rajendran, "A Comprehensive Study on Challenges and Solutions in Massive MIMO-Detection Methods," in *2024 3rd Edition of IEEE Delhi Section Flagship Conference (DELCON)*, 2024, pp. 1–7.

[21] M. Jian, G. C. Alexandropoulos, E. Basar, C. Huang, R. Liu, Y. Liu, and C. Yuen, "Reconfigurable intelligent surfaces for wireless communications: Overview of hardware designs, channel models, and estimation techniques," *Intelligent and Converged Networks*, vol. 3, no. 1, pp. 1–32, 2022.

[22] M. Ahmed, F. Xu, A. Wahid, K. Ali, M. A. Mirza, W. Khan, K. Dev, S. A. Hassan, and Z. Han, "A comprehensive survey of artificial intelligence advances in RIS-assisted wireless networks," *Authorea Preprints*, 2024.

[23] B. Li, Q. Zheng, X. Tian, M. Yang, G. Gui, W. Jiang, H. Lei, J. Jiang, F. Shu, A. Elhanashi, and S. Saponara, "A survey of artificial intelligence enabled channel estimation methods: Recent advance, performance, and outlook," *Artif. Intell. Rev.*, vol. 58, no. 6, Apr. 2025.

[24] M. A. Albreem, A. Alhabbash, A. M. Abu-Hudrouss, and T. A. Almohamad, "Data detection in decentralized and distributed massive MIMO networks," *Computer Communications*, vol. 189, pp. 79–99, 2022.

[25] N. K. Kundu and M. R. McKay, "Channel Estimation for Reconfigurable Intelligent Surface Aided MISO Communications: From LMMSE to Deep Learning Solutions," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 471–487, 2021.

[26] S. N. Sur and R. Bera, "Intelligent reflecting surface assisted MIMO communication system: A review," *Physical Communication*, vol. 47, p. 101386, 2021.

[27] A. Melgar, A. de la Fuente, L. Carro-Calvo, O. Barquero-Perez, and E. Morgado, "Deep Neural Network: An Alternative to Traditional Channel Estimators in Massive MIMO Systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 657–671, 2022.

[28] M. A. Albreem, A. H. Alhabbash, S. Shahabuddin, and M. Juntti, "Deep learning for massive MIMO uplink detectors," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 741–766, 2021.

[29] X. Wei, D. Shen, and L. Dai, "Channel Estimation for RIS Assisted Wireless Communications—Part I: Fundamentals, Solutions, and Future Opportunities," *IEEE Communications Letters*, vol. 25, no. 5, pp. 1398–1402, 2021.

[30] M. Ahmed, S. Raza, A. Amin Soofi, F. Khan, W. Ullah Khan, S. Zain Ul Abideen, F. Xu, and Z. Han, "Active Reconfigurable Intelligent Surfaces: Expanding the Frontiers of Wireless Communication-A Survey," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 2, pp. 839–869, 2025.

[31] H. Dahrouj, R. Alghamdi, H. Alwazani, S. Bahanshal, A. A. Ahmad, A. Faisal, R. Shalabi, R. Alhadrami, A. Subasi, M. T. Al-Nory, O. Kittaneh, and J. S. Shamma, "An Overview of Machine Learning-Based Techniques for Solving Optimization Problems in Communications and Signal Processing," *IEEE Access*, vol. 9, pp. 74 908–74 938, 2021.

[32] M. U. Zia, W. Xiang, G. M. Vitetta, and T. Huang, "Deep Learning for Parametric Channel Estimation in Massive MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 4, pp. 4157–4167, 2023.

[33] M. Alrabeiah and A. Alkhateeb, "Deep Learning for TDD and FDD Massive MIMO: Mapping Channels in Space and Frequency," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 1465–1470.

[34] J. Guo, C.-K. Wen, and S. Jin, "CAnet: Uplink-Aided Downlink Channel Acquisition in FDD Massive MIMO Using Deep Learning," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 199–214, 2022.

[35] B. Warhade and R. Patil, "Massive MIMO Issues and Solutions-An Overview," in *2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, 2024, pp. 1–6.

[36] S. Basharat, S. A. Hassan, H. Pervaiz, A. Mahmood, Z. Ding, and M. Gidlund, "Reconfigurable Intelligent Surfaces: Potentials, Applications, and Challenges for 6G Wireless Networks," *IEEE Wireless Communications*, vol. 28, no. 6, pp. 184–191, 2021.

[37] R. Q. Shaddad, E. M. Saif, H. M. Saif, Z. Y. Mohammed, and A. H. Farhan, "Channel Estimation for Intelligent Reflecting Surface in 6G Wireless Network via Deep Learning Technique," in *2021 1st International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, 2021, pp. 1–5.

[38] A. M. Elbir, A. Papazafeiropoulos, P. Kourtessis, and S. Chatzinotas, "Deep channel learning for large intelligent surfaces aided mm-wave massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 9, no. 9, pp. 1447–1451, 2020.

[39] Y. Jin, J. Zhang, X. Zhang, H. Xiao, B. Ai, and D. W. K. Ng, "Channel estimation for semi-passive reconfigurable intelligent surfaces with enhanced deep residual networks," *IEEE transactions on vehicular technology*, vol. 70, no. 10, pp. 11 083–11 088, 2021.

[40] M. Liu, X. Li, B. Ning, C. Huang, S. Sun, and C. Yuen, "Deep Learning-Based Channel Estimation for Double-RIS Aided Massive MIMO System," *IEEE Wireless Communications Letters*, 2022.

[41] M. H. Rahman, M. A. S. Sejan, M. A. Aziz, J.-I. Baik, D.-S. Kim, and H.-K. Song, "Deep Learning Based Improved Cascaded Channel Estimation and Signal Detection for Reconfigurable Intelligent Surfaces-Assisted MU-MISO Systems," *IEEE Transactions on Green Communications and Networking*, pp. 1–1, 2023.

[42] J. Liao, J. Zhao, F. Gao, and G. Y. Li, "A Model-Driven Deep Learning Method for Massive MIMO Detection," *IEEE Communications Letters*, vol. 24, no. 8, pp. 1724–1728, 2020.

[43] E. C. Vilas Boas, J. D. e Silva, F. A. de Figueiredo, L. L. Mendes, and R. A. de Souza, "Artificial intelligence for channel estimation in multicarrier systems for B5G/6G communications: a survey," *EURASIP Journal on Wireless Communications and Networking*, vol. 2022, no. 1, p. 116, 2022.

[44] Y. Liu, Y. Zhang, M. Liu, M. D. Renzo, Y. Wu, M.-S. Alouini, and Y. Zhang, "Reconfigurable Intelligent Surfaces: Principles and Opportunities," *IEEE Communications Surveys & Tutorials*, 2021, early Access.

[45] H. Lu, Y. Zeng, C. You, Y. Han, J. Zhang, Z. Wang, Z. Dong, S. Jin, C.-X. Wang, T. Jiang, X. You, and R. Zhang, "A Tutorial on Near-Field XL-MIMO Communications Toward 6G," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 4, pp. 2213–2257, 2024.

[46] W.-C. Wu and G.-L. Chen, "Low-Complexity Detection for MIMO Visible Light Communication System Using Generalized Spatial Modulation," *EURASIP Journal on Advances in Signal Processing*, vol. 2025, no. 5, 2025.

[47] W. Xu, J. An, X. Zhou, X. You, Y. Zhang, and G. Huang, "STAR-RISs: Simultaneous Transmission and Reflection Reconfigurable Intelligent Surfaces for Wireless Communication," *IEEE Communications Letters*, vol. 25, no. 9, pp. 3134–3138, 2021.

[48] V. Dala Pegorara Souto, P. S. Dester, M. Soares Pereira Facina, D. Gomes Silva, F. A. P. de Figueiredo, G. Rodrigues de Lima Tejerina, J. C. Silveira Santos Filho, J. Silveira Ferreira, L. L. Mendes, R. D. Souza, and P. Cardieri, "Emerging MIMO Technologies for 6G Networks," *Sensors*, vol. 23, no. 4, p. 1921, 2023.

[49] Z. Wang, J. Zhang, H. Du, D. Niyato, S. Cui, B. Ai, M. Debbah, K. B. Letaief, and H. V. Poor, "A Tutorial on Extremely Large-Scale MIMO for 6G: Fundamentals, Signal Processing, and Applications," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 3, pp. 1560–1605, 2024.

[50] F. Li and B. Liao, "Massive MIMO-ISAC beamforming design via sensing energy maximization," in *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2025, pp. 1–5.

[51] Y. Zhang, S. Jin, Y. Jing, and X. Dong, "Joint Radar and Communication Designs: Applications in Integrated Sensing and Communication (ISAC) Systems," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 1, pp. 178–193, 2023.

[52] D. P. M. Osorio, B. Barua, K.-L. Besser, H. Blue, P. Dass, and P. Porambage, "The Rise of Networked ISAC: Emerging Aspects and Challenges," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 5072–5091, 2025.

[53] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-Free Massive MIMO: A New Next-Generation Paradigm," *IEEE Access*, vol. 7, pp. 99 878–99 888, 2019.

[54] J. Zheng, J. Zhang, H. Du, D. Niyato, B. Ai, M. Debbah, and K. B. Letaief, "Mobile Cell-Free Massive MIMO: Challenges, Solutions, and Future Directions," *IEEE Wireless Communications*, vol. 31, no. 3, pp. 140–147, 2024.

[55] Y. Lin, S. Jin, M. Matthaiou, and X. Yi, "Spherical RIS-Enabled Channel Estimation and User Self-Localization for ISAC Systems," *IEEE Transactions on Communications*, pp. 1–1, 2025.

[56] P. Nidagundi, A. Darvekar, M. Amber, and R. R, "A comparative study of ultra-massive MIMO intelligent receivers with adversarial robustness and energy efficiency for 6G applications," *Results in Engineering*, vol. 26, p. 105409, 2025.

[57] M. Li, Z. Yuan, Y. Lyu, P. Kyösti, J. Zhang, and W. Fan, "Gigantic MIMO Channel Characterization: Challenges and Enabling Solutions," *IEEE Communications Magazine*, vol. 61, no. 10, pp. 140–146, 2023.

[58] R. Feng, C.-X. Wang, J. Huang, and X. Gao, "Recent Advances of 6G Ultra-Massive MIMO Technologies in Spatial and Beam Domains," *arXiv*, 2025, arXiv:2501.10429.

[59] L.-S. Tsai, S.-L. Shih, P.-K. Liao, and C.-K. Wen, "MIMO evolution toward 6g: End-user-centric collaborative MIMO," *IEEE Communications Magazine*, vol. 62, no. 7, pp. 104–110, 2024.

[60] C.-K. Wen, Y.-C. Chan, T.-H. Huang, H.-J. Zeng, F.-K. Wang, L.-S. Tsai, and P.-K. Liao, "End-User-Centric Collaborative MIMO: Performance Analysis and Proof of Concept," *IEEE Transactions on Wireless Communications*, 2025.

[61] S. N. Sur, "Massive MIMO," in *Advances in Greener Energy Technologies*, A. K. Bhoi, K. S. Sherpa, A. Kalam, and G.-S. Chae, Eds. Singapore: Springer Singapore, 2020, pp. 833–842.

[62] Ericsson, "Ericsson Unveils New Wave of Radio, Antenna, and RAN Products for Sustainable 6G," , 2025.

[63] S. Liesegang, S. Buzzi, and C. D'Andrea, "Scalable Integrated Sensing and Communications for Multi-Target Detection and Tracking in Cell-Free Massive MIMO: A Unified Framework," *arXiv*, 2025, arXiv:2503.06703.

[64] W. Tang, M. Zeng, Y. Xu, W. Ma, and R. Zhang, "Wireless Communications With Reconfigurable Intelligent Surface: Path Loss Modeling and Experimental Measurement," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 598–602, 2021.

[65] L. Zhang, M. Chen, W. Liu, M. D. Renzo, and H. V. Poor, "RIS-Aided Wireless Communications: From Surface Reflection to Space-Time Modulation," *IEEE Wireless Communications*, vol. 29, no. 6, pp. 130–137, 2022.

[66] J. Wang, Y. Li, Y. Xia, X. Li, D. W. K. Ng, and O. A. Dobre, "Active STAR-RIS-Aided Wireless Powered Communication Networks," *IEEE Transactions on Wireless Communications*, vol. 25, pp. 3406–3421, 2026.

[67] Y. Zhang and et al., "Stacked Reconfigurable Intelligent Surfaces for 6G: Architecture and Beamforming Strategies," *IEEE Transactions on Communications*, vol. 71, no. 6, pp. 3712–3725, 2023.

[68] S. Kiani and Q. H. Abbasi, "Graphene-Based Reconfigurable Intelligent Surfaces for 6G Terahertz Communication," *IEEE Transactions*

on Terahertz Science and Technology, vol. 13, no. 2, pp. 234–246, 2023.

[69] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable Intelligent Surfaces for Energy Efficiency in Wireless Communication," *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 464–480, 2022.

[70] T. J. Cui, S. Liu, and L. Zhang, "Information Metasurfaces for Advanced Wireless Communications," *Nature Reviews Materials*, vol. 8, pp. 41–57, 2023.

[71] U. Zia, I. H. Naqvi, and M. Uppal, "Performance of ultra-wideband time-reversal systems under imperfect channel side information at the transmitter," *Electronics Letters*, vol. 51, no. 7, pp. 585–586, 2015. [Online]. Available:

[72] F. Pasic, M. Hofer, M. Mussbah, S. Caban, S. Schwarz, T. Zemen, and C. F. Mecklenbräuker, "Channel Estimation for mmWave MIMO Using Sub-6 GHz Out-of-Band Information," in *2024 International Conference on Smart Applications, Communications and Networking (SmartNets)*, 2024, pp. 1–6.

[73] U. Zia, M. Uppal, and I. H. Naqvi, "Robust feedback design for time reversal UWB communication systems under CSIT imperfections," *IEEE Communications Letters*, vol. 19, no. 1, pp. 102–105, 2015.

[74] K. Zhi, C. Pan, G. Zhou, H. Ren, M. Elkashlan, and R. Schober, "Is RIS-Aided Massive MIMO Promising With ZF Detectors and Imperfect CSI?" *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 10, pp. 3010–3026, 2022.

[75] X. Yang, C.-K. Wen, and S. Jin, "MIMO Detection for Reconfigurable Intelligent Surface-Assisted Millimeter Wave Systems," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1777–1792, 2020.

[76] J. Chai, H. Zeng, A. Li, and E. W. Ngai, "Deep learning in computer vision: A critical review of emerging techniques and application scenarios," *Machine Learning with Applications*, vol. 6, p. 100134, 2021. [Online]. Available:

[77] M. Tsuneki, "Deep learning models in medical image analysis," *Journal of Oral Biosciences*, vol. 64, no. 3, pp. 312–320, 2022.

[78] M. U. Zia, W. Xiang, T. Huang, J. Ahmad, J. N. Chattha, I. H. Naqvi, and F. A. Butt, "Unifying ground and air: a comprehensive review of deep learning-enabled CAVs and UAVs," *Artificial Intelligence Review*, vol. 59, no. 1, p. 19, 2026.

[79] C. S. Babu and P. Akshara, "Revolutionizing conversational AI: unleashing the power of ChatGPT-Based applications in generative AI and natural language processing," in *Advanced applications of generative AI and natural language processing models*. IGI Global Scientific Publishing, 2024, pp. 228–248.

[80] Y. Liu, Y. Li, Y. Zhu, Y. Niu, and P. Jia, "A Brief Review on Deep Learning in Application of Communication Signal Processing," in *2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP)*, 2020, pp. 51–54.

[81] M. Arnold, S. Dörner, S. Cammerer, J. Hoydis, and S. ten Brink, "Towards Practical FDD Massive MIMO: CSI Extrapolation Driven by Deep Learning and Actual Channel Measurements," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 1972–1976.

[82] D. Löschenbrand, M. Hofer, L. Eller, M. Rupp, and T. Zemen, "Machine Learning-Based Channel Prediction for Widely Distributed Massive MIMO with Real-World Data," in *2023 57th Asilomar Conference on Signals, Systems, and Computers*, 2023, pp. 982–987.

[83] J.-M. Kang, "Deep-Learning-Based Robust Channel Estimation for MIMO IoT Systems," *IEEE Internet of Things Journal*, vol. 11, no. 6, pp. 9882–9895, 2024.

[84] O. Toutsop, T. M. Yimer, K. Kornegay, F. Donfack, and M. Tienteu, "Lightweight Machine Learning-Deep Learning Framework for IoT Devices," in *2025 IEEE World AI IoT Congress (AIIoT)*, 2025, pp. 1093–1101.

[85] D. Thakur, J. K. Saini, and S. Srinivasan, "DeepThink IoT: the strength of deep learning in internet of things," *Artificial Intelligence Review*, vol. 56, no. 12, pp. 14 663–14 730, 2023.

[86] K. S. Zaman, M. B. I. Reaz, S. H. Md Ali, A. A. A. Bakar, and M. E. H. Chowdhury, "Custom Hardware Architectures for Deep Learning on Portable Devices: A Review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 11, pp. 6068–6088, 2022.

[87] A. Salh, L. Audah, N. S. M. Shah, A. Alhammadi, Q. Abdullah, Y. H. Kim, S. A. Al-Gailani, S. A. Hamzah, B. A. F. Esmail, and A. A. Almohammedi, "A Survey on Deep Learning for Ultra-Reliable and Low-Latency Communications Challenges on 6G Wireless Systems," *IEEE Access*, vol. 9, pp. 55 098–55 131, 2021.

[88] S. S. Jabarani and J. Jacob, "Comparative Analysis of MLP-Based and Traditional Channel Estimation Strategies for Millimeter Wave MIMO Networks," in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, 2024, pp. 1–6.

[89] I. Ioannou, P. Nagaradjane, A. Khalifeh, R. Dhivyaharshini, A. Harini, and V. Vassiliou, "Deep Learning-Aided Channel Estimation for Large Scale Fading Decoding in Massive MIMO Systems," in *2024 IEEE 30th International Conference on Telecommunications (ICT)*, 2024, pp. 01–08.

[90] M. U. Zia, W. Xiang, T. Huang, and I. Haider Naqvi, "Deep Learning-Aided TR-UWB MIMO System," *IEEE Transactions on Communications*, vol. 70, no. 10, pp. 6579–6588, 2022.

[91] P. Dong, H. Zhang, G. Y. Li, N. NaderiAlizadeh, and I. S. Gaspar, "Deep CNN for Wideband Mmwave Massive MIMO Channel Estimation Using Frequency Correlation," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4529–4533.

[92] Y. Jin, J. Zhang, B. Ai, and X. Zhang, "Channel Estimation for mmWave Massive MIMO With Convolutional Blind Denoising Network," *IEEE Communications Letters*, vol. 24, no. 1, pp. 95–98, 2020.

[93] P. Jiang, C.-K. Wen, S. Jin, and G. Y. Li, "Dual CNN-Based Channel Estimation for MIMO-OFDM Systems," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5859–5872, 2021.

[94] T. Peng, R. Zhang, X. Cheng, and L. Yang, "LSTM-Based Channel Prediction for Secure Massive MIMO Communications Under Imperfect CSI," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[95] W. Xu, J. An, Y. Xu, C. Huang, L. Gan, and C. Yuen, "Time-varying channel prediction for RIS-assisted MU-MISO networks via deep learning," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 4, pp. 1802–1815, 2022.

[96] C. Nguyen, T. M. Hoang, and A. A. Cheema, "Channel Estimation Using CNN-LSTM in RIS-NOMA Assisted 6G Network," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 1, pp. 43–60, 2023.

[97] S. Ravula and S. Jain, "Deep Autoencoder-based Massive MIMO CSI Feedback with Quantization and Entropy Coding," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[98] Y. Jang, G. Kong, M. Jung, S. Choi, and I.-M. Kim, "Deep Autoencoder Based CSI Feedback With Feedback Errors and Feedback Delay in FDD Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 833–836, 2019.

[99] V. Kumar, D. Mercado-Perez, F. Ye, R. Q. Hu, and Y. Qian, "A Reconstructed Autoencoder Design for CSI Processing in Massive MIMO Systems," in *ICC 2024 - IEEE International Conference on Communications*, 2024, pp. 5111–5116.

[100] R. R. Singh, K. Mawatwal, and B. J. Prakash, "An Autoencoder Based Channel Estimate Denoising Network (CEDNet) for OFDM Systems," in *2025 IEEE 22nd Consumer Communications & Networking Conference (CCNC)*, 2025, pp. 1–6.

[101] H. Ye, F. Gao, J. Qian, H. Wang, and G. Y. Li, "Deep Learning-Based Denoise Network for CSI Feedback in FDD Massive MIMO Systems," *IEEE Communications Letters*, vol. 24, no. 8, pp. 1742–1746, 2020.

[102] M. Hussien, "PRVNet: Variational autoencoders for massive MIMO CSI feedback," *Authorea Preprints*, 2023.

[103] A. Swain, S. M. Hiremath, and S. K. Patra, "Variational AutoEncoder Based CSI Feedback for Massive MIMO Systems," *Wireless Personal Communications*, pp. 1–18, 2023.

[104] S. Mourya, P. Reddy, S. Amuru, and K. K. Kuchi, "Spectral Temporal Graph Neural Network for Massive MIMO CSI Prediction," *IEEE Wireless Communications Letters*, vol. 13, no. 5, pp. 1399–1403, 2024.

[105] Y. Yang, S. Zhang, F. Gao, J. Ma, and O. A. Dobre, "Graph Neural Network-Based Channel Tracking for Massive MIMO Networks," *IEEE Communications Letters*, vol. 24, no. 8, pp. 1747–1751, 2020.

[106] M. Ye, X. Liang, C. Pan, Y. Xu, M. Jiang, and C. Li, "Channel Estimation for mmWave Massive MIMO Systems Using Graph Neural Networks," in *2023 IEEE/CIC International Conference on Communications in China (ICCC)*, 2023, pp. 1–5.

[107] Q. Peng, Y. Huang, and W. Zhou, "Graph Neural Network-Based Low-Complexity Channel Estimation for Massive MU-MIMO Systems," in *2024 9th International Conference on Computer and Communication Systems (ICCCS)*, 2024, pp. 623–628.

[108] B. Banerjee, R. C. Elliott, W. A. Krzymień, and H. Farmanbar, "Downlink Channel Estimation for FDD Massive MIMO Using Conditional Generative Adversarial Networks," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 122–137, 2023.

[109] M. Momen-Tayefeh, M. Momen-Tayefeh, S. A. G. Ghahramani, and A. M. A. Hemmatyar, "Channel estimation for Massive MIMO systems aided by intelligent reflecting surface using semi-super resolution GAN," *Signal Processing*, vol. 227, p. 109710, 2025.

[110] B. Tolba, M. Elsabrouty, M. G. Abdu-Aguye, H. Gacanin, and H. M. Kasem, "Massive MIMO CSI Feedback Based on Generative Adversarial Network," *IEEE Communications Letters*, vol. 24, no. 12, pp. 2805–2808, 2020.

[111] A. Faisal, I. Al-Nahhal, K. Lee, O. A. Dobre, and H. Shin, "Conditional Generative Adversarial Networks for Channel Estimation in RIS-Assisted ISAC Systems," *IEEE Transactions on Communications*, pp. 1–1, 2025.

[112] J. Zhao, Y. Wu, Q. Zhang, and J. Liao, "Two-Stage Channel Estimation for mmWave Massive MIMO Systems Based on ResNet-UNet," *IEEE Systems Journal*, vol. 17, no. 3, pp. 4291–4300, 2023.

[113] Al-Imran, M. S. Nazim, H. Nguyen, and Y. M. Jang, "Channel estimation of massive MIMO FSO communication system using deep attention residual U-Net," *ICT Express*, vol. 11, no. 2, pp. 287–292, 2025.

[114] W. Xie, J. Xiao, P. Zhu, C. Yu, and L. Yang, "Deep Compressed Sensing-Based Cascaded Channel Estimation for RIS-Aided Communication Systems," *IEEE Wireless Communications Letters*, vol. 11, no. 4, pp. 846–850, 2022.

[115] H. Ju, S. Jeong, S. Kim, B. Lee, and B. Shim, "Transformer-Assisted Parametric CSI Feedback for mmWave Massive MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 23, no. 12, pp. 18 774–18 787, 2024.

[116] H. Ju, S. Jeong, B. Lee, and B. Shim, "Transformer-based Predictive Channel Estimation for mmWave Massive MIMO Systems," in *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, 2024, pp. 1–5.

[117] Y. Jin, Y. Wu, Y. Gao, S. Zhang, S. Xu, and C.-X. Wang, "LinFormer: A Linear-based Lightweight Transformer Architecture For Time-Aware MIMO Channel Prediction," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2025.

[118] Y. Zeng, L. Qiao, Z. Gao, T. Qin, Z. Wu, E. Khalaf, S. Chen, and M. Guizani, "CSI-GPT: Integrating Generative Pre-Trained Transformer With Federated-Tuning to Acquire Downlink Massive MIMO Channels," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 3, pp. 5187–5192, 2025.

[119] Y. Wang, Z. Gao, D. Zheng, S. Chen, D. Gündüz, and H. V. Poor, "Transformer-Empowered 6G Intelligent Networks: From Massive MIMO Processing to Semantic Communication," *IEEE Wireless Communications*, vol. 30, no. 6, pp. 127–135, 2023.

[120] X. Fang, W. Feng, Y. Chen, N. Ge, and Y. Zhang, "Joint communication and sensing toward 6G: Models and potential of using MIMO," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4093–4116, 2022.

[121] S. D. Bast, A. P. Guevara, and S. Pollin, "CSI-based Positioning in Massive MIMO systems using Convolutional Neural Networks," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1–5.

[122] X. Chen, Z. Feng, J. Andrew Zhang, Z. Wei, X. Yuan, and P. Zhang, "Sensing-Aided Uplink Channel Estimation for Joint Communication and Sensing," *IEEE Wireless Communications Letters*, vol. 12, no. 3, pp. 441–445, 2023.

[123] D. Neumann, T. Wiese, and W. Utschick, "Deep Channel Estimation," in *WSA 2017; 21th International ITG Workshop on Smart Antennas*, 2017, pp. 1–6.

[124] X. Ma and Z. Gao, "Data-Driven Deep Learning to Design Pilot and Channel Estimator for Massive MIMO," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5677–5682, 2020.

[125] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-Based Channel Estimation for mmWave Massive MIMO Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 989–1000, 2019.

[126] M. Belgiovine, K. Sankhe, C. Bocanegra, D. Roy, and K. R. Chowdhury, "Deep Learning at the Edge for Channel Estimation in Beyond-5G Massive MIMO," *IEEE Wireless Communications*, vol. 28, no. 2, pp. 19–25, 2021.

[127] D. Yarotsky, A. Ivanov, R. Bychkov, A. Osinsky, A. Savinov, M. Trefilov, and V. Lyashev, "Machine Learning-Assisted Channel Estimation in Massive MIMO Receiver," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 2021, pp. 1–5.

[128] J. Gao, M. Hu, C. Zhong, G. Y. Li, and Z. Zhang, "An attention-aided deep learning framework for massive MIMO channel estimation," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1823–1835, 2021.

[129] Y. Liao, Y. Hua, and Y. Cai, "Deep Learning Based Channel Estimation Algorithm for Fast Time-Varying MIMO-OFDM Systems," *IEEE Communications Letters*, vol. 24, no. 3, pp. 572–576, 2020.

[130] X. Wei, C. Hu, and L. Dai, "Deep Learning for Beamspace Channel Estimation in Millimeter-Wave Massive MIMO Systems," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 182–193, 2021.

[131] Y. Zhang, Y. Mu, Y. Liu, T. Zhang, and Y. Qian, "Deep Learning-Based Beamspace Channel Estimation in mmWave Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 9, no. 12, pp. 2212–2215, 2020.

[132] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Deep Learning-Based Channel Estimation for Beamspace mmWave Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 852–855, 2018.

[133] L. Ge, Y. Guo, Y. Zhang, G. Chen, J. Wang, B. Dai, M. Li, and T. Jiang, "Deep Neural Network Based Channel Estimation for Massive MIMO-OFDM Systems With Imperfect Channel State Information," *IEEE Systems Journal*, vol. 16, no. 3, pp. 4675–4685, 2022.

[134] W. Gao, M. Yang, W. Zhang, Y. Zhao, Y. Zhou, and K. Zhang, "FFDNet-Based Channel Estimation for Beamspace mmWave Massive MIMO Systems," in *2021 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, 2021, pp. 1167–1170.

[135] D. Li, Y. Xu, M. Zhao, S. Zhang, and J. Zhu, "Knowledge-Driven Machine Learning-based Channel Estimation in Massive MIMO System," in *2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2021, pp. 1–6.

[136] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Machine Learning-Based Channel Estimation in Massive MIMO with Channel Aging," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2019, pp. 1–5.

[137] E. Balevi, A. Doshi, and J. G. Andrews, "Massive MIMO Channel Estimation With an Untrained Deep Neural Network," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2079–2090, 2020.

[138] A. Abboud, A. H. Jaber, J.-P. Cances, and V. Meghdadi, "CSI map for indoor massive MIMO," in *2017 Computing Conference*, 2017, pp. 1305–1311.

[139] H. Hirose, T. Ohtsuki, and G. Gui, "Deep Learning-Based Channel Estimation for Massive MIMO Systems With Pilot Contamination," *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 67–77, 2021.

[140] P. Sabeti, A. Farhang, I. Macaluso, N. Marchetti, and L. Doyle, "Blind Channel Estimation for Massive MIMO: A Deep Learning Assisted Approach," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[141] C. Chun, J. Kang, and I. Kim, "Deep Learning-Based Channel Estimation for Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 8, no. 4, pp. 1228–1231, Aug 2019.

[142] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep Learning for Super-Resolution Channel Estimation and DOA Estimation Based Massive MIMO System," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.

[143] C. Huang, G. C. Alexandropoulos, A. Zappone, C. Yuen, and M. Debbah, "Deep Learning for UL/DL Channel Calibration in Generic Massive MIMO Systems," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6.

[144] S. Gao, P. Dong, Z. Pan, and G. Y. Li, "Deep Learning Based Channel Estimation for Massive MIMO With Mixed-Resolution ADCs," *IEEE Communications Letters*, vol. 23, no. 11, pp. 1989–1993, Nov 2019.

This article has been accepted for publication in IEEE Open Journal of the Communications Society. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJCOMS.2026.3705258

[145] E. Balevi and J. G. Andrews, "Two-stage learning for uplink channel estimation in one-bit massive MIMO," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2019, pp. 1764–1768.

[146] A. M. Elbir, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, "A Family of Deep Learning Architectures for Channel Estimation and Hybrid Beamforming in Multi-Carrier mm-Wave Massive MIMO," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 642–656, 2022.

[147] W. Ma, C. Qi, Z. Zhang, and J. Cheng, "Deep Learning for Compressed Sensing Based Channel Estimation in Millimeter Wave Massive MIMO," in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2019, pp. 1–6.

[148] J. Gao, C. Zhong, G. Y. Li, and Z. Zhang, "Deep Learning-Based Channel Estimation for Massive MIMO With Hybrid Transceivers," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 5162–5174, 2022.

[149] J. Gao, C. Zhong, G. Y. Li, J. B. Soriaga, and A. Behboodi, "Deep Learning-Based Channel Estimation for Wideband Hybrid MmWave Massive MIMO," *IEEE Transactions on Communications*, vol. 71, no. 6, pp. 3679–3693, 2023.

[150] X. Ma, Z. Gao, F. Gao, and M. Di Renzo, "Model-Driven Deep Learning Based Channel Estimation and Feedback for Millimeter-Wave Massive Hybrid MIMO Systems," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2388–2406, 2021.

[151] C. Wen, W. Shih, and S. Jin, "Deep Learning for Massive MIMO CSI Feedback," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 748–751, Oct 2018.

[152] W. Ma, C. Qi, Z. Zhang, and J. Cheng, "Sparse Channel Estimation and Hybrid Precoding Using Deep Learning for Millimeter Wave Massive MIMO," *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 2838–2849, 2020.

[153] H. Sun, Z. Zhao, X. Fu, and M. Hong, "Limited Feedback Double Directional Massive MIMO Channel Estimation: From Low-Rank Modeling to Deep Learning," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018, pp. 1–5.

[154] B. Liu, X. Liu, S. Gao, X. Cheng, and L. Yang, "LLM4CP: Adapting Large Language Models for Channel Prediction," *Journal of Communications and Information Networks*, vol. 9, no. 2, pp. 113–125, 2024.

[155] S. Fan, Z. Liu, X. Gu, and H. Li, "Csi-LLM: A Novel Downlink Channel Prediction Method Aligned with LLM Pre-Training," *arXiv preprint arXiv:2409.00005*, 2024.

[156] F. O. Catak, M. Kuzlu, and U. Cali, "BERT4MIMO: A Foundation Model using BERT Architecture for Massive MIMO Channel State Information Prediction," *arXiv preprint arXiv:2501.01802*, 2025.

[157] W. Zheng, R. Wang, Y. Yuan, J. Li, B. He, and F. Lin, "Wireless Channel Estimation based on Transformer and Super-Resolution," in *Proceedings of the 2024 3rd Asia Conference on Algorithms, Computing and Machine Learning*, 2024, pp. 419–423.

[158] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, "Deep Residual Learning for Channel Estimation in Intelligent Reflecting Surface-Assisted Multi-User Communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 898–912, 2022.

[159] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "RIS-Aided mmWave MIMO Channel Estimation Using Deep Learning and Compressive Sensing," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3503–3521, 2023.

[160] H. Jiang, L. Dai, M. Hao, and R. MacKenzie, "End-to-End Learning for RIS-Aided Communication Systems," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6778–6783, 2022.

[161] N. Ginige, A. S. de Sena, N. H. Mahmood, N. Rajatheva, and M. Latva-aho, "Efficient Channel Prediction for Beyond Diagonal RIS-Assisted MIMO Systems with Channel Aging," *IEEE Transactions on Vehicular Technology*, pp. 1–15, 2025.

[162] W. Xiang, M. U. Zia, J. Ahmad, P. Cheng, K. Yu, and T. Huang, "Deep Learning-Enabled RIS Massive MIMO Systems for Industrial IoT: A Joint Communication and Computation Approach," *IEEE Journal on Selected Areas in Communications*, vol. 43, no. 9, pp. 2981–2996, 2025.

[163] A. L. Swindlehurst, G. Zhou, R. Liu, C. Pan, and M. Li, "Channel estimation with reconfigurable intelligent surfaces—A general framework," *Proceedings of the IEEE*, vol. 110, no. 9, pp. 1312–1338, 2022.

[164] W. Shen, Z. Qin, and A. Nallanathan, "Deep Learning for Super-Resolution Channel Estimation in Reconfigurable Intelligent Surface Aided Systems," *IEEE Transactions on Communications*, vol. 71, no. 3, pp. 1491–1503, March 2023.

[165] S. Liu, Z. Gao, J. Zhang, M. D. Renzo, and M.-S. Alouini, "Deep Denoising Neural Network Assisted Compressive Channel Estimation for mmWave Intelligent Reflecting Surfaces," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9223–9228, 2020.

[166] M. Xu, S. Zhang, C. Zhong, J. Ma, and O. A. Dobre, "Ordinary Differential Equation-Based CNN for Channel Extrapolation Over RIS-Assisted Communication," *IEEE Communications Letters*, vol. 25, no. 6, pp. 1921–1925, June 2021.

[167] K. Li, C. Huang, Y. Gong, and G. Chen, "Double Deep Learning for Joint Phase-shift and Beamforming based on Cascaded Channels in RIS-Assisted MIMO Networks," *IEEE Wireless Communications Letters*, 2023.

[168] W. Shen, Z. Qin, and A. Nallanathan, "Federated Learning Enabled Channel Estimation for RIS-Aided Multi-User Wireless Systems," in *2022 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2022, pp. 812–817.

[169] A. M. Elbir and S. Coleri, "Federated learning for channel estimation in conventional and RIS-assisted massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 4255–4268, 2021.

[170] N. N. Bhat, R. Berkvens, and J. Famaey, "CSI4Free: GAN-Augmented mmWave CSI for Improved Pose Classification," in *2024 IEEE 4th International Symposium on Joint Communications & Sensing (JC&S)*, 2024, pp. 1–6.

[171] Y.-H. Lu, K.-H. Ou, H.-Y. Chen, M.-H. Wu, T.-W. Yang, H.-H. Tsai, and C.-F. Chou, "On LSTM Autoencoder-Based Hybrid Precoding for Reconfigurable Intelligent Surface-Aided Multiuser Millimeter-Wave Massive MIMO 6G Systems," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–6.

[172] K. Kim and C. S. Hong, "Pruned Autoencoder based mmWave Channel Estimation in RIS-Assisted Wireless Networks," in *2022 23rd Asia-Pacific Network Operations and Management Symposium (APNOMS)*, 2022, pp. 1–4.

[173] M. Wu, Z. Gao, Y. Huang, Z. Xiao, D. W. K. Ng, and Z. Zhang, "Deep Learning-Based Rate-Splitting Multiple Access for Reconfigurable Intelligent Surface-Aided Tera-Hertz Massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1431–1451, May 2023.

[174] T.-H. Vu, T.-V. Nguyen, D. B. da Costa, and S. Kim, "Reconfigurable Intelligent Surface-Aided Cognitive NOMA Networks: Performance Analysis and Deep Learning Evaluation," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 10 662–10 677, 2022.

[175] H. Ren, C. Pan, L. Wang, W. Liu, Z. Kou, and K. Wang, "Long-term CSI-based design for RIS-aided multiuser MISO systems exploiting deep reinforcement learning," *IEEE Communications Letters*, vol. 26, no. 3, pp. 567–571, 2022.

[176] P.-H. Chou, B.-R. Zheng, W.-J. Huang, W. Saad, Y. Tsao, and R. Y. Chang, "Deep Reinforcement Learning-Based Precoding for Multi-RIS-Aided Multiuser Downlink Systems With Practical Phase Shift," *IEEE Wireless Communications Letters*, vol. 14, no. 1, pp. 23–27, 2025.

[177] W. Tong, W. Xu, F. Wang, W. Ni, and J. Zhang, "Diffusion Model-Based Channel Estimation for RIS-Aided Communication Systems," *IEEE Wireless Communications Letters*, vol. 13, no. 9, pp. 2586–2590, 2024.

[178] J. Xiao, J. Wang, Z. Chen, and G. Huang, "U-MLP-Based Hybrid-Field Channel Estimation for XL-RIS Assisted Millimeter-Wave MIMO Systems," *IEEE Wireless Communications Letters*, vol. 12, no. 6, pp. 1042–1046, 2023.

[179] J. Guo, G. Liu, Q. Wu, and P. Fan, "Parallel Attention-Based Transformer for Channel Estimation in RIS-Aided 6G Wireless Communications," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 11, pp. 15 927–15 940, 2024.

[180] D. Luan and J. S. Thompson, "Channelformer: Attention Based Neural Solution for Wireless Channel Estimation and Effective Online Training," *IEEE Transactions on Wireless Communications*, vol. 22, no. 10, pp. 6562–6577, 2023.

[181] Z. Qi, D. Liu, and J. Zhang, "Channel Estimation for RIS-Assisted Time-Varying MIMO System: An Attention-Based Learning Ap-

proach," *IEEE Transactions on Green Communications and Networking*, vol. 9, no. 1, pp. 140–151, 2025.

[182] S. Zhang, S. Zhang, F. Gao, J. Ma, and O. A. Dobre, "Deep Learning Optimized Sparse Antenna Activation for Reconfigurable Intelligent Surface Assisted Communication," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6691–6705, Oct 2021.

[183] M. Xu, S. Zhang, J. Ma, and O. A. Dobre, "Deep Learning-Based Time-Varying Channel Estimation for RIS Assisted Communication," *IEEE Communications Letters*, vol. 26, no. 1, pp. 94–98, Jan 2022.

[184] M. A. Mohsin, S. M. Jameel, H. Rizwan, M. Iqbal, T. Ashraf, and J.-Y. Pan, "Transformer-based Distributed Machine Learning for Downlink Channel Estimation in RIS-Aided Networks," in *2025 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*, 2025, pp. 1–5.

[185] S. Singh, A. Trivedi, and D. Saxena, "Channel Estimation for Intelligent Reflecting Surface Aided Communication via Graph Transformer," *IEEE Transactions on Green Communications and Networking*, vol. 8, no. 2, pp. 756–766, 2024.

[186] G. Xia, H. Liu, and K. Long, "Transformer-Empowered Parallel Channel Prediction for Fast-Paced and Dynamic RIS-Aided Wireless Communication Systems," *IEEE Communications Letters*, vol. 28, no. 6, pp. 1347–1351, 2024.

[187] A. Siebert, G. Ferre, B. Le Gal, and A. Fourny, "The Smart Kalman Filter: A Deep Learning-Based Approach for Time-Varying Channel Estimation," in *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2023, pp. 1–6.

[188] S. Gao, P. Dong, Z. Pan, and X. You, "Lightweight Deep Learning Based Channel Estimation for Extremely Large-Scale Massive MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 7, pp. 10 750–10 754, 2024.

[189] Q. Hu, F. Gao, H. Zhang, S. Jin, and G. Y. Li, "Deep Learning for Channel Estimation: Interpretation, Performance, and Comparison," *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2398–2412, 2021.

[190] N. Van Huynh and G. Y. Li, "Transfer Learning for Signal Detection in Wireless Networks," *IEEE Wireless Communications Letters*, vol. 11, no. 11, pp. 2325–2329, 2022.

[191] M. Akrout, A. Feriani, F. Bellili, A. Mezghani, and E. Hossain, "Domain Generalization in Machine Learning Models for Wireless Communications: Concepts, State-of-the-Art, and Open Issues," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 4, pp. 3014–3037, 2023.

[192] C. Zhang, P. Patras, and H. Haddadi, "Deep Learning in Mobile and Wireless Networking: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019.

[193] X. Shi, J. Wang, and J. Song, "Triple-Structured Compressive Sensing-Based Channel Estimation for RIS-Aided MU-MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 11 095–11 109, 2022.

[194] W. Chen, R. Yu, and Z. Ye, "An Attention-Based Denoising Neural Network for mmWave RIS-Assisted SIMO Channel Estimation," *IEEE Wireless Communications Letters*, vol. 13, no. 7, pp. 1933–1937, 2024.

[195] M. Dai, T. Gong, and S. Zhang, "Graph Convolutional Network-Based Channel Extrapolation for Hybrid RIS-Aided Communication," *IEEE Wireless Communications Letters*, vol. 13, no. 12, pp. 3558–3562, 2024.

[196] K. Zhi, C. Pan, H. Ren, K. K. Chai, and M. Elkashlan, "Active RIS versus passive RIS: Which is superior with the same power budget?" *IEEE Communications Letters*, vol. 26, no. 5, pp. 1150–1154, 2022.

[197] E. Shi, J. Zhang, H. Du, B. Ai, C. Yuen, D. Niyato, K. B. Letaief, and X. Shen, "RIS-aided cell-free massive MIMO systems for 6G: Fundamentals, system design, and applications," *Proceedings of the IEEE*, vol. 112, no. 4, pp. 331–364, 2024.

[198] E. G. Larsson, "MIMO detection methods: How they work [lecture notes]," *IEEE signal processing magazine*, vol. 26, no. 3, pp. 91–95, 2009.

[199] M. H. Abid, I. A. Talin, and M. I. Kadir, "Reconfigurable intelligent surface-aided orthogonal time frequency space and its deep learning-based signal detection," *IEEE Access*, vol. 11, pp. 47 321–47 338, 2023.

[200] N. Samuel, T. Diskin, and A. Wiesel, "Learning to detect," *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2554–2564, 2019.

[201] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Model-driven deep learning for MIMO detection," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1702–1715, 2020.

[202] F. Jiang, C. Li, Z. Gong, and R. Su, "Stair Matrix and Its Applications to Massive MIMO Uplink Data Detection," *IEEE Transactions on Communications*, vol. 66, no. 6, pp. 2437–2455, 2018.

[203] X. Tan, W. Xu, K. Sun, Y. Xu, Y. Be'ery, X. You, and C. Zhang, "Improving massive MIMO message passing detectors with deep neural network," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 1267–1280, 2019.

[204] A. Kumar, N. Gaur, M. Gupta, and A. Nanthaamornphong, "Implementation of the deep learning method for signal detection in massive-MIMO-NOMA systems," *Heliyon*, vol. 10, no. 3, 2024.

[205] M. Ahmad and S. Y. Shin, "Wavelet-based massive MIMO-NOMA with advanced channel estimation and detection powered by deep learning," *Physical Communication*, vol. 61, p. 102189, 2023.

[206] Y. Yu, J. Ying, P. Wang, and L. Guo, "A data-driven deep learning network for massive MIMO detection with high-order QAM," *Journal of Communications and Networks*, vol. 25, no. 1, pp. 50–60, 2023.

[207] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "A model-driven deep learning network for MIMO detection," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2018, pp. 584–588.

[208] C. Keawin, A. Innok, and P. Uthansakul, "Optimization of signal detection using deep CNN in ultra-massive MIMO," in *Telecom*, vol. 5, no. 2. MDPI, 2024, pp. 280–295.

[209] H. Yao, T. Li, Y. Song, W. Ji, Y. Liang, and F. Li, "Low-complexity signal detection networks based on Gauss-Seidel iterative method for massive MIMO systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2022, no. 1, p. 51, 2022.

[210] Z. Zhang, Y. Li, X. Yan, and Z. Ouyang, "A low-complexity AMP detection algorithm with deep neural network for massive MIMO systems," *Digital Communications and Networks*, vol. 10, no. 5, pp. 1375–1386, 2024.

[211] A. Kumar, A. Nanthaamornphong, M. H. Alsharif, and M. Masud, "SIC based RL for massive MIMO NOMA signal detection for different modulation schemes under diverse channel conditions," *Scientific Reports*, vol. 15, no. 1, p. 24641, 2025.

[212] S. Brennsteiner, T. Arslan, J. S. Thompson, and A. McCormick, "LAMANet: A Real-Time, Machine Learning-Enhanced Approximate Message Passing Detector for Massive MIMO," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 31, no. 3, pp. 382–395, 2023.

[213] M. Khani, M. Alizadeh, J. Hoydis, and P. Fleming, "Adaptive neural signal detection for massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5635–5648, 2020.

[214] L. Li, J. Xu, L. Zheng, and L. Liu, "Real-Time Machine Learning for Multi-User Massive MIMO: Symbol Detection Using Multi-Mode StructNet," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 9172–9186, 2023.

[215] H. Ye and L. Liang, "On Purely Data-Driven Massive MIMO Detectors," *IEEE Transactions on Signal Processing*, vol. 73, pp. 3079–3093, 2025.

[216] G. Gao, C. Dong, and K. Niu, "Sparsely connected neural network for massive MIMO detection," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*. IEEE, 2018, pp. 397–402.

[217] N. T. Nguyen and K. Lee, "Deep learning-aided tabu search detection for large MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 4262–4275, 2020.

[218] L. Jin, Y. Tu, J. Yang, and B. Shen, "Interference cancellation assisted enhanced sparsely connected neural network for signal detection in massive MIMO systems," *Physical Communication*, vol. 66, p. 102438, 2024.

[219] Y. Liu, C. Zhang, B. K. Ng, and C.-T. Lam, "Complex-Valued Neural Network Detection for RIS-Assisted Generalized Spatial Modulation," in *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*. IEEE, 2024, pp. 1–7.

[220] B. Sagir, E. Aydin, and H. Ilhan, "Deep-learning-assisted IoT-based RIS for cooperative communications," *IEEE Internet of Things Journal*, vol. 10, no. 12, pp. 10 471–10 483, 2023.

[221] S. P. Azizi, H.-Y. Lu, and S.-C. Cheng, "Intelligent Detection for RIS-Assisted MIMO Systems: A First-and-Second Momentum Approach," *IEEE Wireless Communications Letters*, 2025.

[222] L. Wang, N. Shlezinger, G. C. Alexandropoulos, H. Zhang, B. Wang, and Y. C. Eldar, "Jointly learned symbol detection and signal reflection in RIS-aided multi-user MIMO systems," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2021, pp. 715–721.

[223] X. Jin, X. Li, Y. Feng, M. Wen, C. Huang, and Y. Yao, "RIS-based opposite phase index modulation and its DNN-assisted joint sequence detection," *IEEE Wireless Communications Letters*, vol. 12, no. 9, pp. 1518–1522, 2023.

[224] K. Singh, A. F. Makarim, H. Albinsaid, C.-P. Li, and Z. J. Haas, "Passive beamforming design and DNN-based signal detection in RIS-assisted MIMO systems with generalized spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 2, pp. 1879–1892, 2022.

[225] I. I. Ioannou, M. Raspopoulos, P. Nagaradjane, C. Christophorou, W. A. Aziz, V. Vassiliou, and A. Pitsillides, "DeepRISBeam: deep learning-based RIS beam management for radio channel optimization," *IEEE Access*, vol. 12, pp. 81 646–81 681, 2024.

[226] J. Liu and M. Di Renzo, "Data-driven and model-driven deep learning detection for RIS-aided spatial modulation," in *2021 IEEE 4th 5G World Forum (5GWF)*. IEEE, 2021, pp. 88–92.

[227] J. Du, X. Luo, X. Li, M. Zhu, K. M. Rabie, and F. Kara, "Semi-blind joint channel estimation and symbol detection for RIS-empowered multiuser mmWave systems," *IEEE Communications Letters*, vol. 27, no. 1, pp. 362–366, 2022.

[228] Y. Xiong, S. Ding, S. Sun, L. Liu, Z. Zhang, S. Yang, and N. Wei, "Separate Channel Estimation and Data Detection for Semi-Passive RIS-Empowered Massive MIMO," *IEEE Communications Letters*, 2024.

[229] S. Khan, K. S. Khan, N. Haider, and S. Y. Shin, "Deep-learning-aided detection for reconfigurable intelligent surfaces," *arXiv preprint arXiv:1910.09136*, 2019.

[230] L. V. Nguyen, N. T. Nguyen, N. H. Tran, M. Juntti, A. L. Swindlehurst, and D. H. Nguyen, "Leveraging deep neural networks for massive MIMO data detection," *IEEE Wireless Communications*, vol. 30, no. 1, pp. 174–180, 2022.

[231] L. Jiao, Y. Shao, L. Sun, F. Liu, S. Yang, W. Ma, L. Li, X. Liu, B. Hou, X. Zhang, R. Shang, Y. Li, S. Wang, X. Tang, and Y. Guo, "Advanced Deep Learning Models for 6G: Overview, Opportunities, and Challenges," *IEEE Access*, vol. 12, pp. 133 245–133 314, 2024.

[232] M. Wu, K. Guo, X. Li, A. Nauman, K. An, and J. Wang, "Optimization design in RIS-assisted integrated satellite-UAV-served 6G IoT: A deep reinforcement learning approach," *IEEE Internet of Things Magazine*, vol. 7, no. 1, pp. 12–18, 2024.

[233] J. Jiang, Y. Gao, X. Wu, and S. Xu, "Towards channel foundation models (CFMs): Motivations, methodologies and opportunities," *arXiv preprint arXiv:2507.13637*, 2025.

[234] B. Narottama, S. Aïssa, and H. Murata, "Channel Estimation and Tracking in STAR-RIS Aided Systems: A Recurrent Quantum Learning Framework," *IEEE Transactions on Green Communications and Networking*, vol. 9, no. 4, pp. 2129–2143, 2025.

[235] O. M. Abusaid and M. R. Alaqtaa, "Toward Security Enhancement for Deep Learning-Based Channel Estimation Model Under Adversarial Attacks in Next-G Wireless Networks," in *2024 IEEE 4th International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)*, 2024, pp. 500–504.

[236] X. Bai and Q. Peng, "An online deep learning based channel estimation method for mmWave massive MIMO systems," in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 2023, pp. 1–5.

[237] M. A. Cheema, A. Chawla, V. C. Gogineni, and P. Salvo Rossi, "Region-specific reliable channel estimation in RIS-enabled wireless communications via clustered federated learning," *Physical Communication*, vol. 71, p. 102698, 2025.

[238] B. Qiu, X. Chang, X. Li, H. Xiao, and Z. Zhang, "Federated Learning-Based Channel Estimation for RIS-Aided Communication Systems," *IEEE Wireless Communications Letters*, vol. 13, no. 8, pp. 2130–2134, 2024.

[239] Q. Yi, P. Yang, Z. Liu, Y. Huang, and S. Zammit, "Model-Driven Federated Learning for Channel Estimation in Millimeter-Wave Massive MIMO Systems," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, 2024, pp. 1–6.

[240] S. Norouzi, E. Samikwa, M. Rahmani, T. Braun, and A. Burr, "Decentralized Federated Learning for GNN-Based Channel Estimation With DM-RS in O-RAN," in *2025 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, 2025, pp. 1–7.

[241] S. Pala, K. Singh, C.-P. Li, and O. A. Dobre, "Empowering ISAC Systems With Federated Learning: A Focus on Satellite and RIS-Enhanced Terrestrial Integrated Networks," *IEEE Transactions on Wireless Communications*, vol. 24, no. 1, pp. 810–824, 2025.

[242] W. Ma, L. Zhuo, L. Li, Y. Liu, and H. Ren, "Deep reinforcement learning for RIS-aided multiuser MISO system with hardware impairments," *Applied Sciences*, vol. 12, no. 14, p. 7236, 2022.

[243] J. A. Soares, K. S. Mayer, and D. S. Arantes, "Semi-Supervised ML-Based Joint Channel Estimation and Decoding for m-MIMO With Gaussian Inference Learning," *IEEE Wireless Communications Letters*, vol. 12, no. 12, pp. 2123–2127, 2023.

[244] C. Qing, W. Hu, L. Wang, G. Ling, Z. Liu, X. Cai, and J. Wang, "Superimposed Pilot-based Channel Estimation for RIS-assisted IoT Systems Using Lightweight Network," *Wireless Personal Communications*, vol. 134, no. 2, pp. 755–779, 2024.

[245] Y. Li, W. Gao, W. Jiang, and Y. Li, "Joint Direct and Cascaded Channel Estimation Using TCNNs in XL-RIS-Aided Massive MIMO," in *2025 4th International Conference on Electronic Electrical Engineering and Automatic Control (EEEAC)*, 2025, pp. 213–217.

Muhammad Umer Zia (Member, IEEE) received two master's degrees in communication engineering, one from the University of Management and Technology, Pakistan, and the other from Chalmers University of Technology, Sweden. He received the Ph.D. degree in Communication Engineering and Artificial Intelligence from James Cook University, Australia.

He has held research positions at Friedrich-Alexander University Erlangen-Nürnberg, Germany, the University of Modena and Reggio Emilia, Italy, and Lahore University of Management Sciences, Pakistan. His research interests include deep-learning-aided channel estimation and detection, massive MIMO, RIS-enabled massive MIMO, IoT communications, and CSI estimation and feedback design for ultra-wideband systems. He has published in leading journals, including IEEE Transactions on Vehicular Technology, IEEE Transactions on Communications, IEEE Journal on Selected Areas in Communications, IEEE Open Journal of the Communications Society, and IEEE Communications Letters. He has taught several telecommunication courses and conducted labs along with his research at James Cook University.

Wei Xiang (S'00–M'04–SM'10) Professor Wei Xiang is La Trobe Distinguished Professor and Cisco Research Chair of AI and IoT at La Trobe University. He founded two world-first research centres, namely the Cisco-La Trobe Centre for AI and IoT in 2020, and the Australian Centre for AI and Medical Innovation (ACAMI) in 2024. Previously, he was Foundation Chair and Head of Discipline of IoT Engineering at James Cook University, Cairns, Australia. For the past six consecutive years (2020- 2025), has consistently ranked among Stanford University World's Top 2% Scientists for both his single-year and career-long impacts. Due to his instrumental leadership in establishing Australia's first accredited Internet of Things Engineering degree program, he was inducted into Pearcy Foundation's Hall of Fame in October 2018. He is a TEDx speaker and an elected Fellow of the IET in UK and Engineers Australia. He received the TNQ Innovation

Award in 2016, and Pearcey Entrepreneurship Award in 2017, and Engineers Australia Cairns Engineer of the Year in 2017. He was a co-recipient of four Best Paper Awards at WiSATS'2019, WCSP'2015, IEEE WCNC'2011, and ICWMC'2009. He has been awarded several prestigious fellowship titles. He was named a Queensland International Fellow (2010-2011) by the Queensland Government of Australia, an Endeavour Research Fellow (2012-2013) by the Commonwealth Government of Australia, a Smart Futures Fellow (2012-2015) by the Queensland Government of Australia, and a JSPS Invitational Fellow jointly by the Australian Academy of Science and Japanese Society for Promotion of Science (2014-2015). He was the Vice Chair of the IEEE Northern Australia Section from 2016-2020. He is currently an Associate Editor for IEEE Communications Surveys & Tutorials, IEEE Transactions on Vehicular Technology, IEEE Internet of Things Journal, IEEE Access, and Nature journal of Scientific Reports. He has published over 450 peer-reviewed papers including 3 books and over 350 journal articles. He has severed in a large number of international conferences in the capacity of General Co-Chair, TPC Co-Chair, Symposium Chair, etc. His research interest includes the Internet of Things, wireless communications, machine learning for IoT data analytics, and computer vision.

Tao Huang (Senior Member, IEEE) received the B.Eng. degree in Electronics and Information Engineering from Huazhong University of Science and Technology, Wuhan, China, in 2003, the M.Eng. degree in Sensor System Signal Processing from The University of Adelaide, Adelaide, SA, Australia, in 2007, and the Ph.D. degree in Electrical Engineering from the University of New South Wales, Sydney, NSW, Australia, in 2016. He was an Endeavour Australia Cheung Kong Research Fellow. He is currently a Senior Lecturer with James Cook University (JCU), Cairns, QLD, Australia, where he also serves as Co-Deputy Head of the JCU Centre for AI and Data Science Innovation (CADSI). Dr Huang has received several awards and recognitions, including the IEEE WCNC Best Paper Award, the IEEE Outstanding Leadership Award, the IEEE Access Outstanding Associate Editor Recognition, the Springer Nature Editor of Distinction Awards, and the JCU Citation for Outstanding Contribution to Student Learning. He currently serves as Vice Chair of the IEEE Northern Australia Section and Chair of the local IEEE MTT-S/ComSoc Chapter. He has held various leadership roles at international conferences and is an Associate Editor of Scientific Reports, IEEE Transactions on Emerging Topics in Computing, IEEE Open Journal of the Communications Society, IEEE Access, and IET Communications.

Jameel Ahmad is an Associate Professor in the Department of Computer Science at SST, working as Director FYP-CS at the University of Management and Technology (UMT), Lahore, Pakistan. He holds a Ph.D. in Electrical Engineering from UET Lahore, an M.Sc. in Systems Engineering from Quaid-e-Azam University (CNS Campus), and an M.Sc. in Electrical Engineering from the University of Southern California (USC), Los Angeles, USA. He received the MOST scholarship from Pakistan's Ministry of Science and Technology (2002–2006). From 2007 to 2010, he worked at Qualcomm and Broadcom Corporation in San Diego, USA, focusing on mobile communication systems. At UMT, he also served in the Department of Electrical Engineering for a decade, was Director of Final Year Projects in the BSEE program for four years, and was honored with UMT's Best Researcher Award in 2019 for a publication in IEEE Transactions on Industrial Informatics. His research interests include deep learning techniques applied to wireless and energy-efficient systems, digital image processing, computer vision, and autonomous vehicles. Professor Ahmad serves as a reviewer for numerous journals and conferences.

Jawwad Nasar Chattha (Senior Member, IEEE) is a Course Leader for Computing in the Department of Computer Arts and Technology at Norwich University of the Arts (NUA). He holds a Ph.D. in Communication Theory from the Lahore University of Management Sciences (LUMS). During his Master's at Virginia Tech (2009), he contributed to a National Science Foundation (NSF)-funded project. He has also worked in industry, including at Qualcomm in San Diego, California, focusing on various communication standards before transitioning to academia. Previously, he served as Chair and Associate Professor in the Department of Electrical Engineering at the University of Management and Technology (UMT), Lahore. He has developed numerous undergraduate courses, including Internet of Things (IoT) and Applied Machine Learning, and implemented an Outcome-Based Education (OBE) model in the Electrical Engineering department. Dr. Chattha was a member of the National Working Group for the International Telecommunication Union's World Radiocommunication Conference 2023 (ITU WRC-23). His research interests include communication theory and computer networking. He has reviewed for several renowned journals in these areas.

Rana Abbas is a transport infrastructure strategist and researcher with over a decade of experience translating advanced technologies into real-world infrastructure solutions. Her expertise spans both fundamental and applied research, including channel coding, random multiple access, V2X communications, point cloud segmentation, and robotics for construction and infrastructure applications. She completed her Ph.D. in Telecommunications at the University of Sydney in 2018, where she received the IEEE PIMRC Best Paper Award (2017) and was recognized as an Exemplary Reviewer by the IEEE Communications Society. Her career spans academia, industry, and government, where she has led multidisciplinary research programs, established strategic industry partnerships, and provided advisory input on innovation and technology adoption. She is the founder and manager of Transport for New South Wales' Innovation Hub, where she leads initiatives focused on integrating emerging technologies into next-generation transport infrastructure, aligning with national strategy and policy priorities.

Asad Mahmood is a PhD student at James Cook University (JCU), Australia, working in the broad areas of Materials Science and Electronics, with a focus on charge transport phenomena in organic semiconductor devices. He has over 10 years of experience as a Lecturer in the Higher Education Department, Punjab, Pakistan. He received the MS degree in Computational Fluid Dynamics and has strong expertise in Computational Methods, Numerical Modelling, Optimization Techniques, and Transport Physics. His research interests include Image Processing-based Morphology Reconstruction, 3D Charge Transport Modelling, and Structure-Property Relationships in organic electronic materials, employing Master Equation, Kinetic Monte Carlo, and Drift-Diffusion frameworks. He has authored several peer-reviewed publications spanning fluid dynamics, including computational and transport-driven flow modelling, as well as experimental and theoretical charge-transport studies in organic semiconductors and related material systems.