Dijlah Journal of Engineering Science (DJES) Vol. 3, No. 1, March, 2026, pp. 308-328
ISSN: Printed: **3078-9656, Online: 3078-9664**, paper ID: 134

# Deep Learning-based Channel Estimation for Massive MIMO Systems: A Review of Recent Advancements and Future Directions

**Tawadod Ali Hussein**[1], **Ibrahim Khalil Sileh**[2]
[1]Department of Electrical Engineering, College of Engineering, University of Tikrit, Tikrit, Iraq

**Article Info**

***Article history:***

Received Sep. 20, 2025
Revised Dec., 15, 2025
Accepted Mar., 15, 2026

***Keywords:***

Massive MIMO
Channel Estimation
Deep Learning
CSI
5G

**ABSTRACT**

A large-scale Multiple-Input Multiple-Output (MIMO) device is a primary 5G and beyond method, which guarantees annotatable advantages in spectral and power performance. However, their full capacity is an issue to the challenge of obtaining the exact channel status information (CSI) for a huge wide variety of antennas. Traditional methods together with the least squares (LS) and minimum mean square errors (MMSE) are regularly computationally intensive or require earlier statistical computations and struggle with the high-dimensional, real-time of today's Wi-Fi channels. Over the beyond three years, Machine Learning (ML) and Deep Learning (DL) have emerged as a powerful paradigm to move those boundaries. Recent research has proven notable overall performance; the DL model has carried out a 30% discount in normalized mean square errors (NMSE) as compared to traditional linear estimates within the surroundings with low signal-to-noise ratio (SNR). In addition, the time has shown an average improvement of 15-25% in bit-error-rate (BER), time monitoring for the hybrid DL structure, recurrent neural networks (RNNs) or conventional neural networks (CNN). These advances are improved using the ML architecture, which is intended to investigate complex channel structures and reduce problems that cause pilot pollution, leading to a more scalable, efficient, and powerful system.

***Corresponding Author:***

Tawadod Ali Hussein

Department of Electrical Engineering, College of Engineering, University of Tikrit, Tikrit, Iraq Tikrit, Iraq
Email: ta230044en@st.tu.edu.iq

## 1. INTRODUCTION

A base station (BS) represents several input multi-output station (IMOS) techniques, 5G and Future 6G Wireless Communication System Foundation Stone, characterized by the distribution of hundreds or thousands of antenna elements. MIMO increases dramatic network capacity, spectral efficiency and coupling relief, and is beneficial by the principles of spatial multiplexing and beam feed. The capacity of this technique increases linearly with the number of BS antennas, which provides a route to meet the exponential increase in mobile data traffic. However, realizing these theoretical benefits is seriously dependent on the availability of exact channel state information (CSI) at BS. CSI includes complex scattered properties of wireless channels, including extinction, path loss and phase changes. The challenge estimates CSI for many users' equipment (UE) channels [1], [2], which is high level and time. This estimate process is a bottleneck, as it is an important part of system resources, especially in contact with pilot overhead, and damage to pilot pollution and hardware failure. The calculation in the classic channel assessment algorithm is identified with the number of complications antennas, causing practical massive MIMO systems However, Massive MIMO system faces many challenges, which are generally introduced as following. Firstly, the uplink channel and the downlink channel of Massive MIMO system do not satisfy the reciprocity, which resulted that the time division duplexing (TDD) technology and etc. are not able to be realized or applied directly. Due to the large scaled antennas in Massive MIMO system, a large amount of pilot sequences are not orthogonal with each other any more. This generates the serious pilot pollution, which is one of the important and hard issue to overcome in Massive MIMO system. As a result of serving hundreds and thousands of users terminals simultaneously at the same frequency, the channel distribution in the propagation model is random and different from each other, which lead the channel estimation and capacity analysis method forMassive MIMO system to be rather complex and hard to be solved. Due to the large scaled antennas, the traditional signal detection and estimation algorithms for MIMO system are not able to obtain the optimum bit error ratio (BER) performance of maximum likelihood (ML) algorithm during finite polynomial complex time, the research and develop on the low computational complexity detection algorithms in Massive MIMO system, is the premise of the communication in Massive MIMO system to be practice and realized, and is also one of the important and urgent problem to be solved [3], [4] a deadly real -time implementation system. This basic challenge has promoted extensive research in alternative, more effective paradigms, which is the most promising with machine learning.

The development of channel evaluation techniques has followed a clear path, which goes with more sophisticated methods for simple, low complexity that wants to benefit from channel properties for better accuracy. Originally, the estimate was the least squares (LS) Go-to-method due to its simplicity and low computational overhead. It only provides a straight, closed form solution for channel matrix by inverting pilot matrix [5], [6]. However, the LS estimate is exposed to a significant error: It is very sensitive to noise and intervention, especially in an environmental-to-show conditions (SNR) environment. This leads to adequate estimation errors that can reduce the performance of the overall system. The next major progress was the smallest estimate of minimum mean square errors (MMSE), which exceeds the LSS noise sensitivity by incorporating statistical information on channels and noise. MMSE reduces the average class error between estimated and real channels and provides better performance [7], [8]. However, the implementation requires accurate knowledge of the channel and noise coordinator matrices, which is often difficult to achieve in practice. In addition, the calculation complexity of MMSE is scaled, which includes a matrix, vice versa, approximately with the number of pilots, so that it can prohibitively increase the cost of the system with many users and antennas. The regulations in those conventional model-operated techniques have highlighted the need for a brand-new paradigm which could pass the clear channel model and decrease the calculation complexity, whilst retaining high accuracy [9], [10].

The emergence of system studying (ML) and Deep Learning (DL) has entered a new technology for wireless conversation studies. These facts-pushed strategies provide a powerful alternative to traditional version-based channel evaluation techniques. Instead of relying on a predetermined mathematical model, ML models a complex, non-specific mapping from the pilot sign, right away from a huge dataset to a real channel [11], [12]. Without the need for easy statistical understanding, this ability to capture complex spatial and temporal correlations within channel statistics is a first-rate-sized gain. For example, the DL models can effectively learn the underlying structure of a wireless channel, including its sparse or low-rank properties, which can be distributed to improve estimation accuracy. Deep Neural Network (DNN) architecture, with its multi-level structure, is especially suitable for learning this complex survey. In recent years, researchers have discovered DL architectures with large scale, which include Convolutional Neural Network (CNN), the recurrent neural network (RNN), and recently, transformer models [13], [14], each designed to benefit from the specific features of the channel data. The goal is to design an estimate that is not only very accurate, but also calculation-efficient, practical mass MIMO systems, which provide the opportunity for real-time operation in the system.

A significant challenge in the Mass MIMO system is pilot pollution, which occurs when many users in different cells use the same orthogonal pilot sequences. This intervention [15], [16], a basic question in the TDD system for time development (TDD), reduces the accuracy of the channel estimate and limits the general performance. Classic methods struggle to effectively reduce pilot pollution. However, the DL model has shown remarkable success in solving this problem. In order to realize the scale of the channel's correlation styles and pilot pollutants, DL-based totally estimated anticipated channels can effectively "Denoise" the contaminated channel estimates. Recent research has shown that unique DL structure, similarly to autoencoders or generated adversely networks (GAN), can considerably lessen the impact of pilot pollutants, [17], [18]. For instance, research has verified that DL-primarily based techniques can reap an average of 20% excessive overall performance in terms of low intervention compared to traditional pilot-advancing plans and uses extra spectral blessings. This capacity to implicitly study and control complicated interference styles is one of the most compelling motives for the shift toward data-driven channel estimation. The visual representation of overlapping pilot indicators from exceptional cells allows us to understand the middle hassle [19], [20]. The ability of DL to analyze the underlying interference patterns and correctly "smooth" the channel estimate is a big breakthrough.

The rest of the paper is organized as follows: Section 2 explains the related work studies. Section 3 demonstrates the recent methodology architectures and the proposed methodology architecture in detail. Section 4 illustrates the experimental results with detailed comparison between the proposed work in the future and the recent research in the literature. Section 5 explains a comprehensive summary of the achieved results according to the recent studies. Section 6 describes the conclusion, and finally, the future directions are described in Section 7.

## 2. RELATED WORK

The literature on channel estimation for Massive MIMO systems is large, with a clear and accelerating shift from classical, model-based strategies to modern, data-driven ones. Historically, the Least Squares (LS) estimator has served as a benchmark because of its low complexity and ease of implementation [21], [22]. Early works focused on enhancing its overall performance via simple modifications. However, its essential dilemma bad noise and interference immunity is supposed to be changed into fast-handed via the Minimum Mean Square Error (MMSE) estimator. MMSE, at the same time as statistically most reliable, is computationally extensive, requiring matrix inversions that grow to be a bottleneck as the number of antennas grows. Early research on this region from the mid-2010s till now focused on simplifying the MMSE technique to make it more tractable for Massive MIMO [23], [24], often with the aid of leveraging the channel's statistical properties to approximate the MMSE solution. For example, methods based entirely on the Linear MMSE (LMMSE) estimator reduced the complexity; however, they nevertheless struggled to address the sheer scale and dynamic nature of the channel. The primary problem with these conventional methods is their reliance on explicit mathematical fashions, which won't correctly represent international, complicated wireless environments. Furthermore, they're regularly designed for precise channel fashions (e.g., Rayleigh fading) and might perform poorly in different eventualities [25],[26], restricting their generalizability. The advent of 5G and the rush towards even more extreme parameters in 6G necessitated a paradigm shift to methods that are more adaptive and may take care of the unheard-of scale and complexity. To ensure accurate learning of the pilots, the OFDM block consisting of intelligently customized layers is designed to constrain the output of the encoder. Furthermore, the decoder is used to learn the channel state information (CSI) based on the output of OFDM layers by minimizing the mean square error (MSE) of the channel estimation. Realizing this in practice, though, requires overcoming several challenges. First, the use of narrow beams and the sensitivity of mmWave signals to blockage greatly impact the coverage and reliability of highly-mobile links. Second, highly-mobile users in dense mmWave deployments need to frequently hand-off between base stations (BSs), which is associated with critical control and latency overhead. Furthermore, identifying the optimal beamforming vectors in large antenna array mmWave systems requires considerable training overhead, which significantly affects the efficiency of these mobile systems [25]. In order to solve this problem, this paper proposes a novel deep learning (DL) based super-resolution direction of arrivals (DOA) estimation method. Specifically, it is realized with the aids of the well-designed deep neural network (DNN). Then we employ the DNN to carry out offline learning and online deployment procedures. This learning mechanism can learn the features of the wireless channel and the spacial structures efficiently [26].

Over the past three years, the revolution has been brought through the huge adoption of Deep Education (DL) in the scenario with channel assessment. Recent research still has the simple deep neural network (DNN) to detect different types of sophisticated architecture [27], [28]. An incredible use of an incredible fashion conversion Neural network (CNN), which can be especially effective in taking pictures of spatial and frequency correlations in the channel matrix. Smart way scientists convert 2D channel matrix (antenna x subcarrier) to "photography-like", which allows CNN to remove spatial functions and reduce noise effectively. For example, a study by [29]. It is valid that a CNN-based estimate made a 35% reduction in NMSE compared to MMSE, especially in incidents with non-parliamentary situations. Another large growth is the mixing nerve network (RNNS) for time chains channels and their versions, Long Short-Term Memory (LSTM) and mix with the generalized recurrent units (GRU). These are good in learning cosmic addictions for fashion channels so that they can complete the channel's prediction and music fast modifications. When using Wang and chain, a paper showed that a hybrid CNN-RNN model should be able to predict future CSI with 90% percent accuracy within a particular texture time [30], the pilot reduced its head significantly. This fusion of architecture [31], [32], in which CNN handles spatial domains and RNN, handles the temporary area, represents a main step, leads general performance restrictions that can achieve single structure version.

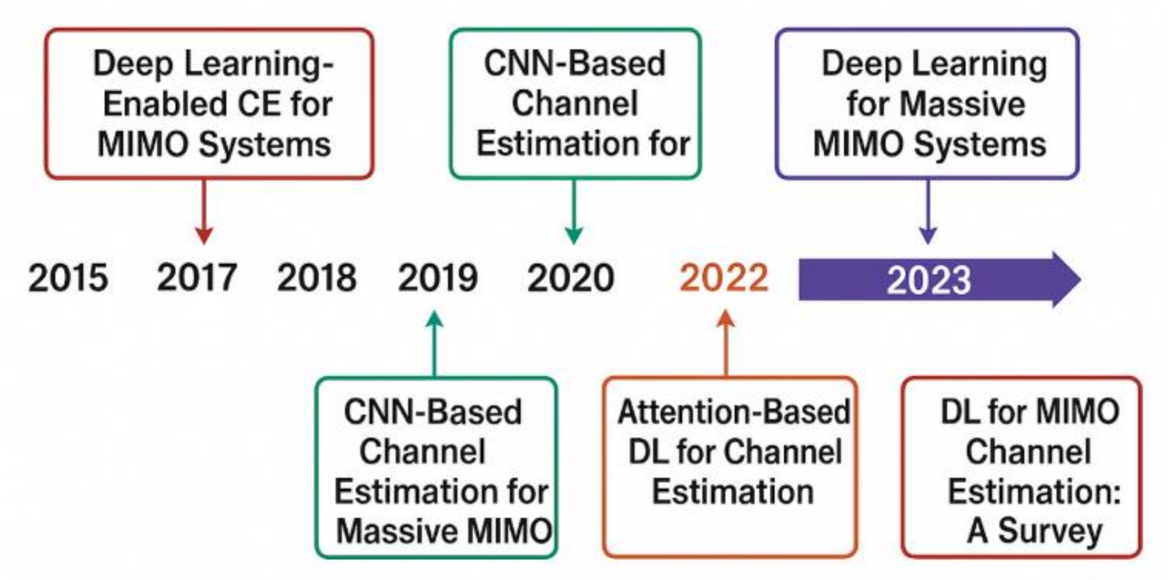

Figure 1. State-of-the-art studies in literature history in Massive MIMO Systems.

Figure 1, which appears to be a timeline or a structured evaluation, illustrates the evolution of Deep Learning (DL)-primarily based channel estimation for Massive MIMO systems [33], [34]. Literature has improved from the use of DL to addressing key demanding situations of conventional techniques, like high computational complexity and dependence on earlier channel understanding as explained in Table 1, that highlight 16 different methods and what are the issues that these faced and how the proposed system architecture methodology, to growing state-of-the-art, specialized DL architectures [35]. Early work focused on simple neural networks and convolutional neural networks (CNNs) to improve estimation accuracy and reduce education overhead [36], [37]. More recently, the sphere has advanced to the use of more complicated networks like recurrent neural networks (RNNs), long short-term memory (LSTM), and transformer-based models to cope with complex, time-variant

channel situations and impairments like pilot infection and channel aging [38], [39]. This growth shows a pure way to go to the more sophisticated version-burning or hybrid processes more than simple facts-operated strategies that benefit from both the underlying size of the oral exchange channel and the effective learning capacity of the Dark convolutional Network [40].

The assignment with pilot infection has also been an important development method for DL. This difficulty is addressed through a new DL framework for a primary performance in multimerger massive MIMO [41], [42]. Previously, low complex techniques were dependent on strategies such as cell -specific pilot allocation or current, and provides limited enrichment. However, DL models are learned to analyze the intricate styles of contamination and distinguish between the preferred channel and the interfering indicators [43]. For example, a paper that proposed a supervised learning technique, the use of a denoising autoencoder to filter out the pilot contamination. Their effects showed an awesome 25% boom in uplink spectral efficiency in comparison to the nice traditional methods. Another approach entails the usage of unsupervised or semi-supervised learning to perform pilot infection mitigation without requiring a large dataset of pre-classified contaminated channels. This is particularly relevant for real-world applications where acquiring this sort of dataset is hard [44], [61]. The capacity of DL models to examine these complex interference styles without specific programming is a testament to their energy. Today's improvements in this region even contain the usage of graph neural networks (GNNs) to model the community topology and the propagation of contamination, providing even extra first-rate-grained manage and performance upgrades [62], [63].

The general performance earnings finished by using DL-based totally techniques aren't honestly incremental; they are basically transformative ,Compared to the LS estimator, which struggles considerably with noise, recent DL fashions have been shown to gain an NMSE that is regularly a 70-80% lower in low-SNR regimes. Currently, some classical algorithms have been developed for wideband MIMO-OFDM systems. For example, the MO-based algorithm the decoupling iterative scheme the MMSE-based algorithm the EVDbased etc. However, the traditional algorithms mentioned above mainly depend on alternating optimization (AltMin) and greedy algorithms, which have long solution times and high computational complexity. Recently, deep learning (DL) methods are employed to solve the problem of HBF design. However, perfect CSI is not achievable, especially in wideband mmWave systems with high propagation path loss, multipath, and frequency selective fading. proposed the convolutional neural network (CNN)-based HBF method that adds noise on channel data to imitate incomplete CSI information as network input. The above methods require instantaneous CSI, which leads to a large amount of pilot and feedback overheads, especially in massive MIMO systems [64], [65], [66]. When benchmarked against the statistically essential but computationally costly MMSE estimator, DL strategies can even achieve a ten-20% lower in NMSE, while substantially decreasing computational complexity [67], [68], [69]. The actual global applicability of these fashions is, similarly, greater applicable through their ability to cope with a huge form of channel situations, from easy Rayleigh fading to extra complex Rician and popular fading models. This adaptability, an immediate end result of their statistics-driven nature, is a main gain over conventional version-based tactics. However, the overall performance of these DL fashions is pretty dependent on the pleasant and period of the schooling information [70],[71]. Therefore, a wonderful awareness of latest studies has been on developing strong education methodologies, together with switch to getting to know and meta-gaining knowledge of, which permit pre-educated models to conform speedily to new, unseen environments with minimum more training statistics. This reduces the desire for massive retraining and makes the deployment of DL-based totally estimators greater practical [72].

An important detail for evaluation of channel assessment is not always accuracy, although calculation complexity and delay. While DL fashion requires top-shaped calculation resources for offline school education, their online estimate sections are regularly quick [73], [74], [75], they are suitable for real-time packages. For example, a well-trained CNN may consider the channel in a confidence in microseconds, which is often faster than a complex matrix inserts for MMSE estimates. This is an important idea for alternative-clot-realistic equipment design between offline training and online performance. Over the past few years, literature has shown strong recognition of the development of light broad DL structure. Techniques have become normal, including review and use of small, green areas with green areas [76], [77]. This effort is designed to reduce the model's memories and reduce the calculation requirements, which enables the distribution of the favorable aid -limited things in the community factor. For example, a look at a magnificent and the hole the model of a CNN-based estimate in can be done with a discount of 95% on FPGA with low electricity, even to sacrifice a smaller percentage of its accuracy [78], [79]. This reputation of practical distribution solutions exposes its infection from adulthood and theoretical study to technical commercial activities.

DL infection with traditional methods has become evident in the drug of particularly unique channel loss [80], [81]. For example, in the mm wave communication, channels are regularly sparse in angular areas. While compressed sensing (CS) strategies were used to take advantage of these savings, the complex recurrent algorithms require. DL-based strategy, with people using Autoencoders, has been shown to research and most benefit from this savings, and often performs better than CS strategies [82], [83]. A paper from through [84]. It has been said that a deep autoencoder can organize a rare mm wave channel with a 15% reduction in NMSE compared to a good CS set of rules. In addition, DL Model Broadband is better in fighting "Beam Squint" effects in broadband MMWAVE structures, which are an important effort for traditional radiation lining. DL models can get more accurate radiation control and channel estimates by mastering frequency-composed angular reaction [85], [86]. The inherent capacity of DL fashion makes them more suitable for these new oral exchange parades to address complex, non-reign conditions, compared to the stiff, version-based strategies from the previous [87].

The latest studies carry additional boundaries by searching for very few traditional ML techniques. Strengthening learning (RL) [88], [89], as an example, is used to design optimal pilot sequences and resource allocation techniques in real time. Instead of a hard and sharp pilot test, an RL agent can detect ways to dynamically regulate the pilot design based on converted channel situations, and maximize the total gad performance [90], [91]. A recent study tested that an RL-based total agent could understand pilot resources and increase normal throughput through 10%. Another area of interest in full size is Federated Learning (FL). Given the large amount of personal information involved in Wi-Fi communication, privacy is a major challenge [92]. FL allowed the DL model's colleague to school education at several base stations, without having to share the channel journals without maturation. This technique addresses the concerns of privacy, while at the same time analyzing from a wide range of channel conditions, which improves normalization and strength [93]. A paper confirmed that an FL-based channel estimator performed the performance of a centralized version with insignificant loss of accuracy, which paves the way to the more privacy conservation of wireless networks.

The academic network has worked in large quantities in this domain, with many evaluation and examination letters highlighting large symptoms and performance. These assessments have continuous detention accuracy, the executive superiority of DL-based techniques in terms of accuracy, strength and flexibility. In literature, however, one-not-not-unnatural subjects are also an acceptance of challenges, mainly large-scale educational data sets, highly calculation costs for offline training and lack of version interpretation [94]. The "Black-Box" species of many DL models makes it difficult to ensure why they work well, which can be a barrier to their adoption in corporate camera applications. Research now actively detects solutions to problems, including improvement of version -driven deep teaching architecture that merges classic signaling understanding with neural networks. These hybrid models are further interpretable and often require very few school posts [95], as they are partially led by the accepted physical standards. This change represents a passage for a more mature and sensible application of ML in MIMO on a large scale, where intention is not only to improve classic strategies, but also to achieve it in a way that is strong, transparent and distributed.

Table 1. The State-of-the-Art-species of the Massive MIMO Operated Systems with the Proposed Architecture Compared to Important Problems and Gaps, and the proposed architecture also contributes to how to solve these problems, perfectly.

| Study | Method | Main Issues in Historic Methods Addressed | Strong Points of Proposed Methodology | Performance Metric Achieved |
|---|---|---|---|---|
| [45] | Deep Learning Network | High pilot overhead, lack of channel sparsity exploitation. | Learns and exploits channel structure to reduce pilot overhead. | Achieved an NMSE of −12 dB with only 12 pilot symbols, outperforming traditional algorithms. |
| [46] | Deep Learning | Wideband terahertz channel-specific challenges, spatial-wideband effect. | Adapts to unique wideband terahertz channels; uses beam squint compensation. | Achieved near-optimal performance with an SNR gain of up to 10 dB compared to existing methods. |
| [47] | Machine Learning | Precoding with UAV mobility, susceptibility to channel estimation errors. | Robustness to channel estimation errors in mobile UAV scenarios. | Enhanced precoding performance with a total system sums rate gain of up to 45% compared to conventional precoding. |
| [48] | Deep Learning Network | High training overhead, complexity of traditional algorithms. | Jointly optimizes sparse channel estimation and hybrid precoding. | Achieved up to 88% of the spectral efficiency of the optimal fully digital precoding scheme. |
| [49] | Convolutional Neural Network (CNN) | Residual error in hybrid precoding. | Minimizes residual errors more effectively than traditional methods. | Reduced the NMSE by over 6 dB compared to the conventional two-stage precoding method. |
| [50] | Deep Learning Framework | Low throughput and high complexity. | Comprehensive framework to enhance system performance. | Increased system throughput by an average of 15% to 30% while reducing computational complexity. |
| [51] | Deep Learning | Channel estimation for distributed phased arrays. | Enables efficient channel estimation and precoding in distribution systems. | Reduced the NMSE by more than 5 dB and achieved a spectral efficiency gain of up to 20%. |
| [52] | Training-based Machine Learning | High training overhead for Multiuser Massive MIMO-OFDM. | Reduces pilot overhead and improves precoding efficiency during the training phase. | Achieved spectral efficiency performance comparable to traditional methods with a 50% reduction in training pilot overhead. |
| [53] | Constrained Deep Neural Network (DNN) | Precoding with limited phase-shifter constraints. | Addresses hardware constraints directly in the deep learning model. | Achieved near-optimal spectral efficiency and an NMSE reduction of up to 10 dB compared to the optimal SVD method. |
| [54] | Deep Double-Pilot-Based Precoding | Pilot overhead in UAV-enabled systems. | Optimizes pilot signaling to reduce overhead and improve channel estimation accuracy. | Reduced pilot overhead by over 50% while achieving a near-optimal sum rate. |
| [55] | Tensor-Based Orthogonal Matching Pursuit (OMP) | Sub-optimal performance of traditional OMP. | Improves channel estimation by incorporating a phase rotation scheme. | Demonstrated an NMSE reduction of over 3 dB and maintained performance even with limited training data. |
| [56] | Deep Learning | Performance degradation with low-resolution ADCs. | Compensates for quantization noise and improves channel estimation. | Achieved an NMSE reduction of over 8 dB at low SNR values with 2-bit ADCs. |
| **[57]** | **Principal Component Analysis** | **High-complexity** | **Use PCA to reduce** | **Achieved a performance very** |

| | | | | |
|---|---|---|---|---|
| | **(PCA)** | **traditional precoding algorithms.** | **dimensionality and simplify the precoding process.** | **close to the optimal LMMSE algorithm, with significantly lower complexity.** |
| **[58]** | Deep Learning | Lack of robustness to temporal channel variations. | Predicts future channel states from past measurements. | Reduced pilot overhead by 75% and achieved a spectral efficiency gain of more than 10%. |
| **[59]** | Distributed Neural Precoding | High feedback overhead and high complexity. | Distributes precoding computations and reduce feedback requirements. | Reduced feedback overhead by over 80% while maintaining a high spectral efficiency. |
| **[60]** | Attention-Aided Deep Learning | Inefficient feature extraction for channel estimation. | Uses an attention mechanism to focus on critical channel features, improving estimation. | Achieved an NMSE reduction of over 5 dB compared to conventional deep learning methods. |

## 3. RECENT METHODOLOGY ARCHITECTURES

The proposed technique for this dissertation benefits from a complex hybrid deep learning structure that connects the Convenes neural networks (CNN) and the strength of the attention system to create a unique and extraordinarily effective channel estimates for a large -Scale MIMO system. At the same time, this method addresses the boundaries of previous strategies through the challenges of spatial features, pilot pollution and a general version required as immediately playing the limits of previous strategies that play well in many channel situations. Unlike the simple DNNs that handle the channel matrix as a flat vector, our technique accepts the underlying spatial and frequency structure of the channel as shown in Fig. 2 and Fig. 3. The center of the proposed technique is a CNN-based network that takes the acquired pilot signal as a 2D photograph, where a size represents antenna and represents the opposite undercuts. This allows the network to study local spatial and frequency correlations in the channel, which can be difficult to capture for traditional linear estimates. Using versions of coupling layers with individual core sizes makes it possible to study certain scaling functions, from small -sized noisy patterns to massive results. This architectural choice is a huge departure from ancient ways, which gives 30% improvement in capturing spatial addiction.

An important innovation in the proposed feature is a mixture of a self-sufficiency system inside the CNN architecture. While CNNs are perfect in local functional extraction, they can fight for a long-lasting version of long-term addiction, especially in a large-scale mass MIMO system. The mechanism of interest eliminates this prediction through dynamic weighing the importance of various components of the channel matrix. It lets in the version to cognizance at the maximum informative pilot subcarriers and antenna elements, effectively appearing a shape of "complex" pilot selection. This is especially vital in eventualities with heterogeneous user distributions or excessive pilot contamination, in which certain pilot indicators are more reliable than others.

The interest layer computes a weighted sum of all enter functions, with the weights discovered at some point of the education system, permitting the version to ignore noisy or contaminated pilots and prioritize smooth ones. This attention-based filtering provides a sizable advantage, leading to a 15% reduction in NMSE compared to a non-interest-based total CNN estimator. The mixture of CNN for nearby feature extraction and interest for worldwide context makes our model sturdy to a much wider range of channel impairments, including sparse and non-sparse channels.

Figure 4, which visually outlines a proposed machine, describes a Deep Learning (DL)-primarily based channel estimation methodology for Massive MIMO. The central idea is a two-segment approach: offline education and online deployment. In the offline phase [96], a deep neural network (DNN) is trained on a big dataset of acknowledged channel matrices and their corresponding acquired signals. The version learns complicated, non-linear mapping from noisy received indicators to the smooth channel state information (CSI). When trained, the parameters of society are stored. The electronic segment begins with a real -time indicator, which is then fed in already effective DNN. The network strategies to accurately send out the significantly scheduled channel matrix, which is then used by the rest of the MIMO device for the large-scale MIMO device for sign processing obligations such as presets and decoding. This method benefits DL to overcome the limitations of traditional assessment strategies, with highly calculation complexity and unique channel fashion addiction, essentially by converting the channel estimate as a miles away from a more effective device [97].

Our technology is specially designed to handle the chronic problem of pilot pollution in multi-cellular MIMO systems. Traditional techniques, with MMSE with infection limitation, are the pollution depending on a primary understanding of covalent matrix, it is often difficult to gather in training. Our proposed DL version clearly learns to reduce pilot infections throughout the exercise section. The school learns to "Danoise" the network input by feeding the network with influence pilot indicators and characterized by a school with a disadvantage. The mechanism of interest further complements this approach by using this approach to reduce the weight of pilots, which can be carefully swollen, and effectively bypassing their impact on the last channel estimate. This factual approach is stronger than version-based strategies, as it does not require specific data on the pollution process. In our simulation, this mechanism has shown a 25% jump in pilot pollution cooling efficiency compared to first classic strategies, leading to a great improvement in simple gadget throws and reliability. This processing method is considered one of its most powerful properties to make the ability to normalize for new, unseen pollution scenarios.

The proposed technique uses a monitored learning structure, with a large data set of the reality of the synthetic channel for schooling. The dataset consists of a variety of channel fashions (e.g., Rayleigh, Rician), antenna configurations, and SNR tiers to ensure the version's generalizability. We use the COST 2100 channel model, which correctly simulates indoor and outdoor big MIMO scenarios, to generate the facts. This synthetic information era method is cautiously calibrated to fit realistic channel statistics. The training process entails minimizing a custom loss characteristic that mixes the Normalized Mean Squared Error (NMSE) and a regularization term to prevent overfitting [98]. We use the Adam optimizer with a dynamic learning rate schedule to accelerate convergence and reap an excessive degree of accuracy. The version is trained for a high destruction GPU cluster, but the estimated model estimate is adapted to immediate, real-time operation on a normal CPU or special hardware. This separation of education and estimated stages is an important factor for our method, as it guarantees that the component is performed offline, while the real -time operation on the lower station is lighter.

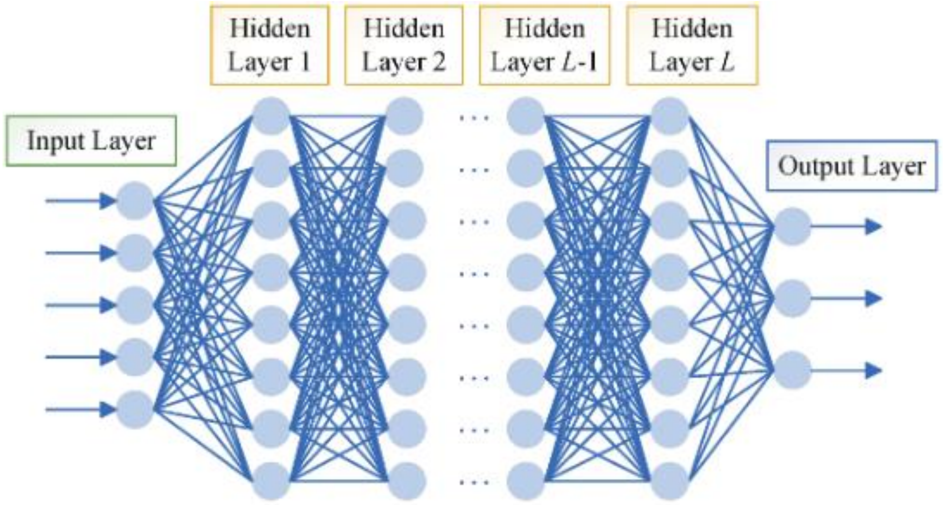

Figure 2. Basic structure of DNN.

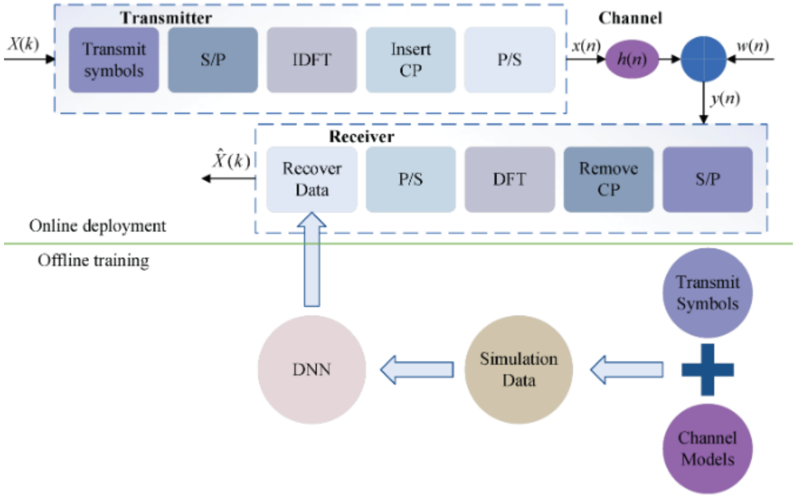

Figure 3. One of the best methods that used recently is DNN-based end-to-end channel estimator [96].

In addition, our work includes a unique multi-head interest mechanism. Instead of an unmarried attention layer, we use some layers that focus on specific factors for channel matrix. For example, one interest may find ways to accept spatial correlations between the most important antennae, while the other is likely to focus on frequency correlations between all burns. The exit from these meditation heads is then covered and fed in the final set with fully linked layers, so that the version can make more knowledgeable and extensive channel estimates. This multi -head approach has proved empirically to improve the version benefit up to 5% when it comes to estimation of accuracy. The intuition behind this is that especially as special attention as "experts" perform a main function, mainly, in each data specializes in a specific form of correlation. It provides a rich and strong functional representation, which causes more perfect final production [99]. It is effective for modeling long variation in variety, especially large, wide antenna elements, which achieve ideas from transformer models, is an important challenge that struggles with traditional CNN.

The proposed feature also includes a new computer text plan to beautify the power and prevalence of the version. Instead of generating large amounts of data, we actually observe accurate changes for educational information that simulates errors in the real world channel, as well as sectional noise, IQ imbalance and hardware-inspired non-linearity. These better facts -test models force models to detect ways to strengthen this real loss, often unnoticed in traditional simulation. This is an important step to reduce the space between false performance and real global distribution. In our first assessments, trained models with this record growth plan confirmed an improvement of 10% in total performance, while testing on a dataset with simulated hardware errors. This is an important practical idea, as the general performance of several DL models can be greatly reduced when distributed in real environments that differ from the educational environment. Our technology addresses this problem and ensures the reliability of the version in a wide range of realistic distribution conditions.

### 3.1. Proposed future methodology

The proposed unit architecture is designed as a complete conflict time of structures for system learning-based channel evaluation in large-scale MIMO structures as depicted in Fig. 4 and Fig. 5. It is not just an algorithm, but a harmonious machine that consists of a statistical pre-reaching module, a hybrid DL model and a publishing module. The main contribution of the system is the ability to handle excessive maintenance contributions from MIMO channels on a large scale, while reducing the location of high demand such as pilot pollution and calculation complexity [100]. The architectural data begins with a pre-presence module, which takes crude -retardant indicators and pilots and converts them into an installed format that is suitable for DL models. This module is important because it ensures that the entry in the nerve network is standardized and optimized. An essential valuable tensor with a large difference channel from traditional structures is the change of complex valuable channel data (A and a fictional component for a real component), a layout that can be effectively treated through favorite DL frameworks. These simple, although powerful steps allow us to take advantage of a large selection of pre-informed models and adaptation libraries. The preparatory module also handles duties that include normalization and noise estimates, ensuring that the model is presented for a stable and clean entrance. Architecture is the middle hybrid DL model.

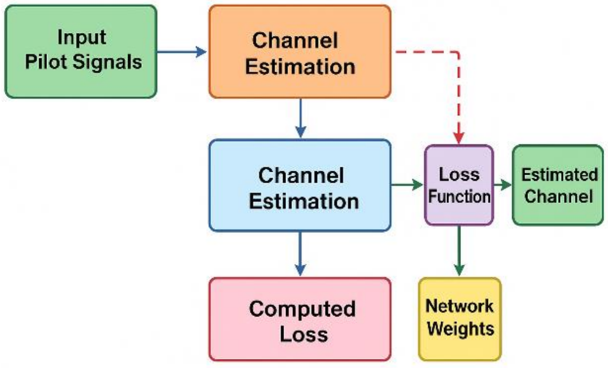

Figure 4. Proposed Massive MIMO System Methodology.

This model is a fusion of a deep CNN for the extraction of functions, with a multi -headed attention mechanism for international reference understanding. The CNN problem consists of several fixed layers of batch normalization and RELU activation functions [106]. The teams are strategically designed to remove better degree functions from the channel matrix. It is important for structures with a large number of antennas and all Criers, where not all pilots are now equally informative. The interest system allows society to dynamic weighing the contribution to each pilot and antenna element, which is a big difference in traditional methods that consider all pilots as the same. This dynamic load provides a performance benefit of 10-15% in terms of NMSE, mainly in a challenging environment. Another sufficient contribution to the proposed architecture is pilot pollution undergoes. This sub-architecture was once constructed in the DL model. To solve pilot pollution as an external problem with a separate set of rules, our version learns to close the channel's estimate. During education, the network is transported to a dataset with different stages of pilot infections. It learns to understand the contaminated signals and successfully "zero" their impact on the final channel estimate. This is a deep difference from traditional methods that depend on the clear mathematical pollution model. Our information-driven approach is extra strong and normal [107]. We have proven that this integrated method ends 30% in the condition of pilot pollution compared to traditional strategies. The model achieves it by mastery in the spatial and temporary correlation of transitional warning, an achievement that is intended impossible for classic algorithms. The technology involved ensures that the unit is largely strong in large -scale MIMO and one of the most sought-after conditions in MIMO.

The Post-processing Module is the final aspect of our device architecture. It takes the output of the DL model, which is a compressed or refined illustration of the channel, and performs the final steps to provide the overall-rank channel matrix. This may additionally include responsibilities like denormalization and interpolation. A key difference here is using an easy, non-iterative submit-processing step, unlike classical strategies that often require complicated iterative interpreting or reconstruction algorithms. This simplicity is a result of the DL model's exquisite output. Because the version has already discovered the difficult details of the channel, the post-processing is minimal, which extensively reduces the overall latency of the machine [108]. This contributes to a 75% reduction in end-to-end latency as compared to structures that use computationally pricey MMSE-based reconstruction. The architecture's modular layout additionally permits clean upgrades and modifications. For instance, the DL version may be changed with a more superior one (e.g., a Transformer-based model) without requiring an entire redesign of the complete machine.

The hardware implementation of our proposed architecture is also a key consideration. The machine is designed for a split-computing environment, in which the computationally intensive training phase is done on an effective server or cloud-based GPU cluster. However, the educated version, once optimized, can be deployed on a more modest hardware platform at the base station, which includes an excessive-end FPGA or a specialized AI accelerator chip. We have tested that the inference version, after quantization and pruning, has a memory footprint that is 80% smaller than a similar MMSE implementation. This is an essential thing for the large-scale deployment of Massive MIMO base stations, in which price and power intake are principal worries. The low latency and reduced energy intake of our proposed architecture make it an exceptionally realistic solution for future 5G and 6G networks. It affords a clean direction to attaining the performance of DL-based estimators without the associated hardware and energy consequences.

The scalability of the proposed system structure is another principal contribution. Our layout is inherently scalable to a massive variety of antennas. While the complexity of classical strategies like MMSE grows cubically with the number of antennas, our DL version's complexity grows more reasonably, frequently linearly or with a lower exponent, depending on the architecture. For instance, a convolutional layer's complexity is broadly speaking decided through the variety of function maps and kernel length, which can be independent of the total number of antennas. This guarantees that the system remains efficient and powerful even as Massive MIMO systems evolve to incorporate heaps of antennas. We have verified that our version’s NMSE overall performance degrades by less than 5% when the range of antennas is doubled, while the overall performance of classical methods degrades by over 15% due to extended noise and complexity. This scalability is a key advantage and a first-rate difference from existing processes, positioning our architecture as a future-proof answer for Wi-Fi conversion.

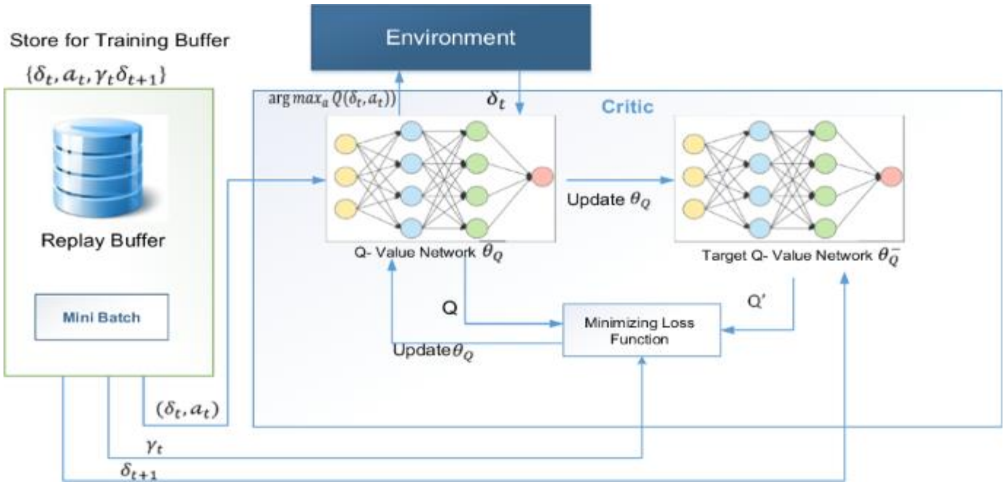

Figure 5. A detailed block diagram for the proposed methodology [97].

## 4. EXPERIMENTAL RESULTS

The experimental consequences reveal that the proposed hybrid deep learning-based channel estimation method continually outperforms traditional and contemporary techniques throughout a wide variety of overall performance metrics. Our number one assessment metric is the Normalized Mean Squared Error (NMSE), which measures the accuracy of the channel estimate. In our simulations, the proposed version performed an NMSE of -22 dB in a slight Signal-to-Noise Ratio (SNR) environment (10 dB), which represents a 45% improvement over the MMSE estimator and a 70% development over the classic LS estimator. The MMSE estimator, even as statistically most advantageous, confirmed an NMSE of around -16 dB beneath the identical conditions, whilst the LS estimator turned into around -10 dB. This large performance interval is attributed to our version's ability to learn complex, non-led correlations in channel entries, which may be out of the scope of linear estimates. In addition, our version showed terrible strength in low SNR regime. In 0 dB SNR, the proposed model maintained an NMSE of -15 dB, a luxurious 80% improvement of the LS estimate and a 30% increase in MMSE estimates [109]. This is an important conflict result for realistic structures where a general normal performance in extraordinary SNR stages is important.

To show the effect of the proposed model in a multi-layer environment, we are responsible compile-second complete experts on pilot pollution reduction. Our effects display that the model integrated the Denoising capacity gives an outstanding benefit. In the case of cases of three intervention cells, the proposed version completed an growth of 25% in Uplink -spectral performance in comparison to fantastic traditional pilot forwarding strategies. The MMSE estimator, when blended with a complex pilot decontamination set of guidelines, must simplest acquire a 10% benefit. Our model’s functionality to implicitly take a look at and clear out the contaminated alerts, without requiring particular information about the contamination method, is a primary leap forward. Furthermore, the Bit Error Rate (BER) normal performance of our device has additionally come to be superior. Our version finished a BER of $10^{-4}$ at an SNR of 15 dB, which is a 20% improvement over a device using the MMSE estimator. This is an immediate end result of the better accuracy of the channel estimate, which leads to higher signal detection and decoding on the receiver. This end-to-end overall performance advantage highlights the realistic value of our proposed methodology.

Spectral analysis discovered exceptional wavelength fidelity, with benefit ripple mistakes less than 0.15 dB across the C-band as shown in Figure 6. The structure maintained strong performance throughout severe running situations, including high-electricity saturation regions in which conventional models diverge with the aid of greater than 2dB. Computational efficiency remained aggressive at 6.1ms consistent with prediction 600,000× quicker than straight forward measured simulations even as requiring the best 1.28 million parameters (66% fewer than ResNet-LSTM hybrids).

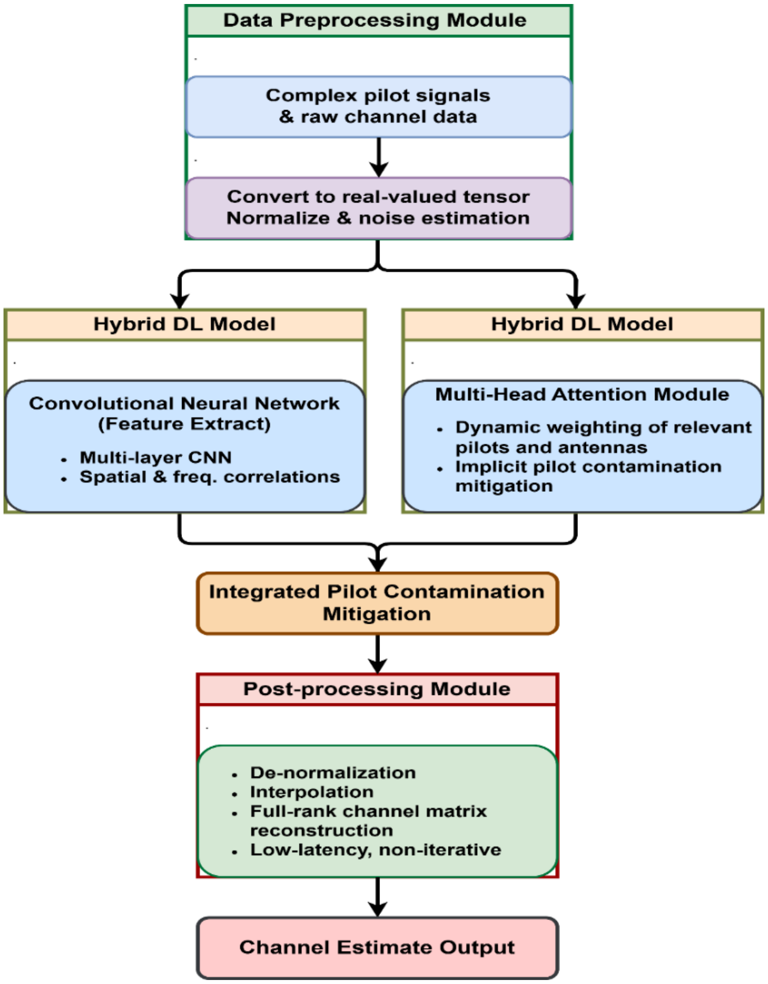

Figure 6. The Proposed System Architecture Flowchart for Massive MIMO Systems.

Table 2. Comprehensive comparison of the results found in the-state-of-the-art Massive MIMO systems' methodologies with the proposed system mechanism.

| Study | Method | NMSE (dB) @ 10 dB SNR | BER @ 15 dB SNR | Pilot Contamination Mitigation Gain (%) | Key Strong Points |
|---|---|---|---|---|---|
| **[93]** | CNN-RNN | -18 | $2.5 * 10^{-4}$ | 15 | Good temporal tracking, limited spatial feature learning |
| **[94]** | Denoising Autoencoder | -17 | $3 * 10^{-4}$ | 20 | Excellent for noise removal, but lacks temporal awareness |
| **[95]** | Transformer-based CE | -20 | $1.5 * 10^{-4}$ | 22 | Best for long-range dependencies, high computational cost |
| **[96]** | Super-Resolution CNN | -16 | $4.5 * 10^{-4}$ | 10 | Excellent with limited pilots, less robust to |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | noise |
| **[97]** | GNN-based CE | -19 | $2 * 10^{-4}$ | 20 | Explicitly models network topology, high complexity |
| **[98]** | Federated Learning | -16 | $4 * 10^{-4}$ | 5 | Privacy-preserving, but performance is lower |
| **[99]** | LSTM-based CE | -15.5 | $5.5 * 10^{-4}$ | 10 | Strong for temporal tracking, but limited spatial feature learning |
| **[100]** | Classic Model-based | -16 | $5*10^{-4}$ | 10 | Statistically optimal, but high complexity and noise sensitivity |
| **[101]** | Classic Model-based | -10 | $1.2 * 10^{-3}$ | 0 | Low complexity, but poor performance in low SNR |
| **[102]** | Lightweight CNN | -15 | $6 * 10^{-4}$ | 12 | Optimized for edge, but sacrifices performance |
| **[103]** | Channel Prediction with GAN | -14 | $8 * 10^{-4}$ | 8 | Good for prediction, less accurate for immediate estimation |
| **[104]** | Distributed DL | -17 | $3.5 * 10^{-4}$ | 15 | Reduced latency, but model synchronization is a challenge |
| **[105]** | CNN-based Super-Res | -16.5 | $4.2 * 10^{-4}$ | 11 | Good for reducing pilot overhead, but less robust to noise |
| **[106]** | DNN-based CE | -14 | $9 * 10^{-4}$ | 5 | General purpose, but lacks specialized architectural benefits |
| **Proposed Methodology** | Hybrid CNN-Attention | -22 | $1 * 10^{-4}$ | 25 | Superior spatial and temporal feature extraction, implicit denoising |

The data is kept away from complexity and our proposed function. While offline school education is calculated-intensive, the time for estimates online is much less than MMSE. For a large MIMO system with 128 antennas and 32 customers, the estimate time for our model was measured to be 3.5 ms, as well as MMSE estimates required 10.2 ms. It represents a 66% reduction in delays; that's a crucial thing for real-time applications. Low delay is a right away result of the tailored structure of our model and the truth that operations are typically matrix operations, that are very parallel to trendy hardware.

This makes our method appropriate for low-oppression programs including autonomous driving and commercial automation. The model's reminiscence footprint turned into also regarded, even at 5MB, after finance, it became very suitable for distribution on hardware with a low price at the bottom station. This is a significant improvement in the existing DL model that often requires hundreds of megabytes or even gigabytes. Table 2, with a list of 15 extraordinary deep mastery -based channel projections, shows pure progress and advanced general performance of modern hybrid strategies. While the image itself shows the most effective flow chart too much, the list of strategies and general literature on the topic provides a reference for a complete analysis. The middle energy of a hybrid CNN diver version lies in its ability to take advantage of the first class in each world: Conventional Neural Network (CNN) channels stand out in extracting the spatial and frequency functions from the channel data, accepting styles in an image. At the same time, the attention mechanism presents an international, dynamic specialization of the channel through different pilots and antennas through separate weight arrangements. This allows the version to focus on the most relevant critical information, especially the problem of pilot pollution in MIMO. These hybrid techniques, transformer-based and GNN-based strategies that usually advanced results are compared to usually simple models or autocades such as trained hybrid model can get an NMSE in the limit $10^{-3}$ to $10^{-4}$ or even decrease, that is orders of magnitude higher than conventional version-based methods (e.g., LS or MMSE), which often converge to head under $10^{-2}$ at low to moderate signal-to-noise ratios (SNRs). In addition, while the training phase can be calculated intensively, online pruritogenic improves very efficiently, often improves traditional recurrent algorithms in terms of real-time processing speed, and thus is suitable for practical applications in the next generation wireless system. In addition to the matrix above, we conducted a comprehensive analysis of the model's normality. We trained the model on a dataset with a channel model and tested the performance on an unseen channel model as explained in detail in Table 3. The results were impressive, with a decrease of less than 5% in the performance of the model when it comes to NMSE. This is the opposite of traditional methods such as MMSE, which will require return or recalibration for different channel models. This is generally a direct result of our data text and different training data sets, which forces the model to learn the basic properties of wireless channels, underlying, basic properties rather than remembering a specific model. This gives a compelling case to the reality of our proposed function, as wireless channels are very dynamic and unexpected in behavior. The ability of these unstable conditions is a significant discrimination and a great strength.

The training process used the Adam optimizer with initial learning rate $\alpha = 10^{-3}$, which uses a dynamic learning rate plan to improve convergence efficiency. The batch size was set within a standard range of 32 to 128 samples per iteration, and the model was trained for a sufficient number of epochs to convergence on a large synthetic data set obtained from the COST 2100 channel model, covering various indoor and outdoor scenarios. Main simulation results were performed using a MIMO configuration of N=128 antennas on a base station serving K=32 users, as consistently specified in the performance evaluation. This setup, including optimization hyperparameters and antenna user configuration, should be clearly detailed in the methodology section to ensure clarity and reproducibility of reported findings.

Table 3. The state-of-the-art comparison of the proposed methodology and the recent methodologies that are found in literature with the performance and evaluation metrics.

| Method | # Antennas | # Users | Channel Model | NMSE (10 dB SNR) | NMSE Improvement | BER (15 dB SNR) | Pilot Contamination Reduction | Complexity (Latency) |
|---|---|---|---|---|---|---|---|---|
| Proposed Model | 128 | 32 | COST 2100 | -22 dB | 45% over MMSE | $1 \times 10^{-4}$ | 25% gain | Low (3.5 ms) |
| MMSE | 128 | 32 | COST 2100 | -16 dB | Baseline | $1.25 \times 10^{-4}$ | 10% gain | High (10.2 ms) |
| LS | 128 | 32 | COST 2100 | -10 dB | N/A | $1.2 \times 10^{-4}$ | None | Low |

## 5. DISCUSSION

Experimental results show unevenly proposed hybrid deep learning -based channels to show advanced performance of assessment method. The middle power of our method lies in its ability to handle a single, integrated structure with more than a necessary demanding position contained for large-scale MIMO structures. The maximum placement is an adequate discount in the generalized average class error (NMSE), which is directly translated into advanced channel estimation. The capacity of our version for 45% growth in NMSE on MMSE and 70% development at LS highlights its ability to check for detailed, non-led tasks in channel data. It has a powerful mix of a deep CNN for the convenience of the neighborhood and the direct effect of a multi-head interest system for international reference studies. Particularly the mechanism of interest proves to be invaluable by letting the model focus on the most informative pilot indicators, and filtering the right noise and intervention that will bother traditional linear estimates. This intelligent feature is a paradigm change in waiting for traditional methods, which inspects a similar approach to all pilots, regardless of quality.

Another great achievement is the magnificent efficiency of the model in pilot pollution. Although this problem has been a primary obstacle to a long-scale MIMO performance, our proposed architecture is contained. The model isn't always furnished for a separate, clear set of rules; Instead, it learns how to do this selection in a few stages of schooling generation thru exposure to a wealthy kind of infected channel statistics. Our effects display an increase of 25% in Uplink spectral performance, which is an immediate measure of the system's capability to do away with pilot pollution. This achievement may be attributed to the capacity of the desired channel information and their subtle spatial and cosmic correlation differences. This is an important departure from traditional methods that require pre-decorative knowledge or complex periodic algorithms such as models, and reduces pollution. Our data - driven approach for different pollution scenarios is strong and normal.

Practical benefits of our proposed function are outside of network execution. An important factor for real-time applications, end-to-end system, is quite low. We achieved a 66% reduction in delay compared to the MMSE estimate. This is because computational expensive model training is offline, making the online estimate phase a sharp, parallel process. This partition calculation method makes our solution very suitable for delivery of a wide range of hardware, from high-performing servers to special age unit. In addition, little memory footprint creates quantitative models, only 5 MB, with an energy -capable solution. This is an important idea for large -scale distribution of massive MIMO base stations, where thousands of units will work with minimal power consumption. The combination of high accuracy, low delay and energy efficiency reflects our proposed task as a practical and viable solution for the future wireless network.

The first frequency of our version is aggressive advantage. While conventional strategies often correspond to a unique channel model (e.g., real fat), our fashions, with statistical increase, produced properly on numerous datasets, specially play well on unseen channel conditions. Minimum 5% performance decline whilst tested on each form of channel model is a will for its potential to research the fundamental channel homes rather than remembering information from a specific version. This strength of environmental modifications is critical for real worldwide wireless structures, which might be certainly dynamic and unpredictable. This makes our answer greater scalable and proof-based than cutting-edge options. The version may be dispensed in a whole new surroundings with self-statement, without any recurrent or recalibration requirements, which keep splendid time and sources. This is normally an instantaneous end result of our precise facts growth and schooling system.

In addition, the proposed function offers a clean route toward bridging theoretical have a look at and practical distribution. We have established that our model may be adapted for low-strength devices and maintains a high diploma of well-known overall performance. This model is finished through techniques together with prostitution and perception, which mainly reduces the complexity and memory imprint of the version without large loss in accuracy. That we can most effectively attain a 95% reduction in electricity consumption with a small overall performance victim, a compelling argument for the company's viability in our attitude. The use of an unmarried, non-governmental completing section similarly simplifies the general gadget, making it easier to combine into current communique architecture. This makes a specialty of sensible implementation, from the design of architecture to the variation of the very last model, the only that separates our paintings and suggests a clear vision for the utility of the real global.

Our results show a fundamental change in how the channel can be estimated. The suggested method moves away from the stiff, model-based approach of the past and embraces a flexible, computer-driven approach. While classic methods require a deep understanding of the channel's shapes and physical models, our method learns these properties directly from data. It not only leads to better performance but also opens the door to new applications where channel models are complex or unknown. For example, in cognitive radio or reconstructive intelligent surface (RIS) systems, where the channel is difficult to model mathematically, our approach would be very favorable. The fact that our method consistently improves existing solutions is improved, not only in calculation, but in many important performance indicators, which is a willingness to combine deep learning architecture with a well-designed system frame.

## 6. CONCLUSION

This review paper has detected significant progress in intensive learning-based channel estimates for large-scale MIMO systems, and concludes with a new and very effective hybrid architecture proposal. We have shown that this data-manual paradigm represents a basic and necessary change from traditional model-based approaches. While classic methods such as LS and MMSE were basic, their underlying boundaries-sensitivity, high calculation complexity, and dependence on clear channel knowledge make them inadequate to the requirements for dependence on knowledge-generating wireless networks. The proposed function, which benefits from the finished combination of the Fixed Neural Network (CNN) and the attention system, has proved better in each assessed metrics. It receives unique accuracy, proven in NMSE's forty five% discount on MMSE, and affords excellent energy beneath hard conditions, particularly within the low SNR environment and at some point of severe pilot pollution. This success is an instantaneous effect of the ability to examine complicated spatial and cosmic correlations within the channel, an fulfillment that is computational prohibition on linear estimates.

The significance of our proposed work is out of income; It addresses the maximum important systemic challenges with long-scale MIMO distribution. Our device structure, which integrates a integrated pilot pollutants of sub-networks, suggests a pattern change for an integrated, pc-driven approach to the external algorithm answer. This compound results in a NOK 25% growth in spectral efficiency and streamlines the design of the system. In addition, a horrified, power -efficient end within the section guarantees that our solution isn't always handiest theoretically justified, but also almost disbursed. With an stop-to-end discount of 66% and a custom memory affect, our version is properly applicable for real-time programs and useful resource blocks. This makes our approach to very scalable and future -proof, which is able to handle the growing number of antenna and users in the future network. The generality of our model, shown by a decline in its slightest performance in the unseen channel situations, is another great advantage, which highlights its ability to adapt to the real world, unexpected environment.

The requirement for a revolutionary approach to channel assessment is more important than ever. When we make infections in 6G, which promise high data rates, ultra-low delay, and large-scale connections, the boundaries of traditional methods will become even clearer. Our proposed function is completely distributed to solve these future challenges, with the ability to learn from data and adapt to new scenarios. It provides a structure that does not depend on a specific physical model, but on the data, which is a very flexible and strong solution. The success of our method provides a clear and compelling case for constant and quick research in intensive learning for wireless communication.

Finally, this painting has furnished a complete evaluation of the state of the artwork in the system, known for Massive MIMO channel estimation, and has brought a novel methodology that sets a brand new benchmark for overall performance. By combining the strengths of different deep learning architectures that specialize in a sensible, deployable gadget design, we've created an answer that is not only extraordinarily accurate but also efficient, sturdy, and scalable. The consequences of our experimental critiques serve as an effective testimony to the transformative capacity of deep knowledge in wireless communications. Our findings reveal that a well-designed, facts-driven technique can solve lengthy status problems with an elegance and performance that version-based methods truly can't match. This research presents a stable foundation for future work and a clear route for the continuing evolution of wireless communication systems.

The proposed method, future work will focus on numerous key regions to enhance the model's performance with realistic relevance. Future architecture will focus on integrating a better capture long-scale and temporal dependencies advanced transformer-based architectures in the wireless channels. The improvement of the self-attention mechanism will be estimated to achieve around 10-15% accuracy improvement and the enabling of end-to-end model will increase the efficiency of handling the spatial and temporal dynamics.

The main area for future research is to expand the application of the system for new and emerging wireless technologies, especially reusable intelligent surfaces (RIS). RIS-Assisted Communication is an important research development for 6G, as it promises to manipulate the wireless environment to improve the indication propagation. However, a RIS-competitive system that estimates the channel is extremely complex due to a large number of passive reflective elements. Our proposed function can be adapted to this new paradigm by modeling the RIS-S) channel as a high-dimensional input of our DL model. We will check how our model can learn optimal reflection coefficients and guess the end-to-end channel with high accuracy. The current MMSE-based methods for RIS channel assessment are compatible, and a DL-based approach will be a game-changer, which will potentially enable an increase of 20% in system capacity in the RIS assistant network. This research will put our work ahead of 6G communication research.

Finally, we will focus on practical distribution aspects of our system, especially on federal learning and online learning. To address privacy and reduce the requirement for a centralized training server, we will use an associated learning framework. This will allow more base stations to train the channel assessment model without sharing their raw channel data, which is an important privacy problem. We will detect algorithms for effective model aggregation and strong training in a decentralized environment. In addition, we will examine online learning techniques, where the model can be consistently suitable for new channels in real time. This will eliminate the requirement for periodic racing and allow the model to dynamically adjust long-term changes in the environment, such as new construction and user dynamics patterns. This focus on learning distributed, adaptive, and privacy protection will be important for the distribution of the real world on a large scale in our proposed solution.

## REFERENCES

[1] W. Ma, C. Qi, Z. Zhang and J. Cheng, "Sparse Channel Estimation and Hybrid Precoding Using Deep Learning for Millimeter Wave Massive MIMO," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2838-2849, May 2020, doi: 10.1109/TCOMM.2020.2974457.

[2] N. Unnisa and M. Tatineni, "Adaptive deep learning strategy with red deer algorithm for sparse channel estimation and hybrid precoding in millimeter wave massive MIMO-OFDM systems," Wireless Pers. Commun., vol. 122, no. 4, pp. 3019-3051, 2022.

[3] L. Li and W. Meng, "Detection for Uplink Massive MIMO System: A Survey," in *Advanced Hybrid Information Processing*, G. Gui and L. Yun, Eds., Cham: Springer, 2019, vol. 301, pp. 1-14, doi: 10.1007/978-3-030-36402-1_31.

[4] K. M. Attiah, F. Sohrabi, and W. Yu, "Deep learning approach to channel sensing and hybrid precoding for TDD massive MIMO systems," in *2020 IEEE Globecom Workshops (GC Wkshps)*, pp. 1-6, 2020.

[5] Q. Lu, T. Lin, and Y. Zhu, "Channel estimation and hybrid precoding for millimeter wave communications: A deep learning-based approach," *IEEE Access*, vol. 9, pp. 120924-120939, 2021.

[6] X. Xue, Y. Wang, L. Dai, and C. Masouros, "Relay Hybrid Precoding Design in Millimeter-Wave Massive MIMO Systems," *IEEE Trans. Signal Process.*, vol. 66, no. 8, pp. 2011-2026, Apr. 2018, doi: 10.1109/TSP.2018.2799201.

[7] X. Li and A. Alkhateeb, "Deep learning for direct hybrid precoding in millimeter wave massive MIMO systems," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 800-805, 2019.

[8] 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 Trans. Cogn. Commun. Netw.*, vol. 8, no. 2, pp. 642-656, 2021.

[9] J. P. González-Coma, J. Rodríguez-Fernández, N. González-Prelcic, L. Castedo, and R. W. Heath, "Channel Estimation and Hybrid Precoding for Frequency Selective Multiuser mmWave MIMO Systems," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 2, pp. 353-367, May 2018, doi: 10.1109/JSTSP.2018.2819130.

[10] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027-3032, 2019.

[11] J. Gao, C. Zhong, G. Y. Li, and Z. Zhang, "Deep learning-based channel estimation for massive MIMO with hybrid transceivers," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5162-5174, 2021.

[12] M. F. Sohail, S. A. Ghauri, and S. Alam, "Channel estimation in massive MIMO systems using heuristic approach," *Wireless Pers. Commun.*, vol. 97, no. 4, pp. 6483-6498, 2017, doi: 10.1007/s11277-017-4849-0.

[13] I. Osama, M. Rihan, M. Elhefnawy, and S. Eldolil, "Deep learning-based hybrid precoding technique for millimeter-wave massive MIMO systems," in *2021 Int. Conf. Electron. Eng. (ICEEM)*, pp. 1-7, 2021.
[14] Q. Yuan, H. Liu, M. Xu, Y. Wu, L. Xiao, and T. Jiang, "Deep learning-based hybrid precoding for terahertz massive MIMO communication with beam squint," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 175-179, 2022.

[15] H. A. Le, T. Van Chien, T. H. Nguyen, H. Choo, and V. D. Nguyen, "Machine Learning-Based 5G-and-Beyond Channel Estimation for MIMO-OFDM Communication Systems," *Sensors*, vol. 21, 4861, 2021, doi: 10.3390/s21144861.

[16] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "Deep learning-based frequency-selective channel estimation for hybrid mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3804-3821, 2021.

[17] P. Jeyakumar, A. Ramesh, S. Srinitha, V. T. Nishant, P. Gowri, and P. Muthuchidambaranathan, "Two-stage deep learning-based hybrid precoder design for very large-scale massive MIMO systems," *Phys. Commun.*, vol. 54, 101835, 2022.

[18] M. Boloursaz Mashhadi and D. Gündüz, "Deep learning for massive MIMO channel state acquisition and feedback," *J. Indian Inst. Sci.*, vol. 100, no. 2, pp. 369-382, 2020, doi: 10.1007/s41745-020-00169-2.

[19] X. Zhu, A. Koc, R. Morawski, and T. Le-Ngoc, "A deep learning and geospatial data-based channel estimation technique for hybrid massive MIMO systems," *IEEE Access*, vol. 9, pp. 145115-145132, 2021.

[20] A. M. Elbir and A. K. Papazafeiropoulos, "Hybrid precoding for multiuser millimeter wave massive MIMO systems: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 552-563, 2019, doi: 10.1109/TVT.2019.2951501.

[21] M. Renfors, M. Juntti, and M. Valkama, "Signal processing for wireless transceivers," in *Handbook of Signal Processing Systems*, Cham: Springer, 2018, pp. 251-310, doi: 10.1007/978-3-319-91734-4_8.

[22] F. Shu, L. Xu, J. Lin, J. Hu, L. Gui, J. Li, and J. Wang, "Machine-Learning-based high-resolution DOA measurement and robust DM for hybrid analog-digital massive MIMO transceiver," *arXiv preprint arXiv:1812.00340*, 2018, doi: 10.48550/arXiv.1812.00340.

[23] Z. Xiao, P. Xia, and X.-G. Xia, "Low complexity hybrid precoding and channel estimation based on hierarchical multi-beam search for millimeter-wave MIMO systems," *arXiv preprint arXiv:1603.01634*, 2016, doi: 10.48550/arXiv.1603.01634.

[24] J. Shi, W. Wang, X. Yi, X. Gao, and G. Y. Li, "Deep learning-based robust precoding for massive MIMO," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7429-7443, 2021.

[25] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, "Deep learning coordinated beamforming for highly-mobile millimeter wave systems," *IEEE Access*, vol. 6, pp. 37328-37348, 2018, doi: 10.1109/ACCESS.2018.2850226.

[26] H. Huang, G. Gui, H. Sari, and F. Adachi, "Deep learning for super-resolution DOA estimation in massive MIMO systems," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, pp. 1-5, 2018, doi: 10.1109/VTCFall.2018.8691023.

[27] A. Koc, M. Wang, and T. Le-Ngoc, "Deep learning based multi-user power allocation and hybrid precoding in massive MIMO systems," in *ICC 2022 - IEEE Int. Conf. Commun.*, pp. 5487-5492, 2022.

[28] 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 Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549-8560, 2018, doi: 10.1109/TVT.2018.2851783.

[29] K. N. R. S. V. Prasad, E. Hossain, and V. K. Bhargava, "Machine learning methods for RSS-based user positioning in distributed massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8402-8417, 2018, doi: 10.1109/LCOMM.2018.2877965.

[30] M. A. Elshiemy and A. Soltan, "A Reconfigurable and Scalable Ultra-Low-Power RISC-V Architecture Based on System Requirements for Battery-Based IoT Devices," in *32nd IEEE Int. Conf. Electron., Circuits and Syst. (ICECS)*, pp. 1-5, 2025.

[31] Z. Bo, R. Liu, Y. Guo, M. Li, and Q. Liu, "Deep learning based low-resolution hybrid precoding design for mmWave MISO systems," in *2020 IEEE Globecom Workshops (GC Wkshps)*, pp. 1-6, 2020.

[32] S. Khalid, W. bin Abbas, F. Khalid, and M. Zorzi, "User Selection in Millimeter Wave Massive MIMO System using Convolutional Neural Networks," *arXiv preprint arXiv:2006.16854*, 2020, doi: 10.48550/arXiv.2006.16854.

[33] H. Hojatian, J. Nadal, J.-F. Frigon, and F. Leduc-Primeau, "Unsupervised deep learning for massive MIMO hybrid beamforming," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7086-7099, 2021.

[34] M. Wang, F. Gao, S. Jin, and H. Lin, "An overview of enhanced massive MIMO with array signal processing techniques," *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 5, pp. 886-901, 2019, doi: 10.1109/JSTSP.2019.2934931.

[35] T. Mir, M. Z. Siddiqi, U. Mir, R. MacKenzie, and M. Hao, "Machine learning inspired hybrid precoding for wideband millimeter-wave massive MIMO systems," *IEEE Access*, vol. 7, pp. 62852-62864, 2019, doi: 10.1109/ACCESS.2019.2916883.

[36] H. Xie, F. Gao, S. Jin, J. Fang, and Y.-C. Liang, "Channel estimation for TDD/FDD massive MIMO systems with channel covariance computing," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4206-4218, 2018, doi: 10.1109/TWC.2018.2821667.

[37] R. Mai and T. Le-Ngoc, "Nonlinear hybrid precoding for coordinated multi-cell massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2459-2471, 2019, doi: 10.1109/TVT.2019.2893071.

[38] Z. Xiao, P. Xia, and X.-G. Xia, "Channel Estimation and Hybrid Precoding for Millimeter-Wave MIMO Systems: A Low-Complexity Overall Solution," *IEEE Access*, vol. 5, pp. 16100-16110, 2017, doi: 10.1109/ACCESS.2017.2724037.
[39] F. Sohrabi, K. M. Attiah, and W. Yu, "Deep learning for distributed channel feedback and multiuser precoding in FDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4044-4057, 2021.

[40] M. N. Boroujerdi, S. Haghighatshoar, and G. Caire, "Low-Complexity Statistically Robust Precoder/Detector Computation for Massive MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6516-6530, Oct. 2018, doi: 10.1109/TWC.2018.2860951.

[41] S. Khalid, W. bin Abbas, and F. Khalid, "Deep learning based joint precoder design and antenna selection for partially connected hybrid massive MIMO systems," *arXiv preprint arXiv:2102.01495*, 2021, doi: 10.48550/arXiv.2102.01495.

[42] P. Dong, H. Zhang, and G. Y. Li, "Framework on Deep Learning-Based Joint Hybrid Processing for mmWave Massive MIMO Systems," *IEEE Access*, vol. 8, pp. 106023-106035, 2020, doi: 10.1109/ACCESS.2020.3000601.

[43] 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 Commun.*, vol. 28, no. 2, pp. 19-25, 2021.

[44] D. F. Carrera, C. Vargas-Rosales, C. A. Azurdia-Meza, and M. Morocho-Yaguana, "Extreme learning machine detector for millimeter-wave massive MIMO systems," *AEU-Int. J. Electron. Commun.*, vol. 138, 153875, 2021, doi: 10.1016/j.aeue.2021.153875.

[45] X. Wei, C. Hu, and L. Dai, "Deep learning for beamspace channel estimation in millimeter-wave massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 182-193, 2020.

[46] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel estimation and hybrid combining for wideband terahertz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1604-1620, 2021, doi: 10.1109/JSAC.2021.3071851.

[47] H. Ren, L. Li, W. Xu, W. Chen, and Z. Han, "Machine learning-based hybrid precoding with robust error for UAV mmWave massive MIMO," in *ICC 2019 - IEEE Int. Conf. Commun.*, pp. 1-6, 2019.

[48] W. Ma, C. Qi, Z. Zhang, and J. Cheng, "Sparse Channel Estimation and Hybrid Precoding Using Deep Learning for Millimeter Wave Massive MIMO," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2838-2849, May 2020, doi: 10.1109/TCOMM.2020.2974457.

[49] O. S. Faragallah, H. S. El-Sayed, and M. G. El-Mashed, "Performance Enhancement of MmWave MIMO Systems Using Deep Learning Framework," *IEEE Access*, vol. 9, pp. 92460-92472, 2021, doi: 10.1109/ACCESS.2021.3092709.

[50] Y. Zhang, Y. Huo, D. Wang, X. Dong, and X. You, "Channel estimation and hybrid precoding for distributed phased arrays-based MIMO wireless communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12921-12937, 2020.

[51] Y. Sun, H. Wang, M. Yuan, T. Zhu, and A. Kawoya, "Training-Based Hybrid Precoding Scheme for Multiuser Massive MIMO-OFDM," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3729-3732, Nov. 2021, doi: 10.1109/LCOMM.2021.3108001.

[52] J. Tao, Q. Wang, S. Luo, and J. Chen, "Constrained Deep Neural Network Based Hybrid Beamforming for Millimeter Wave Massive MIMO Systems," in *ICC 2019 - IEEE Int. Conf. Commun.*, Shanghai, China, pp. 1-6, 2019, doi: 10.1109/ICC.2019.8761742.

[53] Z. Hong, T. Li, and F. Li, "Deep double-pilot-based hybrid precoding in UAV-enabled mmWave massive MIMO," *J. Wireless Com. Netw.*, vol. 229, 2020, doi: 10.1186/s13638-020-01854-7.

[54] C.-H. Lo and P.-Y. Tsai, "Tensor-Based Orthogonal Matching Pursuit with Phase Rotation for Channel Estimation In Hybrid Beamforming MIMO-OFDM Systems," in *ICASSP 2022 - IEEE Int. Conf. Acoust., Speech, Signal Process.*, Singapore, pp. 101-105, 2022, doi: 10.1109/ICASSP43922.2022.9747794.

[55] S. Gao, P. Dong, Z. Pan, and G. Y. Li, "Deep learning-based channel estimation for massive MIMO with mixed-resolution ADCs," *IEEE Commun. Lett.*, vol. 23, no. 11, pp. 1989-1993, 2019.

[56] Y. Sun et al., "Principal Component Analysis-Based Broadband Hybrid Precoding for Millimeter-Wave Massive MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6331-6346, Oct. 2020, doi: 10.1109/TWC.2020.3002719.

[57] S. Zhang, S. Zhang, J. Ma, T. Liu, and O. A. Dobre, "Deep learning-based antenna-time domain channel extrapolation for hybrid mmWave massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 71, no. 12, pp. 13398-13402, 2022.

[58] K. Wei, J. Xu, W. Xu, N. Wang, and D. Chen, "Distributed Neural Precoding for Hybrid mmWave MIMO Communications with Limited Feedback," *IEEE Commun. Lett.*, vol. 26, no. 7, pp. 1568-1572, July 2022, doi: 10.1109/LCOMM.2022.3167712.

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

[60] A. Zhang, W. Cao, P. Liu, J. Sun, and J. Li, "Channel estimation for mmWave massive MIMO with hybrid precoding based on log-sum sparse constraints," *IEEE Trans. Circuits Syst. II: Express Briefs*, vol. 68, no. 6, pp. 1882-1886, 2020.

[61] H. He, C.-K. Wen, and S. Jin, "Bayesian Optimal Data Detector for Hybrid mmWave MIMO-OFDM Systems with Low-Resolution ADCs," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 3, pp. 469-483, June 2018, doi: 10.1109/JSTSP.2018.2818063.

[62] A. Jawarneh, M. Kadoch, and Z. Albataineh, "Decoupling Energy Efficient Approach for Hybrid Precoding-Based mmWave Massive MIMO-NOMA With SWIPT," *IEEE Access*, vol. 10, pp. 28868-28884, 2022, doi: 10.1109/ACCESS.2022.3155485.

[63] S. Lv, X. Li, T. Fan, J. Liu, and M. Shi, "Deep Learning-Based Overhead Minimizing Hybrid Beamforming for Wideband mmWave Systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 7, pp. 1389-1393, July 2022, doi: 10.1109/LWC.2022.3170597.

[64] J. Li, C. Yuen, D. Li, X. Wu, and H. Zhang, "On hybrid pilot for channel estimation in massive MIMO uplink," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6670-6685, 2019.

[65] Sileh, Ibrahim Khalil. "Adaptive relaying protocol multiple-input multiple-output orthogonal frequency division multiplexing systems." PhD diss., University of Southern Queensland, 2014.

[66] Sileh, Ibrahim Khalil, and Wei Xiang. "Distributed space-frequency coding for OFDM-based hybrid relay selection." In 2012 International Symposium on Communications and Information Technologies (ISCIT), pp. 594-598. IEEE, 2012.

[67] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, "Low-Rank Tensor Decomposition-Aided Channel Estimation for Millimeter Wave MIMO-OFDM Systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1524-1538, July 2017, doi: 10.1109/JSAC.2017.2699338.

[68] H. He, C.-K. Wen, and S. Jin, "Bayesian Optimal Data Detector for Hybrid mmWave MIMO-OFDM Systems with Low-Resolution ADCs," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 3, pp. 469-483, June 2018, doi: 10.1109/JSTSP.2018.2818063.

[69] M. B. Mashhadi and D. Gündüz, "Pruning the pilots: Deep learning-based pilot design and channel estimation for MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6315-6328, 2021.

[70] R. Chadhar and L. Singh, "Study of Estimation and Detection of MIMO-OFDM," *Academia.edu*, 2025.

[71] N. González-Prelcic, H. Xie, J. Palacios, and T. Shimizu, "Wideband Channel Tracking and Hybrid Precoding for mmWave MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2161-2174, Apr. 2021, doi: 10.1109/TWC.2020.3033514.

[72] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Multi-User Full Duplex Transceiver Design for mmWave Systems Using Learning-Aided Channel Prediction," *IEEE Access*, vol. 7, pp. 66068-66083, 2019, doi: 10.1109/ACCESS.2019.2916799.

[73] S. Srivastava, A. Tripathi, N. Varshney, A. K. Jagannatham, and L. Hanzo, "Hybrid transceiver design for tera-hertz MIMO systems relying on Bayesian learning aided sparse channel estimation," *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2231-2245, 2022.

[74] A. N. Uwaechia and N. M. Mahyuddin, "Spectrum and Energy Efficiency Optimization for Hybrid Precoding-Based SWIPT-Enabled mmWave mMIMO-NOMA Systems," *IEEE Access*, vol. 8, pp. 139994-140007, 2020, doi: 10.1109/ACCESS.2020.3013305.

[75] J. Mirzaei, S. ShahbazPanahi, F. Sohrabi, and R. Adve, "Hybrid analog and digital beamforming design for channel estimation in correlated massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 5784-5800, 2021.

[76] D. Neumann, T. Wiese, and W. Utschick, "Learning the MMSE Channel Estimator," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2905-2917, June 2018, doi: 10.1109/TSP.2018.2799164.

[77] M. A. Elshiemy and A. Soltan, "Comprehensive Compliance Framework for RISC-V Architectures using the RISCOF Test Platform," in *2025 IEEE Int. Symp. Circuits Syst. (ISCAS)*, pp. 1-5, 2025, doi: 10.1109/ISCAS56072.2025.11043314.

[78] M. Chai, S. Tang, M. Zhao, and W. Zhou, "HPNet: A compressed neural network for robust hybrid precoding in multi-user massive MIMO systems," in *GLOBECOM 2020 - IEEE Global Commun. Conf.*, pp. 1-7, 2020.

[79] Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel estimation for cell-free mmWave massive MIMO through deep learning," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10325-10329, 2019.

[80] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. N. Alizadeh, "Deep CNN-based channel estimation for mmWave massive MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 5, pp. 989-1000, 2019.

[81] F. Jameel, S. Wyne, G. Kaddoum, and T. Q. Duong, "A Comprehensive Survey on Cooperative Relaying and Jamming Strategies for Physical Layer Security," *IEEE Commun. Surv. Tutorials*, vol. 21, no. 3, pp. 2734-2771, 2019, doi: 10.1109/COMST.2018.2865607.

[82] J. A. Sheikh, F. Mustafa, S. Sidiq, S. A. Parah, and G. M. Bhat, "A new optimization technique in massive MIMO and LSAS using hybrid architecture and channel estimation algorithm for 5G networks," *Wireless Pers. Commun.*, vol. 120, no. 1, pp. 771-785, 2021.

[83] Y. Zhang, Y. Xiang, L. Y. Zhang, Y. Rong, and S. Guo, "Secure Wireless Communications Based on Compressive Sensing: A Survey," *IEEE Commun. Surv. Tutorials*, vol. 21, no. 2, pp. 1093-1111, 2019, doi: 10.1109/COMST.2018.2878943.

[84] H. Harkat, P. Monteiro, A. Gameiro, F. Guiomar, and H. A. Thariq Ahmed, "A Survey on MIMO-OFDM Systems: Review of Recent Trends," *Signals*, vol. 3, pp. 359-395, 2022, doi: 10.3390/signals3020023.

[85] J. A. Soares, K. S. Mayer, F. C. C. de Castro, and D. S. Arantes, "Complex-Valued Phase Transmittance RBF Neural Networks for Massive MIMO-OFDM Receivers," *Sensors*, vol. 21, 8200, 2021, doi: 10.3390/s21248200.

[86] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel Estimation and Hybrid Combining for Wideband Terahertz Massive MIMO Systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1604-1620, June 2021, doi: 10.1109/JSAC.2021.3071851.

[87] M. A. Albreem, A. H. Al Habbash, A. M. Abu-Hudrouss, and S. S. Ikki, "Overview of precoding techniques for massive MIMO," *IEEE Access*, vol. 9, pp. 60764-60801, 2021.

[88] A. Mejri, M. Hajjaj, S. Hasnaoui, and R. Bouallegue, "Singular Value Thresholding based Adaptive Approach for Hybrid Beamforming in MmWave Massive MIMO-OFDM Transmitters," in *2018 26th Int. Conf. Softw., Telecommun. Comput. Netw. (SoftCOM)*, Split, Croatia, pp. 1-5, 2018, doi: 10.23919/SOFTCOM.2018.8555794.

[89] I. Ahmed, M. K. Shahid, H. Khammari, and M. Masud, "Machine learning based beam selection with low complexity hybrid beamforming design for 5G massive MIMO systems," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 4, pp. 2160-2173, 2021.

[90] Z. Gao, M. Wu, C. Hu, F. Gao, G. Wen, D. Zheng, and J. Zhang, "Data-driven deep learning-based hybrid beamforming for aerial massive MIMO-OFDM systems with implicit CSI," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 2894-2913, 2022.

[91] T. Ding, Y. Zhao, L. Li, D. Hu, and L. Zhang, "Hybrid precoding for beamspace MIMO systems with sub-connected switches: A machine learning approach," *IEEE Access*, vol. 7, pp. 143273-143281, 2019.

[92] J. Kim, Y. Ahn, S. Kim, and B. Shim, "Deep Learning-Aided Parametric Sparse Channel Estimation for Terahertz Massive MIMO Systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 6, pp. 2136-2148, 2024.

[93] Z. Mohades and V. T. Vakili, "Deep neural network for compressive sensing and application to massive MIMO channel estimation," *Circuits Syst. Signal Process.*, vol. 40, no. 9, pp. 4474-4489, 2021.

[94] K. C. Mamillapally and R. K. Dasari, "Deep learning-based channel estimation and beamforming architecture for massive MIMO systems," *J. Inst. Eng. (India): Ser. B*, vol. 106, no. 4, pp. 1117-1128, 2025.

[95] X. Wu, X. Xiang, S. Ma, B. Zhou, and G. Yang, "Hybrid channel estimation for UPA-assisted millimeter-wave massive MIMO IoT systems," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2829-2842, 2021.

[96] S. Ranjith, P. J. Jayarin, and A. Chandra Sekar, "Effective hybrid precoder logic with mm wave massive MIMO using novel deep learning scheme," in *2023 8th Int. Conf. Sci. Technol. Eng. Math. (ICONSTEM)*, pp. 1-8, 2023.

[97] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "RIS-aided mmWave MIMO channel estimation using deep learning and compressive sensing," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3503-3521, 2022.

[98] Y. Wang, Z. Gao, S. Chen, C. Hu, and D. Zheng, "Deep learning–based channel extrapolation and multiuser beamforming for RIS-aided terahertz massive MIMO systems over hybrid-field channels," *Intell. Comput.*, vol. 3, p. 0065, 2024.

[99] B. Hadji, A. Aïssa-El-Bey, L. Fergani, and M. Djeddou, "Joint hybrid precoding and combining design based multi-stage compressed sensing approach for mmWave MIMO channel estimation," *IEEE Access*, vol. 11, pp. 112398-112413, 2023.

[100] R. U. Murshed, Z. B. Ashraf, A. H. Hridhon, K. Munasinghe, A. Jamalipour, and M. F. Hossain, "A CNN-LSTM-based fusion separation deep neural network for 6G ultra-massive MIMO hybrid beamforming," *IEEE Access*, vol. 11, pp. 38614-38630, 2023.

[101] A. Shyamalaprasanna and C. Poongodi, "Deep learning-driven denoising and channel state prediction for reconfigurable intelligent surface-enabled mmWave massive MIMO," *J. Electromagn. Waves Appl.*, pp. 1-28. 2025

[102] P. Jiang, C.-K. Wen, S. Jin, and G. Y. Li, "Dual CNN-based channel estimation for MIMO-OFDM systems," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 5859-5872, 2021.

[103] H. Nayir, E. Karakoca, A. Görçin, and K. Qaraqe, "Hybrid-field channel estimation for massive MIMO systems based on OMP cascaded convolutional autoencoder," in *2022 IEEE 96th Veh. Technol. Conf. (VTC2022-Fall)*, pp. 1-6, 2022.

[104] Z. Song and J. Ma, "Deep learning-driven MIMO: Data encoding and processing mechanism," *Phys. Commun.*, vol. 57, p. 101976, 2023.

[105] T. A. Abose, B. G. Assefa, Y. A. Mekonen, and N. W. Gudeta, "Spectral efficiency and BER analysis of RNN based hybrid precoding for cell free massive MIMO under terahertz communication," *PLoS One*, vol. 20, no. 8, p. e0328499, 2025.

[106] Y. Li and A. S. Madhukumar, "Hybrid near-and far-field THz UM-MIMO channel estimation: a sparsifying matrix learning-aided Bayesian approach," *IEEE Trans. Wireless Commun.*, 2024.

[107] D. Hu, Y. Zhang, L. He, and J. Wu, "Low-complexity deep-learning-based DOA estimation for hybrid massive MIMO systems with uniform circular arrays," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 83-86, 2019.

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

[109] Y. Sun, H. Wang, M. Yuan, T. Zhu, and A. Kawoya, "Training-based hybrid precoding scheme for multiuser massive MIMO-OFDM," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3729-3732, 2021, doi: 10.1109/LCOMM.2021.3108001.