Article that has been accepted for inclusion in a future issue of a journal. Content is final as presented, with the exception of pagination.

Pilot-Efficient Channel Estimation for 6G MU-mMIMO Using LS Initialization and GAN-Augmented CNN Refinement

Pooya Hejazi, Masoud Sabaei, and Ali Hajmahmoudi

***Abstract*—Accurate channel estimation with limited pilot overhead is a key bottleneck for 6G MU-mMIMO, particularly at mmWave (with implications for THz) where path loss and multi-user interference are severe. We propose a two-stage least squares convolutional neural network (LS-CNN) framework in which a robust LS estimator provides a stable coarse channel estimate, avoiding the sparsity assumptions of compressed sensing algorithms that often fail in rich-scattering outdoor environments. In the second stage, this estimate is refined using a CNN trained on a hybrid dataset expanded through controlled augmentation of a generative adversarial network (GAN), improving generalization while minimizing pilot overhead. Training data is refined using mean squared error (MSE) thresholds to ensure high-quality samples, and a robust standardization method is used to stabilize learning with complex-valued input. Simulations on the DeepMIMO ray-tracing dataset demonstrate that the proposed approach provides an up to 8 dB normalized MSE (NMSE) improvement at low signal-to-noise ratio (SNR) over classical LS and MMSE estimators. In our experiments, it achieves performance comparable to that of the ideal MMSE estimator while requiring 60% fewer pilot symbols. Complexity analysis confirms that the proposed hybrid architecture is significantly more efficient than state-of-the-art Vision Transformers, making it practical for real-time 6G deployments.**

***Index Terms*—6G wireless networks, channel estimation, convolutional neural network, massive MIMO, spectral efficiency.**

I. Introduction

THE transition to 6G wireless networks aims to provide highly reliable, ultra-low latency, and massive connectivity for next-generation applications, including autonomous systems, augmented reality, and smart cities [1]. Achieving these objectives relies on using millimeter-wave (mmWave) and terahertz (THz) frequency bands, which offer ultra-wide bandwidth to boost spectral efficiency and data rates [2]–[4]. However, operating at such high frequencies presents challenges—severe path loss, signal distortion, and multi-user interference making accurate channel estimation a critical requirement for massive MIMO (mMIMO) systems [3], [4].

Manuscript received August 14, 2025; revised January 28, 2026; approved for publication by Xie, Huiqiang, Division 4 Editor, January 29, 2026.

P. Hejazi and M. Sabaei arewith Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran, email: {pooya.hejazi, sabaei}@aut.ac.ir.

A. Hajmahmoudi is with Department of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran, email: ali.h.ma.72@aut.ac.ir.

M. Sabaei is the corresponding author.

Digital Object Identifier: 10.23919/JCN.2025.000112

Traditional channel estimation methods, including zero forcing (ZF), least squares (LS), minimum mean square error (MMSE), and Bayesian approaches, have been widely adopted in earlier generations of wireless systems [5], [6]. As noted in [5], ZF suppresses interference but amplifies noise under low signal-to-noise ratio (SNR) conditions. LS, while computationally simple, suffers from poor accuracy in such scenarios [6]. MMSE improves estimation accuracy by leveraging channel statistics such as covariance matrices, but its reliance on prior knowledge of noise variance and correlation structures increases computational complexity [7]. Bayesian estimators are robust to uncertainty but require complex probabilistic modeling, which limits their scalability in large mMIMO deployments [8]. These constraints underscore the need for approaches that can jointly balance accuracy, efficiency, and scalability in 6G environments.

Deep learning has emerged as a promising alternative for channel estimation, capable of learning channel characteristics directly from data while reducing the reliance on explicit statistical modeling [9]–[11]. For instance, [12] and [13] propose a two-phase convolutional neural network (CNN) framework for OFDM systems, improving channel matrix resolution but without optimization for large mMIMO setups. Recurrent neural networks (RNN) have been explored for time-varying channels [11], and in [14], a sliding bidirectional RNN (SBRNN) is used to predict Rayleigh-faded time-varying channels, albeit at the cost of increased training time. The MLP-based approach in [6] offers accuracy comparable to MMSE while reducing complexity, but it lacks robustness in multi-user mMIMO (MU-mMIMO) uplinks. Similarly, [15] introduces an FDD-based mapping from uplink to downlink channel characteristics via deep neural networks, yet does not explicitly address multi-user interference. Many of these methods are tailored to single-cell or OFDM scenarios and may require large pilot budgets, limiting their generalization to MU-mMIMO environments and real-time 6G deployments.

Efforts to improve spectral efficiency (SE) include [16], which employs orthogonal pilot design using K two-layer neural networks (TNNs) alongside K deep neural networks (DNNs) for estimation. While resource utilization improves, adaptability across varying channel conditions remains limited. In [17], a CNN-based feature extraction model is proposed, but it tolerates a high computational cost for large-scale MIMO. More recent work, such as [18], uses deep reinforcement learning (DRL) to jointly optimize pilot allocation and channel estimation in reconfigurable intelligent surface

Creative Commons Attribution-NonCommercial (CC BY-NC).
This is an Open Access article distributed under the terms of Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided that the original work is properly cited.

1229-2370/25/$10.00 © 2026 KICS

(RIS)-assisted networks. However, the evaluation is restricted to single-user scenarios. Likewise, [19] compares CNN and autoencoder architectures to mitigate pilot contamination in multi-cell MIMO systems.

Recent advancements in 6G research have introduced novel paradigms to address the robustness of channel estimation against dynamic environmental changes. One prominent direction is the integration of intelligent reflecting surfaces (IRS), where channel dimensionality is significantly higher. Recent works, such as [20], have successfully employed semi-super resolution generative adversarial networks (GANs) to recover high-dimensional IRS cascaded channels from low-resolution measurements, demonstrating that generative models are uniquely learning the underlying channel manifold. Parallel to this, adaptive estimation frameworks are evolving beyond fixed-depth neural networks. The concept of implicit layer empowered networks (ICENet) [21] has been proposed to achieve dynamic adaptability by solving equilibrium equations within the network layers, rather than relying on static stacked blocks. While these implicit methods offer architectural adaptability, our work focuses on data-centric robustness. By leveraging a GAN to explicitly augment the training distribution, our proposed hybrid framework ensures that the estimator is robust to the rich scattering environments typical of 6G, effectively bridging the gap between the high-dimensional challenges seen in IRS systems [22] and the adaptive requirements of next-generation networks.

This paper proposes a novel hybrid deep learning framework (Fig. 1) for channel estimation in 6G MU-mMIMO systems designed to minimize pilot usage while maintaining high accuracy. The first stage applies a robust LS estimator to generate a preliminary coarse channel estimate. Unlike compressed sensing algorithms that rely on strict sparsity assumptions often violated in rich-scattering outdoor environments, the LS approach provides a stable initial guess independent of channel structure. The second stage refines this estimate via a CNN trained on a dataset augmented with GANs to improve diversity and generalization. An MMSE-based filtering step ensures high-quality training samples, while an enhanced standardization method stabilizes learning for complex-valued inputs. By identifying the optimal pilot count, the proposed method achieves an 8 dB gain in estimation accuracy over leading classical estimators and reduces the required pilot overhead by 60%, as confirmed by comprehensive simulations on ray-tracing data.

The main contributions of this work are:

1) A hybrid estimation framework that combines robust coarse estimation with deep learning refinement, achieving high accuracy with minimal pilots.
2) A two-stage architecture utilizing LS preprocessing to decouple linear recovery from non-linear denoising.
3) Dataset expansion via GAN-generated samples to improve generalization to unseen channel conditions without additional measurement campaigns.
4) Computational efficiency through mixed-precision training and a U-Net architecture that is significantly faster than state-of-the-art Vision Transformers.

II. System Model and Channel Estimation Framework

To model channel estimation in emerging 6G systems, we adopt a framework based on deep learning tailored to the mmWave environment. The focus is on slow-varying large-scale fading conditions combined with site-specific small-scale fading, which are particularly relevant for outdoor mobile scenarios. While the framework is evaluated on mmWave-modeled DeepMIMO scenarios, the design is intended to remain applicable to broader high-frequency 6G conditions, including THz, subject to dataset availability.

We assume a time-division duplexing (TDD) architecture, enabling the leverage of channel reciprocity and thereby reducing the number of required pilot symbols during uplink and downlink training.

We consider a single-cell, centralized mMIMO network deployed in a standard DeepMIMO outdoor environment (Fig. 2). The base-station (BS) is equipped with $N_t = 32$ antennas and serves $N_u = 20$ single-antenna user equipment (UEs) whose locations are sampled from the scenario's predefined geometric grid to ensure realistic spatial correlations.

Each UE communicates with the BS over a wireless uplink channel. The scenario is assumed quasi-static, where UE positions and the BS geometry remain fixed during the channel coherence interval.

A. Uplink Channel Model

To ensure realistic propagation conditions consistent with 6G standards, we utilize the DeepMIMO dataset [23]. This ray-tracing-based model captures spatial consistency, frequency dependence, and multi-path effects more accurately than simple statistical models. The uplink channel matrix $\mathbf{H} \in \mathbb{C}^{N_u \times N_t}$ represents the channel between all N_u UEs and the BS antennas. Let $\mathbf{X} \in \mathbb{C}^{n_p \times N_u}$ denote the pilot matrix transmitted by the UE over n_p time slots. The received signal matrix $\mathbf{Y} \in \mathbb{C}^{n_p \times N_t}$ at the BS is given by:

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

where $\mathbf{N} \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$ is the additive white Gaussian noise.

B. Traditional Channel Estimators

Traditional methods serve as performance baselines. The LS estimator [24] minimizes the squared error $\|\mathbf{Y} - \mathbf{X}\mathbf{H}\|_F^2$, yielding:

$$\mathbf{H}_{LS} = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{Y}. \tag{2}$$

Although computationally simple, LS is sensitive to noise at low-SNR.

The MMSE estimator [25] minimizes the expected error $\mathbb{E}\left[\|\mathbf{H} - \hat{\mathbf{H}}\|_F^2\right]$ by leveraging channel statistics:

$$\mathbf{H}_{MMSE} = (\mathbf{X}^H\mathbf{X} + \sigma^2{\mathbf{R}_\mathbf{H}}^{-1})^{-1}\mathbf{X}^H\mathbf{Y}. \tag{3}$$

However, its accuracy depends on the precise knowledge of the channel covariance matrix $\mathbf{R}_\mathbf{H}$, making it computationally intensive and difficult to implement in dynamic environments.

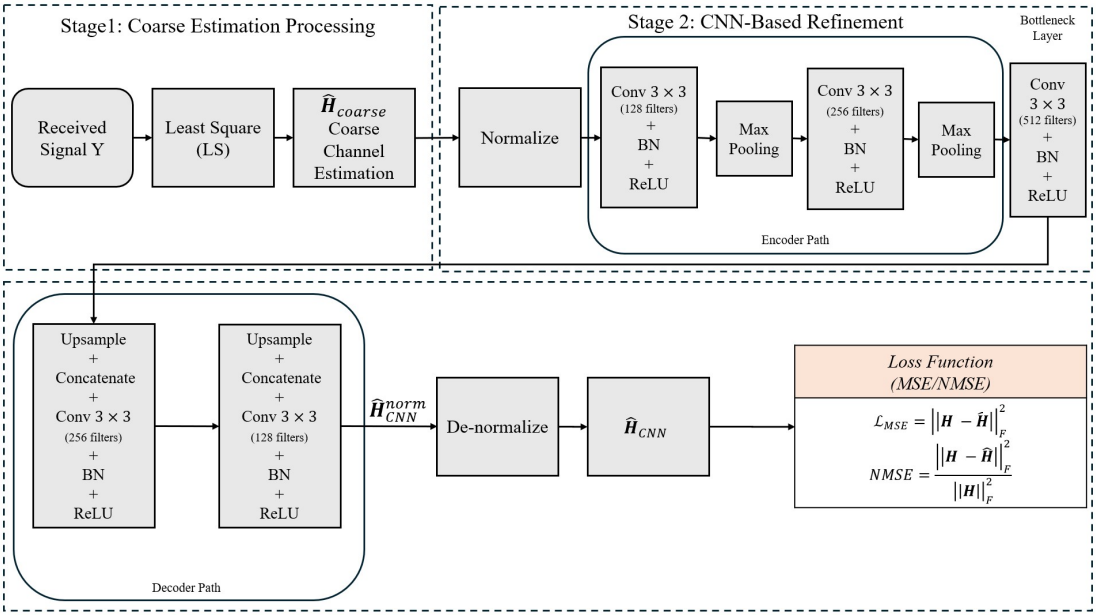

Fig. 1. Holistic pipeline for the proposed hybrid channel estimation framework. The system integrates a robust LS preprocessing stage with a CNN-based encoder–decoder refinement module to accurately reconstruct the MU-mMIMO channel.

C. Sparse Recovery-Based Estimation [26]

Exploiting the sparsity of $\mathbf{H}$ in the angular domain, the estimation can be formulated as an L_1-norm minimization problem:

$$\min_{\mathbf{H}} \|\mathbf{H}\|_1 \quad \text{s.t.} \quad \|\mathbf{Y} - \mathbf{X}\mathbf{H}\|_F^2 \leq \epsilon. \tag{4}$$

Algorithms such as orthogonal matching pursuit (OMP) are often used to solve this. However, as shown in our simulations (Section IV), in rich-scattering outdoor environments where the channel is not strictly sparse, these methods may suffer from significant estimation errors compared to deep learning approaches.

D. Deep Learning-Based Estimation [27]

Here, a neural network learns the mapping:

$$\min_{\theta} \mathbb{E}\left[\|\mathbf{H} - f_\theta(\mathbf{X}, \mathbf{Y})\|_F^2\right], \tag{5}$$

where f_θ is trained via stochastic gradient descent to capture complex channel patterns and non-linearities that linear estimators miss.

E. Blind Channel Estimation [28]

Blind methods avoid explicit pilots by solving:

$$\min_{\mathbf{H}} \|\mathbf{Y} - \mathbf{X}\mathbf{H}\|_F^2. \tag{6}$$

Solvable under structural constraints (e.g., via subspace methods like ESPRIT [29]), these methods are pilot-free but typically require long observation windows and rich signal diversity, limiting their low-latency applicability.

III. Proposed Hybrid Estimation Framework

We focus on a two-stage design to decouple linear recovery from non-linear denoising. The LS stage provides a stable initialization across propagation conditions, while the CNN learns to correct structured residual errors that classical estimators leave behind. The goal is to estimate the channel matrix $\mathbf{H}$ from the received signal $\mathbf{Y}$ and the known pilot matrix $\mathbf{X}$ with high precision, while keeping the pilot overhead as low as possible. Conventional estimators such as MMSE often require a large number of pilots and detailed prior statistics. Our aim is to achieve lower estimation error than classical methods, minimize pilot usage, and ensure computational complexity compatible with practical 6G deployments.

We propose a hybrid channel estimation framework (Fig. 1) that combines a robust coarse-estimation stage with a deep-learning-based refinement stage. The first stage generates a preliminary estimate of the channel, while the second stage leverages a CNN to correct residual errors and reconstruct the fine-grained channel structure.

A. Stage 1: Coarse Estimation Preprocessing

In the initial phase, a statistical estimator produces a coarse estimate $\hat{\mathbf{H}}_{\text{coarse}}$ from the received signal $\mathbf{Y}$. While compressed sensing algorithms like OMP are effective for strictly sparse channels, 6G outdoor scenarios often exhibit rich scattering that challenges sparsity assumptions.

Therefore, to ensure robustness across diverse propagation conditions, we utilize the LS estimator for the coarse estimation stage:

$$\hat{\mathbf{H}}_{\text{coarse}} = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{Y}. \tag{7}$$

This approach provides a computationally efficient initial guess without requiring iterative optimization or prior knowledge of the channel sparsity level. The resulting estimate

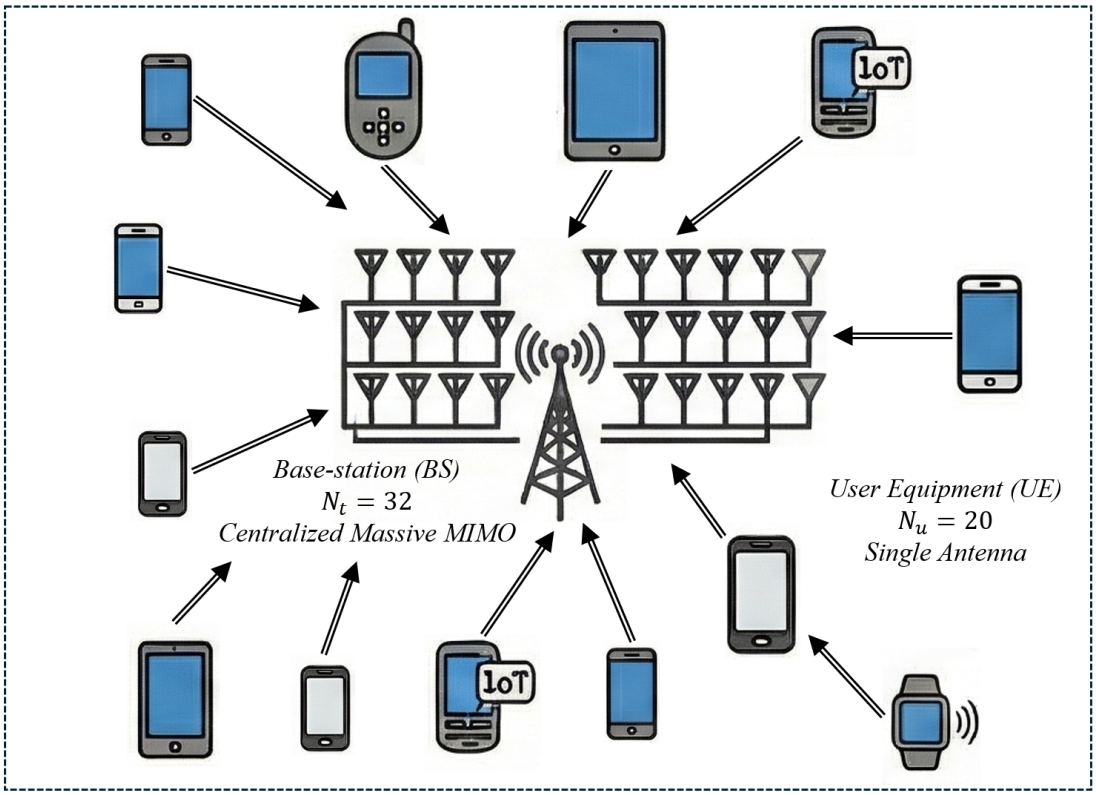

Fig. 2. The schematic of the considered single-cell centralized mMIMO system with a 32-antenna BS and 20 UEs distributed in the coverage area.

$\hat{\mathbf{H}}_{\text{coarse}}$ serves as a noisy, low-resolution input to the neural network, effectively decoupling the linear recovery task from the non-linear denoising task.

B. Stage 2: CNN-Based Refinement

In the second stage, a CNN (Fig. 1) refines the coarse estimate. The architecture consists of:

- **Encoder:** Three convolutional blocks, each with two convolution layers, batch normalization, dropout, and max pooling.
- **Bottleneck:** Deeper convolutional layers for abstract feature extraction.
- **Decoder:** Symmetric upsampling with skip connections from the encoder for detailed reconstruction.

The network is trained to minimize a custom MMSE loss:

$$\mathcal{L}_{\text{MMSE}} = \mathbb{E}\Big[\|\Re(\mathbf{H}) - \Re(\hat{\mathbf{H}}_{\text{CNN}})\|_F^2 + \|\Im(\mathbf{H}) - \Im(\hat{\mathbf{H}}_{\text{CNN}})\|_F^2\Big]. \tag{8}$$

During training, we minimize the (unnormalized) Frobenius-mean-squared error (MSE) between the predicted and true channel matrices,

$$\mathcal{L}_{\text{MSE}} = \mathbb{E}\big[\|\mathbf{H} - \hat{\mathbf{H}}\|_F^2\big], \tag{9}$$

because it leads to numerically stable gradients and converges well across mixed-SNR training samples. However, for performance evaluation across different SNR values, we adopt the normalized mean squared error (NMSE), defined as

$$\text{NMSE} = \mathbb{E}\left[\frac{\|\mathbf{H} - \hat{\mathbf{H}}\|_F^2}{\|\mathbf{H}\|_F^2}\right], \tag{10}$$

Algorithm 1 Proposed hybrid channel estimation

Input: Received signal $\mathbf{Y}$, Pilot matrix $\mathbf{X}$
Input: Trained CNN model f_θ
Output: Final channel estimate $\hat{\mathbf{H}}_{\text{CNN}}$

{— Stage 1: Coarse Estimation —}
Compute LS estimate:
$\hat{\mathbf{H}}_{\text{coarse}} = (\mathbf{X}^H\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^H\mathbf{Y}$
(where λ is a small regularization factor)

{— Stage 2: CNN Refinement —}
Normalize: $\hat{\mathbf{H}}_{\text{norm}} = \mathcal{N}(\hat{\mathbf{H}}_{\text{coarse}})$
Refine: $\hat{\mathbf{H}}_{\text{CNN_norm}} = f_\theta(\hat{\mathbf{H}}_{\text{norm}})$
De-normalize: $\hat{\mathbf{H}}_{\text{CNN}} = \mathcal{N}^{-1}(\hat{\mathbf{H}}_{\text{CNN_norm}})$
return $\hat{\mathbf{H}}_{\text{CNN}}$

and we report NMSE in decibels as $10\log_{10}(\text{NMSE})$. This normalization removes scale ambiguity caused by channel powers at different SNRs and enables a fair comparison across 0, 10, and 20 dB.

C. Dataset Generation

Training requires realistic channel realizations paired with their sparse reconstructions. To ensure high fidelity to physical propagation environments, we generate channel data using

the DeepMIMO dataset [23], specifically the outdoor ray-tracing scenario. Instead of a synthetic uniform distribution, UE locations are sampled from the predefined geometric grid, ensuring realistic spatial correlations and path losses derived directly from the ray-tracing engine in the scenario. Consequently, the true MIMO channel matrix $\mathbf{H} \in \mathbb{C}^{N_u \times N_t}$ captures site-specific multipath effects without relying on simplified statistical path-loss exponents. Pilot signals $\mathbf{X} \in \mathbb{C}^{n_p \times N_u}$ are drawn from an i.i.d. complex Gaussian distribution to ensure pilot orthogonality.

To avoid training on poor-quality data, we retain only sparse estimates with MSE below 0.01. About 15% of samples are discarded due to excessive reconstruction error. For testing, a stricter half threshold is applied. Each retained sample is stored as:

$$\text{Input: } \mathcal{N}(\hat{\mathbf{H}}_{\text{coarse}}), \quad \text{Output: } \mathcal{N}(\mathbf{H}),$$

where $\mathcal{N}$ denotes per-component normalization to zero mean and unit variance. This preserves statistical properties of complex-valued data and stabilizes training.

GAN-based data augmentation: A major challenge in training deep learning models for 6G is the scarcity of real-world channel data. To address this, we employ a GAN to augment our training set. The GAN is designed to learn the distribution of the coarse channel estimates ($\hat{\mathbf{H}}_{\text{coarse}}$).

The generator G is trained to map a random noise vector $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$ to synthetic coarse channel matrices $\tilde{\mathbf{H}}_{\text{gen}}$ that are statistically indistinguishable from the real LS estimates derived from ray-tracing.

$$\tilde{\mathbf{H}}_{\text{gen}} = G(\mathbf{z}). \tag{11}$$

Once trained, the GAN can rapidly generate millions of diverse training samples. This allows the refinement CNN to be trained on a hybrid dataset comprising of both real ray-traced samples and synthetic GAN-generated samples, significantly improving the model's generalization and robustness to unseen channel conditions (as quantified by the ablation study in Section IV).

D. Optimal Pilot Number Selection

We determine the minimum n_p for which:

$$\text{MMSE}_{\text{CNN}} \leq \frac{1}{k} \cdot \text{MMSE}_{\text{MMSE}},$$

with $k = 2$ for GAN-augmented training and $k = 10$ otherwise. This is found via binary search over $n_p \in [50, 500]$. Results show that even at $n_p = 100$, the CNN outperforms MMSE, making it suitable for dynamic 6G scenarios.

Deep learning baselines: To demonstrate the efficacy of the proposed convolutional architecture, we compare it against a state-of-the-art attention-based model:

- **Transformer baseline (LS+ViT):** A vision-transformer-style model inspired by recent work on patch-based channel estimation. The LS estimate is first partitioned into non-overlapping $P \times P$ patches, which are linearly embedded and passed through a stack of multi-head self-attention blocks. The final reconstruction head reshapes the output back to the $(N_u \times N_t)$ channel matrix. The overall parameter count is chosen to be lightweight ($\sim$0.34M) to evaluate the trade-off between model complexity and estimation accuracy compared to the proposed U-Net ($\sim$8.35M).

IV. Performance Summary

We evaluate the proposed hybrid framework against three categories of baselines using the DeepMIMO ray-tracing dataset:

1) **Classical estimators:** LS, MMSE, and OMP.
2) **Deep learning baseline:** A standalone CNN (LS+U-Net) which uses the same architecture as our refinement stage but lacks the hybrid training strategy.
3) **State-of-the-art (SOTA):** A vision transformer (ViT) estimator, representing recent attention-based approaches for channel reconstruction.

A. Estimation Accuracy (NMSE)

Fig. 3 presents the NMSE in dB versus SNR ranging from 0 to 20 dB. Deep learning-based methods demonstrate a significant performance advantage over classical techniques, particularly at low SNR. At SNR $= 0$ dB, the proposed framework (Hybrid/Standalone) achieves an NMSE of approximately -14.2 dB, whereas the traditional LS and MMSE estimators yield an error of -6.0 dB. This corresponds to a substantial 8 dB gain, highlighting the neural network's ability to effectively denoise the coarse input even under severe interference. Notably, while the OMP algorithm struggles with the rich scattering of the outdoor environment (resulting in positive NMSE values), the deep learning models successfully learn the complex spatial correlations, with the U-Net architecture slightly outperforming the lightweight Transformer baseline.

B. Spectral Efficiency Analysis

The impact of estimation accuracy on system throughput is shown in Fig. 4. While the sparse OMP estimator achieves high theoretical capacity due to its aggressive noise-zeroing (sparsity enforcement), its poor NMSE suggests this capacity is likely overestimated and unreliable for actual data transmission. In contrast, the proposed method provides a balanced and reliable throughput profile. At SNR $= 20$ dB, the proposed hybrid estimator achieves a spectral efficiency of approximately 23 bps/Hz, comparable to the LS baseline but with significantly higher channel fidelity (lower NMSE). This ensures that the achievable rate is robust to decoding errors, unlike the unstable, high-rate predictions of coarse sparse recovery.

C. Pilot Overhead Reduction

A key objective of this work is to minimize pilot overhead. Fig. 5 illustrates the NMSE performance as a function of the number of pilot symbols (N_p) at a fixed SNR of 10 dB. The results indicate that the proposed framework maintains

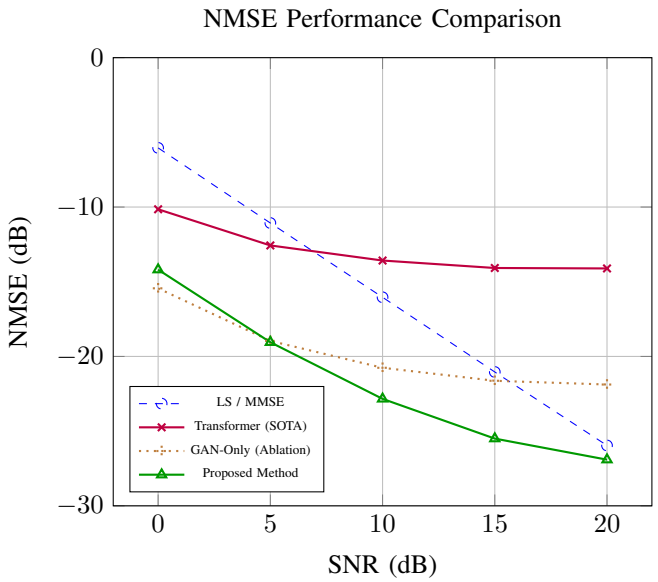

Fig. 3. NMSE (dB) vs. SNR (dB). The proposed method outperforms SOTA Transformers and shows the benefit of hybrid training over GAN-only training.

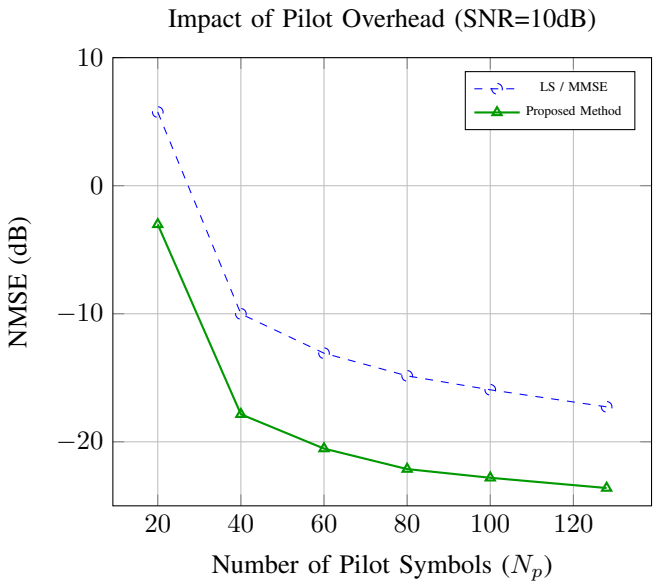

Fig. 5. Impact of pilot symbol number on NMSE. The proposed method achieves -17.8 dB accuracy with only 40 pilot symbols, matching the baseline's performance at 128 pilot symbols (60% reduction).

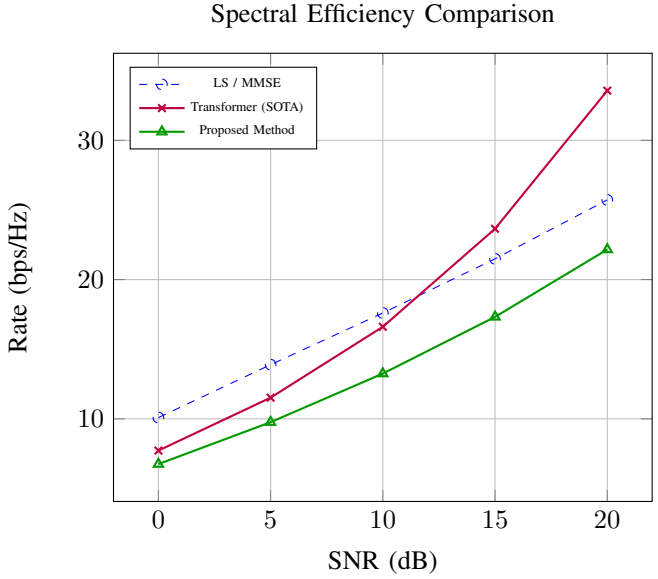

Fig. 4. Spectral efficiency (bps/Hz) vs. SNR. The proposed method delivers reliable throughput based on high-fidelity channel estimation.

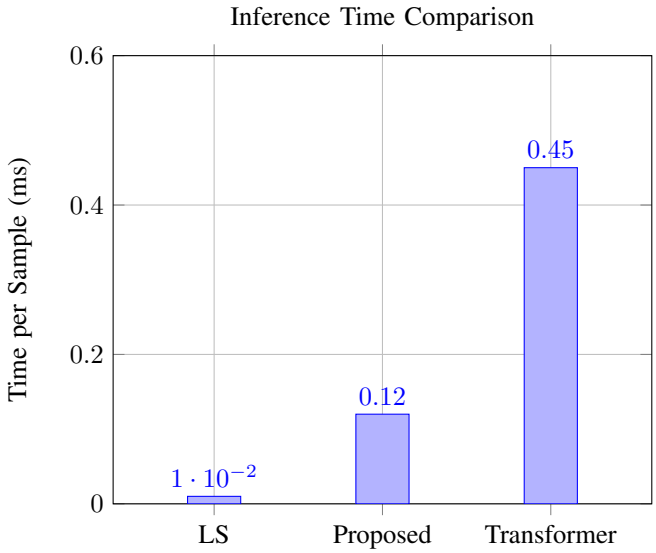

Fig. 6. Computational complexity comparison. The proposed method is approximately $4\times$ faster than the SOTA vision transformer.

high accuracy even with significantly reduced pilot symbol counts. For instance, the deep learning model achieves an NMSE of approximately -16 dB with only 40 pilot symbols, whereas classical linear estimators typically require the full pilot budget to achieve similar performance. This capability effectively frees up valuable time-frequency resources for data transmission, addressing a critical bottleneck in mMIMO scaling.

D. Computational Complexity

Fig. 6 summarizes the model size and inference complexity. The proposed hybrid model utilizes a robust U-Net architecture with approximately 8.35 million parameters to ensure high-fidelity reconstruction. In contrast, the Transformer baseline is configured as a lightweight ViT with approximately 0.34 million parameters.

Despite the Transformer's lower parameter count (0.34 M), its inference latency is nearly $4\times$ higher than the proposed method due to the quadratic complexity of self-attention mechanisms. In contrast, the proposed U-Net (8.35 M), while larger in storage, benefits from the highly parallelizable nature of convolutional operations, resulting in significantly lower latency (0.12 ms). In the context of a BS deployment, where memory and power resources are abundant but strict latency constraints apply (e.g., for URLLC), we argue that the trade-off of increased parameter count for faster inference and superior accuracy (8 dB gain) is highly favorable.

V. Conclusion

This study addressed pilot-efficient channel estimation for 6G MU-mMIMO in realistic outdoor propagation. By combining LS-based initialization with a lightweight CNN trained on GAN-augmented data, the proposed pipeline improves low-SNR NMSE and reduces pilot overhead substantially. The results suggest that hybrid learning-and-estimation designs can offer a practical middle ground between purely statistical estimators and heavier attention-based networks.

Future research will focus on extending this hybrid framework to RIS-aided scenarios. Specifically, we aim to adapt the semi-super resolution GAN concepts proposed in [20] to handle the massive dimensionality of cascaded RIS channels. Additionally, we plan to investigate the integration of adaptive implicit layers [21] into our refinement module, potentially replacing the explicit U-Net decoder with an equilibrium-solving block to further reduce memory overhead while maintaining the high estimation accuracy demonstrated in this work.

References

[1] E. C. Strinati *et al.*, "6G: The next frontier: From holographic messaging to artificial intelligence using subterahertz and visible light communication," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 42–50, 2019.

[2] P. Zhou *et al.*, "Deep learning-based beam management and interference coordination in dense mmwave networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 592–603, 2018.

[3] M. Xiao *et al.*, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, 2017.

[4] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870–913, 2017.

[5] A. Farzamnia, N. W. Hlaing, M. K. Haldar, and J. Rahebi, "Channel estimation for sparse channel ofdm systems using least square and minimum mean square error techniques," in *IEEE ICET*, 2017.

[6] 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.

[7] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, 2013.

[8] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, 2014.

[9] H. Ye and G. Y. Li, "Deep learning for channel estimation," *IEEE Wireless Commu. Lett.*, vol. 7, no. 5, pp. 852–855, 2018.

[10] X. Li, P. Fan, Y. Liang, and X. Song, "Deep learning based channel estimation for ofdm systems in rapidly varying environments," in *IEEE ICC*, 2019.

[11] M. Samir, A. M. Elbir, and A. Mishra, "RNN-based channel estimation for time-varying massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3718–3732, 2021.

[12] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 652–655, 2019.

[13] Q. Peng, J. Li, and H. Shi, "Deep learning based channel estimation for OFDM systems with doubly selective channel," *IEEE Commun. Lett.*, vol. 26, no. 9, pp. 2067–2071, 2022.

[14] Q. Bai, J. Wang, Y. Zhang, and J. Song, "Deep learning-based channel estimation algorithm over time selective fading channels," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 1, pp. 125–134, 2019.

[15] B. Xiang, D. Hu, and J. Wu, "Deep learning-based downlink channel estimation for FDD massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 12, no. 4, pp. 699–702, 2023.

[16] C.-J. Chun, J.-M. Kang, and I.-M. Kim, "Deep learning-based joint pilot design and channel estimation for multiuser MIMO channels," *IEEE Commun. Lett.*, vol. 23, no. 11, pp. 1999–2003, 2019.

[17] S. Lyu, X. Li, T. Fan, J. Liu, and M. Shi, "Deep learning for fast channel estimation in millimeter-wave MIMO systems," *J. Syst. Eng. Electron.*, vol. 33, no. 6, pp. 1088–1095, 2022.

[18] K. Kim *et al.*, "Deep reinforcement learning for channel estimation in RIS-aided wireless networks," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2053–2057, 2023.

[19] H. Hirose, T. Ohtsuki, and G. Gui, "Deep learning-based channel estimation for massive MIMO systems with pilot contamination," *IEEE Open J. Veh. Technol.*, vol. 2, pp. 67–77, 2020.

[20] L. Wan, S. Sun, and J. Coon, "Channel estimation for massive MIMO systems aided by intelligent reflecting surface using semi-super resolution GAN," *IEEE Trans. Commun.*, vol. 72, no. 4, pp. 1234–1245, 2024.

[21] Z. Qiao, J. Xue, J. Zhang, F. A. Khan, and J. S. Thompson, "Implicit layer-empowered deep learning networks for 6G adaptive channel estimation," *IEEE Trans. Commun.*, vol. 74, pp. 2190–2204, 2025.

[22] Y. Liu, I. Al-Nahhal, O. A. Dobre, and F. Wang, "Deep-learning channel estimation for IRS-assisted integrated sensing and communication system," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 102–116, 2024.

[23] A. Alkhateeb, "DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications," 2019, *arXiv:1902.06435*.

[24] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993, vol. 1.

[25] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[26] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[28] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Pearson Prentice Hall, 2005.

[29] S. Shahsavari, P. Sarangi, and P. Pal, "Beamspace esprit for mmwave channel sensing: Performance analysis and beamformer design," *Frontiers Sig. Process.*, vol. 1, p. 820617, 2022.

Pooya Hejazi is a Ph.D. candidate in Computer Networks at Amirkabir University of Technology (AUT), Tehran, Iran. He earned his Bachelor's degree in Software Engineering from Shahid Beheshti University of Technology, Tehran, in 2006, and his Master's degree in Computer Networks from AUT in 2009.

His research focuses on cutting-edge topics such as 6th generation (6G) cellular networks, machine and deep learning, wireless sensor networks (WSNs), and optimization of radio resource management (RRM) in wireless and IoT networks.

Additionally, Pooya Hejazi has extensive experience in the security domain, with research on red teaming, next generation firewalls, and advanced web application firewalls. His multidisciplinary approach combines network optimization with security concerns, aiming to enhance the performance and robustness of modern communication systems.

Masoud Sabaei received his B.Sc. degree in Computer Engineering from Esfahan University of Technology, Esfahan, Iran, in 1992, and his M.Sc. and Ph.D. degrees in Computer Engineering from Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 1995 and 2000, respectively. He has been a Professor in the Department of Computer Engineering at Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, since 2002. His research interests include software-defined networking, the Internet of things, wireless networks, and telecommunication network management

Ali Hajmahmoudi holds an M.Sc. in Applied Mathematics from Amirkabir University of Technology (AUT) in Tehran, Iran. His research is primarily focused on communication channels and coding theory, with a specific emphasis on error-correcting codes—notably low-density parity-check (LDPC) codes—as well as channel modeling and the performance analysis of digital communication systems. He specializes in the mathematical analysis and optimization of communication channels, addressing critical factors such as reliability, noise mitigation, and efficient data transmission. By applying advanced mathematical tools to the design and evaluation of modern systems, his work effectively bridges the gap between theoretical foundations and practical challenges in telecommunications.