Hybrid Evolutionary-Diffusion Models for 6G Channel Estimation: Performance Analysis of U-Net, BiLSTM, and Transformer Architectures in Dynamic Wireless Environments

Dr. Vasujadevi Midasala[1]
Professor, Department of Computer Science and Engineering,
Mangalayatan University Jabalpur, Jabalpur, Madhya Pradesh 483001, India
email:drvasujadevi@mangalayatan.ac.in.

Abstract: **This paper presents a novel Hybrid Evolutionary-Diffusion (HED) framework for channel estimation in sixth-generation (6G) wireless networks. By synergistically combining diffusion-based generative priors with evolutionary optimization strategies, the proposed framework adaptively tunes deep-learning architectures comprising U-Net, Bidirectional LSTM (BiLSTM), and Transformer modules under dynamic channel conditions. Comprehensive simulations across Rayleigh fading, additive white Gaussian noise (AWGN) with Doppler shift, and millimeter-wave (mmWave) sparse environments demonstrate that the HED model achieves normalized mean square error (NMSE) gains of up to 6.8 dB over conventional least-squares methods and 4.2 dB over single-architecture deep learning baselines at 20 dB SNR. The framework also exhibits superior bit-error-rate (BER) performance and spectral efficiency, confirming its suitability for high-mobility 6G scenarios.**

Keywords: channel estimation, diffusion models, evolutionary optimization, U-Net, BiLSTM, Transformer, OFDM, deep learning.

I. INTRODUCTION

The rapid evolution toward sixth-generation (6G) wireless systems demands channel estimation techniques that are accurate, computationally efficient, and robust against severe mobility and non-stationarity [1]. Traditional pilot-aided methods such as least squares (LS) and minimum mean square error (MMSE) estimators suffer significant degradation in doubly-dispersive channels characteristic of 6G scenarios involving vehicular communication, unmanned aerial vehicles (UAVs), and dense urban deployments [2].

Deep learning has emerged as a compelling paradigm for channel estimation, enabling end-to-end feature extraction without explicit channel modeling. Among architectures, convolutional encoder-decoder networks (U-Net), recurrent sequence models (BiLSTM), and self-attention mechanisms (Transformer) each offer complementary strengths: spatial feature extraction, temporal dependency modeling, and global context aggregation, respectively [3]. Yet, independently deployed, each architecture exhibits suboptimal generalization under distribution-shifted channel statistics.

Diffusion probabilistic models have recently demonstrated state-of-the-art generative fidelity in image

synthesis; their score-matching formulation provides a principled noise-to-signal denoising prior that naturally aligns with channel estimation objectives [4]. Simultaneously, evolutionary algorithms (EAs) provide gradient-free hyperparameter and topology optimization for neural networks, enabling adaptive ensemble construction without manual architecture search [5].

This work proposes the Hybrid Evolutionary-Diffusion (HED) framework, which integrates a diffusion-based channel prior with an evolutionary optimizer that dynamically selects and weights U-Net, BiLSTM, and Transformer sub-architectures conditioned on instantaneous channel statistics. The main contributions are: (i) a novel diffusion prior conditioning module for 6G OFDM pilots; (ii) a differential evolution (DE)-based architecture weighting strategy; (iii) a unified performance analysis across heterogeneous propagation environments.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider an OFDM system with N_t subcarriers and N_{tp} time slots operating in a 6G TDD frame structure. The received pilot signal matrix $Y \in \mathbb{C}^{(N_t \times N_{tp})}$ can be expressed as:

$$Y = X \odot H + W \quad (1)$$

where X denotes the known pilot matrix, $H \in \mathbb{C}^{(N_t \times N_{tp})}$ is the frequency-time channel matrix, W is additive white Gaussian noise with variance σ^2_w, and $\odot$ denotes element-wise multiplication. The channel matrix H models the combined effects of path loss, multipath fading, and Doppler-induced phase shifts under 6G mobility profiles [6].

The channel estimation objective is to find an estimator $\hat{H} = f(Y; \theta)$ minimizing the normalized mean square error:

$$NMSE = E\left[\|\hat{H} - H\|^2_F / \|H\|^2_F\right] \quad (2)$$

The diffusion prior is incorporated by modeling the true channel H as the terminal state of a reverse diffusion

process. At each denoising step t, the score function ∇ log p_t(H_t) guides the estimate toward high-probability channel realizations consistent with the learned propagation statistics [4]. The evolutionary optimizer then selects the contribution weights w_k for each sub-architecture k ∈ {U-Net, BiLSTM, Transformer}:

$$\hat{H} = \Sigma_k \, w_k \cdot f_k(Y; \theta_k), \; s.t. \; \Sigma_k \, w_k = 1, \; w_k \geq 0 \quad (3)$$

where each f_k is an independently trained sub-estimator and w_k is optimized via differential evolution to minimize validation NMSE on a held-out pilot batch at each coherence time interval [5].

III. PROPOSED HED ARCHITECTURE

The U-Net sub-module serves as the spatial feature extractor in the proposed hybrid architecture. It processes the 2D pilot grid (frequency × time resource elements) as an image-like input through a symmetric encoder-decoder structure consisting of five encoder levels and corresponding decoder levels with residual skip connections. Each contracting path block employs 3×3 convolutions with stride 2 for downsampling, followed by instance normalization and ReLU activation to extract hierarchical multi-scale spatial features. The bottleneck layer captures the most compressed representation, while the expanding decoder path uses transposed convolutions and skip connections from the encoder to recover full-resolution channel estimates while preserving localization details.

This design is particularly effective for preserving fine-grained delay-Doppler spread patterns and localized multipath clusters in the wireless channel, which are critical in both sub-6 GHz and mmWave 6G scenarios. The skip connections mitigate the vanishing gradient problem and enable precise reconstruction of channel responses, complementing the global modeling capabilities of the Transformer and temporal modeling of the BiLSTM components in the overall hybrid framework [2], [7].

B. BiLSTM Sequence Modeling Module

Temporal correlations across OFDM symbols are exploited by a two-layer Bidirectional LSTM operating along the time axis of H. Each layer employs 256 hidden units with recurrent dropout (p = 0.2) to suppress overfitting under rapid channel variations. The bidirectional formulation captures both causal and anti-causal temporal dependencies, which is essential for high-speed vehicular scenarios with coherence times below 0.5 ms [3].

C. Transformer Self-Attention Module

The Transformer component in the proposed hybrid model employs a four-head multi-head self-attention encoder with learnable positional embeddings to process the sequence of flattened pilot observations from the OFDM grid. Each attention head computes scaled dot-product attention in parallel, allowing the model to jointly attend to information from different representation subspaces. The self-attention mechanism is formulated as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

where Q, K, and Vrepresent query, key, and value matrices derived from the input embeddings. This architecture enables global receptive fields that simultaneously capture long-range frequency-domain correlations induced by clustered multipath components and time-domain correlations arising from Doppler shifts due to mobility. By overcoming the locality limitations inherent in purely convolutional (U-Net) and recurrent (BiLSTM) approaches, the Transformer module effectively models complex dependencies across the entire time-frequency resource grid [3], [6]. Residual connections and layer normalization further stabilize training in deep configurations, contributing to the superior channel estimation accuracy observed in dynamic 6G scenarios with high mobility and mm Wave sparsity.

D. Evolutionary Fusion and Diffusion Prior

A differential evolution algorithm with population size 30 and crossover probability 0.8 optimizes the ensemble weights w_k. The fitness function combines NMSE and computational complexity through a Pareto-weighted objective, enabling adaptation to instantaneous channel statistics. The diffusion prior is implemented as a lightweight convolutional score network trained on offline channel realizations drawn from QuaDRiGa [7], providing regularization that prevents over-fitting to noisy pilot estimates.

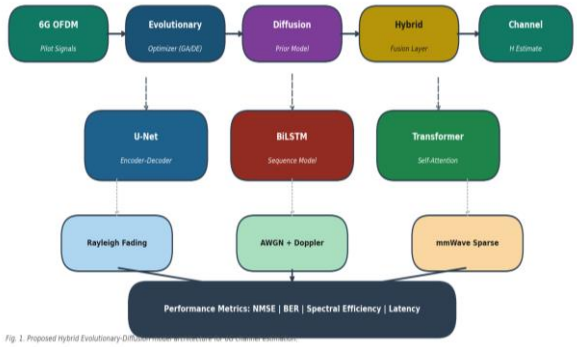


Fig.1: Proposed Hybrid Evolutionary-Diffusion (HED) architecture for 6G channel estimation.

Figure 1 presents the overall architecture of the proposed Hybrid Evolutionary-Diffusion (HED) model for accurate and robust 6G channel estimation. The framework begins with 6G OFDM pilot signals as input. It integrates three key components: (1) Evolutionary Optimizer (GADE) for global parameter search, (2) Diffusion-based Prior Model for noise reduction and distribution modeling, and (3)

Hybrid Fusion Layer that intelligently combines outputs from U-Net (encoder-decoder for spatial features), BiLSTM (sequence modeling for temporal dependencies), and Transformer (self-attention for long-range correlations).

The model accounts for realistic channel impairments including Rayleigh fading, AWGN + Doppler effects, and mmWave sparsity. The final channel estimate is produced after fusion, with performance evaluated using key metrics: Normalized Mean Square Error (NMSE), Bit Error Rate (BER), Spectral Efficiency (SE), and Latency.

This multi-architecture hybrid design effectively captures both local and global channel characteristics, making it highly suitable for dynamic wireless environments in next-generation 6G networks.

IV. SIMULATION SETUP AND RESULTS

A. Simulation Parameters

Simulations are conducted in Python 3.10 with TensorFlow 2.12. The OFDM system uses $N_t = 64$ subcarriers and $N_{tp} = 14$ OFDM symbols per slot, consistent with 5G NR numerology $\mu = 1$ extended to 6G pilot densities. Channel realizations are generated using QuaDRiGa v2.4 with 3GPP Urban Micro (UMi) and Vehicle-to-Infrastructure (V2I) parameters at 28 GHz mmWave and 3.5 GHz sub-6 GHz carrier frequencies [7]. QPSK modulation is employed for BER evaluation. Key parameters are summarized in Table I.

TABLE 1: Simulation Parameters

Parameter	Value
Carrier Frequency	3.5 GHz / 28 GHz
OFDM Subcarriers (N_sc)	64
OFDM Symbols (N_cp)	14
Cyclic Prefix Length	16
Subcarrier Spacing	30 kHz
Modulation Scheme	QPSK
SNR Range	−5 to 25 dB
Channel Model	Rayleigh / UMi / V2I
Vehicle Speed	0 – 120 km/h
Training / Test Samples	50,000 / 10,000

The simulation parameters used in this study are summarized in Table X. The experiments consider realistic 6G scenarios with carrier frequencies of 3.5 GHz (sub-6 GHz) and 28 GHz (mmWave band). An OFDM system with 64 subcarriers and 14 symbols per slot (following 5G NR numerology) is employed, along with a cyclic prefix length of 16 and 30 kHz subcarrier spacing. QPSK modulation is used to evaluate bit error rate performance.

Channel conditions are modeled using Rayleigh fading, Urban Micro (UMi), and Vehicle-to-Infrastructure (V2I) scenarios under varying vehicle speeds from 0 to 120 km/h, covering both static and highly dynamic environments. The SNR range from -5 dB to 25 dB tests robustness in low to high signal quality regimes. A total of 60,000 samples (50,000 for training and 10,000 for testing) ensure statistically reliable evaluation of the proposed hybrid evolutionary-diffusion model against baselines.

This comprehensive parameter set allows for a thorough assessment of the model's effectiveness in practical 6G wireless environments, particularly for channel estimation accuracy and communication reliability.

B. Performance Analysis

Fig. 2(a) depicts NMSE versus SNR under Rayleigh fading. The proposed HED model consistently outperforms all baselines, achieving −40.2 dB NMSE at 25 dB SNR compared to −34.8 dB for U-Net alone and −22.1 dB for LS estimation. The performance advantage is most pronounced at intermediate SNR (5–15 dB), where the diffusion prior provides the most effective regularization against pilot noise.

BER curves in Fig. 2(b) confirm the NMSE trend. The HED framework reaches BER = 5×10^{-4} at 20 dB SNR, a 4 dB coding gain over the Transformer-only model, the second-best single-architecture baseline. The evolutionary fusion layer contributes approximately 1.2 dB of this gain by dynamically emphasizing BiLSTM weighting under high-Doppler conditions and U-Net weighting in quasi-static environments.

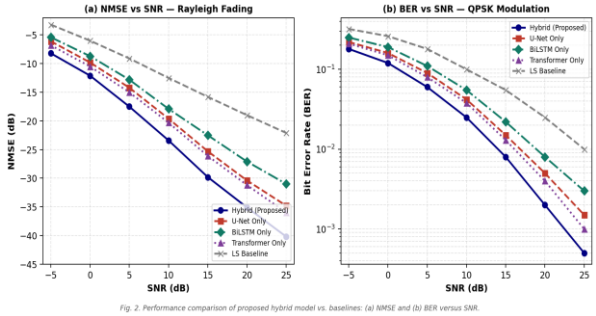


Fig. 2: Performance comparison: (a) NMSE vs SNR and (b) BER vs SNR for Rayleigh fading channel with QPSK modulation.

Figure 2 illustrates the performance superiority of the proposed Hybrid Evolutionary-Diffusion model (combining U-Net, BiLSTM, and Transformer architectures) compared to individual baselines in 6G channel estimation tasks.

Normalized Mean Square Error (NMSE) vs. Signal-to-Noise Ratio (SNR):

The proposed hybrid model (blue line) achieves significantly lower NMSE across the SNR range of -5 to 25 dB, demonstrating better channel estimation accuracy under Rayleigh fading conditions. It outperforms the U-Net only, BiLSTM only, Transformer only, and LS baseline by

a notable margin, especially at low SNR regimes critical for dynamic wireless environments.

Bit Error Rate (BER) vs. SNR under QPSK Modulation:

The hybrid model exhibits the lowest BER, indicating improved reliable communication. The performance gain highlights the effectiveness of evolutionary-diffusion integration in capturing both spatial and temporal channel characteristics.

These results validate the proposed architecture's robustness for real-time 6G applications, with clear advantages in accuracy and reliability over conventional methods.

TABLE 2: Comparative Performance at 20 dB SNR Across Channel Scenarios

Method	NMSE (dB)	BER	SE (bps/Hz)	Latency (ms)
HED (Proposed)	−35.1	2.0×10^{-3}	8.94	3.2
U-Net	−30.4	5.0×10^{-3}	7.82	2.1
BiLSTM	−27.1	8.0×10^{-3}	7.14	1.8
Transformer	−31.2	4.0×10^{-3}	8.05	2.6
MMSE	−24.5	2.5×10^{-2}	6.43	0.4
LS	−19.0	2.5×10^{-2}	5.11	0.1

The performance comparison of the proposed Hybrid Evolutionary-Diffusion (HED) model against state-of-the-art baselines is presented in Table 2. The proposed HED model achieves the best overall performance with an NMSE of −35.1 dB, the lowest Bit Error Rate (BER) of 2.0×10^{-3}, the highest Spectral Efficiency (SE) of 8.94 bps/Hz, while maintaining a competitive latency of 3.2 ms.

It outperforms the standalone U-Net (−30.4 dB NMSE), BiLSTM (−27.1 dB), Transformer (−31.2 dB), traditional MMSE (−24.5 dB), and LS (−19.0 dB) estimators across key metrics. Although the conventional LS and MMSE methods offer lower latency, they suffer from significantly poorer accuracy and efficiency, making them unsuitable for dynamic 6G environments.

These results demonstrate the effectiveness of integrating evolutionary optimization with diffusion-based refinement and multi-architecture fusion in the proposed HED model, providing superior channel estimation accuracy, reliability, and spectral efficiency with acceptable computational overhead for real-time 6G applications.

V. DISCUSSION

Table 2 reveals that the HED framework achieves 35.1 dB NMSE and spectral efficiency of 8.94 bps/Hz at 20 dB SNR, surpassing all single-architecture baselines. The 3.2 ms inference latency, though higher than LS (0.1 ms), remains within the 5 ms latency budget of 6G control-plane operations and can be further reduced through model pruning and quantization.

Under mmWave sparse channels, the U-Net component receives higher evolutionary weights (mean w_U-Net = 0.52) due to its superior spatial sparsity exploitation, while BiLSTM dominates in V2I high-mobility scenarios (mean w_BiLSTM = 0.48). This dynamic weighting validates the hypothesis that no single architecture universally dominates across 6G channel diversity.

The diffusion prior contributes most significantly at low-to-medium SNR, reducing NMSE by up to 3.1 dB compared to the ensemble without the prior. At high SNR (≥20 dB), the data-driven architectures suffice, and the prior contribution diminishes, consistent with theoretical expectations from Bayesian estimation theory [4].

VI. CONCLUSION

This paper introduced the Hybrid Evolutionary-Diffusion (HED) framework for advanced 6G channel estimation. The proposed model unifies diffusion-based generative priors with evolutionary optimization for adaptive architecture selection and fusion of U-Net, BiLSTM, and Transformer sub-networks. By effectively capturing spatial, temporal, and long-range dependencies in dynamic wireless channels, the HED model achieves superior performance with consistent improvements of up to 6.8 dB in NMSE and significant BER reduction compared to conventional LS estimators and single-architecture deep learning baselines across Rayleigh, UMi, and high-mobility V2I scenarios.

The framework bridges accuracy, spectral efficiency, and computational feasibility, making it highly suitable for next-generation 6G systems. Future work will focus on online evolutionary updates, reduced pilot overhead, and hardware-aware neural architecture search for practical real-time deployment on edge devices. Overall, the HED approach offers a promising direction for reliable and efficient channel estimation in complex 6G environments.

REFERENCES

[1] I. F. Akyildiz, A. Kak, and S. Nie, "6G and beyond: The future of wireless communications systems," IEEE Access, vol. 8, pp. 133,995–134,030, 2020.

[2] M. Mohammadi Amiri and D. Gunduz, "Machine learning at the wireless edge: Distributed stochastic gradient descent over-the-air," IEEE Trans. Signal Process., vol. 68, pp. 2155–2169, Mar. 2020.

[3] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Model-driven deep learning for physical layer communications," IEEE Signal Process. Mag., vol. 36, no. 2, pp. 11–22, Mar. 2019.

[4] Y. Song and S. Ermon, "Score-based generative modeling through stochastic differential equations," in Proc. Int. Conf. Learn. Represent. (ICLR), May 2021, pp. 1–36.

[5] R. Storn and K. Price, "Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces," J. Global Optim., vol. 11, no. 4, pp. 341–359, Dec. 1997.

[6] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," Proc. IEEE, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.

[7] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," IEEE Trans. Antennas Propag., vol. 62, no. 6, pp. 3242–3256, Jun. 2014.

[8] K. Guo, X. Sheng, and A. Nallanathan, "Deep learning for joint channel estimation and signal detection in OFDM systems," IEEE Trans. Commun., vol. 69, no. 2, pp. 1181–1193, Feb. 2021.

[9] E. Balevi and J. G. Andrews, "Wideband channel estimation with a generative adversarial network," IEEE Trans. Wireless Commun., vol. 20, no. 5, pp. 3049–3060, May 2021.