Received: 20 April 2023 | Revised: 14 September 2023 | Accepted: 8 December 2023

DOI: 10.1002/ett.4925

RESEARCH ARTICLE

WILEY

Optimized deep learning based hypernet convolution neural network and long short term memory for joint pilot design and channel estimation in MIMO-OFDM model

C. Silpa[1,2] | **A. Vani**[3] | **K. Rama Naidu**[4]

[1]Department of Electronics and Communication Engineering, JNTUA, Anantapur, India

[2]Department of Electronics and Communication Engineering, Malla Reddy Engineering College, Hyderabad, India

[3]Department of Electronics and Communication Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, India

[4]Department of Electronics and Communication Engineering, JNTUA, Anantapur, India

Correspondence
C. Silpa, JNTUA, Anantapur, Andhra Pradesh 515002, India.
Email: srsilpavasu@gmail.com

Abstract

In multiple input multiple output-orthogonal frequency division multiplexing (MIMO-OFDM) systems, efficient pilot design (PD) and channel estimation (CE) greatly influences the reliability and robustness of pilot-based CE methods. But, accurate estimationof the channel remains a challenge in a high-mobility environment with non-linear channel characteristics. Many techniques have been introduced to overcome the pilot contamination and CE problems in the MIMO-OFDM system to overcome this issue. However, these techniques take multiple paths at the receiver, resulting in delay spread and interference in the communication. Hence, this study presents a novel deep learning (DL) based technique for channel estimation based on channel state information (CSI). A DL technique based on a hyper convolutional neural network (Hyper-CNN) is introduced for the optimal pilot design. The selection of the pilot position can be made using the tunicate swarm optimization (TSO) approach. Finally, a DL-based long short-term memory (LSTM) model is proposed for the CE in the MIMO-OFDM system. The proposed method is implemented in the MATLAB platform, and the outcome is compared under different metrics like mean square error (MSE) and bit error rate (BER). In the experimental scenario, the proposed method attains the BER of 0.20 and 0.03 for CP (cyclic prefix) and without CP, respectively. In addition, the proposed method attains the MSE of 1.14, 0.99, 1.08 and 0.97 for 8, 16, 48 and 64 pilots, respectively. The performance of a proposed method is compared with the existing method and proves the efficacy of the proposed method.

1 | INTRODUCTION

The enormous growth in the area of wireless communication (WC) for multiple types of users with better quality service requirements continues in the following years. Further, due to the advancement of vehicular communication and high-speed railways, MIMO-OFDM in high mobility environment has become a hot research topic.[1] fifth generation (5G) have been introduced by hybridizing various technologies like massive-MIMO (M-MIMO), millimetreWave (mm-Wave) communications and reconstructing surfaces for managing fast growth in reliable communication and wireless

data traffic.[2] OFDM is employed in 5G technology to resist the fading effects and provides better communication quality in multiple path propagation environments. Moreover, the OFDM enhances the SE compared to the single carrier model.[3,4]

MIMO-OFDM can exploit space, time and frequency resources to improve the transmission rate, spectral and power efficiency. Recently, it has been a major technology of broadband WC systems. Obtaining exact CSI by CE is essential to realize the high potential of the MIMO-OFDM model. CE should be known in the base station (BS) to perform the pre-coding. Further, CE is required to detect both MS and BS receivers[5] accurately, and it is important to evaluate CSI with better performance.

In addition, it is essential to realize pre-coding, detection of the signal, and physical layer protection.[6–8] Based on the need for a pilot signal, CE is categorized into blind, decision, and pilot-based CE. Blind CE does not require a pilot signal and performs CE by the second-order mathematical information of the received signal.[9] The pilot-based CE is intended to include the pilot symbol in the transmitted signal, and the receiver evaluates CE based on the obtained pilot signals. It is mainly utilized CE compared to blind CE because of their ease of evaluation and the decision based utilize pilot and detect data symbols and update CE.[10]

When MIMO-OFDM is initialized, pilots are periodically transmitted to the entire subcarrier in blocks. To decode the exact signal, the CSI and its effects at the receiver must be identified.[11–15] To reduce overhead in pilots, orthogonal-based comb pilots are generally utilized, where the data in the pilot is evenly distributed in each OFDM with some specific subcarriers. However, it cannot be utilized in selective frequency channels. The block pilot is used for this case, which periodically transmits the pilot in entire subcarriers. Integrating comb pilots and block pilots is also possible but leads to high computational complexity. This integrated pilot achieves less accuracy in CE.[16–20]

The developing trend in future-generation WCs predicts that the integration of different communication designs, transport alternatives, and switching technologies will become more significant. The techniques used in MIMO-OFDM improve the performance of the WC system by increasing transmission speed, channel capacity, and bit rate. MIMO employs multiple antennas for transmission and reception to take advantage of multipath propagation. Only minor research has been reported to improve the OFDM-MIMO system's BER. However, increasing the capability of a system remains a challenging task. Hence, an effective approach to data transmission in the MIMO system is required. Even though peripheral research is conducted in this field, existing MIMO techniques are highly suffered due to transmission circumstances with larger power consumption. These major drawbacks motivate to develop an enhanced approach to overcome the transmission circumstances in the MIMO system.

1.1 | Motivation

For the MIMO-OFDM system, the correct CE with low pilot overhead is necessary for beamforming and channel detection. But it is difficult to estimate the high dimensionality of MIMO-OFDM using low training overheads. Developing an efficient CE model is complex because of non-stationary and fast time-varying channels. The CE process generally utilizes linear minimum MSE (LMMSE) and least squares (LS) algorithms. However, they are unsuitable for high mobility cases. Further, the traditional models suffer from degradation in performance due to a serious lack of the prior CE. Then the traditional interpolation algorithms cannot perform well in high-mobility environments.

During the last two decades, DL methods have been used in WCs like CE, data detection and channel state information (CSI) to achieve better performance. Generally, the MIMO-OFDM model has multi-subcarrier channels and a transmitter antenna pair; hence several channel variables must be estimated. The conventional pilot-based CE models have a high pilot and huge feedback overhead. Due to this, major losses occur, and performance like energy efficiency and spectral efficiency is affected. The CE performance is based on the positions transmitted by pilots in a time-frequency grid. However determining such positions are a challenging task to DL algorithms. But, the HyperNet CNN based DL can potentially utilize the correlation of adjacent components in frequency, time and spatial domains. Hence, in this work, an optimization-based CNN-based model is used to find the optimized pilot pattern and using this pilot output, LSTM based CE process is performed. This model is developed to analyze the efficiency of hyperNet CNN-LSTM over the CE process and its impact over the performance enhancement. The TSO used in the proposed model not only enhances the performance but also achieve efficient result in less cost and complexity for large processing environment. This merit enhances the efficiency of overall model by selecting the optimal pilot. TSO based optimal pilot selection automatically enhances the CE performance in MIMO-OFDM communication.

1.2 | Related works

Some of the recent existing algorithms used in the pilot design and CE are listed below:

Shalavi et al.[21] introduced downlink CE from frequency division duplex (FDD) on M-MIMO using partial spatial and sparsity models in which the vectors were defined in multi-phases. This work establishes the modified Adaptive Structured Subspace Pursuit (M-ASSP) to solve the sparse model, and the estimated set of every phase was utilized for estimation. It was considered that the spatial pattern of antennas in every group has an equal and unequal part. Then, the proper PD was determined based on a tiny correlation among the sensing matrix columns. This model obtained average normalized mean square error (NMSE) and BER performance than the LS model in a similar simulation environment.

Sun et al.[22] introduced a CE model with the least pilot overhead in filter bank multicarrier (FBMC) with offset quadrature amplitude modulation (OQAM). Half of the pilot overhead was used to transmit data and remove the imaginary interference. The proposed model minimizes the pilot overhead without the pilot energy requirement. Compared to traditional models with three-column pilots, the FBMC-OQAM model required only two-column pilots.

Nandi et al.[23] presented a CE model of M-MIMO-OFDM using Elman Recurrent Neural Network (ERNN). This work aimed to resolve ISI in the transmitter and receiver filter. The ERNN was used to estimate the M-MIMO-OFDM channel, which considered scalability and reliability. Performances like BER, throughput, capacity, and peak-to-average power (PARP) ratio were considered for the evaluation. After the epoch of 40, this model obtained a PARP ratio of 0.127. In addition, the CE using ERNN was compared with other DL models.

Le et al.[24] introduced ML based 5G and beyond CE for the MIMO-OFDM model. The major aim of the work was to use the new CE model with the DL model Bi-LSTM to improve the CE obtained using the LS model. This work considered several channel path profiles for network prototypes over the mobility level estimated by Doppler effects. The model was designed for the random transmitter antennas. The performance of the Bi-LSTM was compared over other models like LS and LMMSE by varying the pilot densities.

Ponnaluru and Penke[25] introduced a DL model to estimate the channel in OFDM systems. Here, 1D-CNN was used to estimate the channel and equalization data was recovered. This model used the channel variation features and derived the additional features from received signals and pilots. The performance of 1D-CNN was compared over other DL models for several digital modulations. Further, it was proved that this model overcomes the distortion, and ISI complicates the wireless channels.

Liao et al.[26] presented a 2D-CNN-biLSTM-based fast time-varying MIMO-OFDM system. Initially, there were two phases: offline training and offline prediction. The CSI generated by the training samples were efficiently utilized by conducting the offline training to the network. The experimentation proved that the proposed model was efficient for the high mobility scenarios in the MIMO-OFDM system.

Jiang et al.[27] introduced CNN-based CE for the MIMO-OFDM system. This work compared the benefits and challenges of CNN-based CE in various fields. A new dual CNN (d-CNN) and an additional network, HyperNet, were introduced for CE. The HyperNet was used to detect the variation from a similar input as the d-CNN. The HyperNet was used to update various parameters and combined with the traditional d-CNN to improve robustness. The experimentation proved enhanced estimation performance for the time-varying cases.

Vidhya et al.[28] developed CE for the MIMO-OFDM system using particle swarm optimization (PSO) and genetic algorithm (GA). This PSO-GA model outperformed the conventional LS and MMSE approaches using the fine-tuning ability. Seyman et al.[29] presented adaptive neuro-fuzzy inference system (ANFIS) for CE in a space–time block-coded model. The ANFIS was trained with accurate CSI and utilized as the channel estimator.

Seyman et al. presented PSO,[30] differential evolution (DE)[31] and ant colony optimizer (ACO)[32] models for pilot tones in the MIMO-OFDM model. Here, the PSO and DE were utilized to optimize pilot tones placement, and the LS was utilized for CE. Finally, Doppler shifts' influence on designing the pilot tones was analyzed. Seyman et al.[33] presented the radial basis function (RBF) for CE over Rayleigh fading channel.

Mashhadi et al.[2] developed the DL-based model for pilot design and downlinking CE for MIMO-OFDM systems. The DL model utilized fully connected layers (FC) for pilot design and outperformed the linear minimum mean square error (LMMSE). The DL model utilized a non-local attention mechanism to learn long-range correlations to enhance CE performance. During the training process, the efficient pilot reduction model was introduced.

Cheng et al.[34] developed maximum likelihood (ML) based adaptive iterative model (AIM) for joint carrier frequency offset (JCFO) and CE in MIMO-OFDM. Here, the de-rotation was utilized to estimate JCFO and the outcome was utilized

to apply a frequency domain equalizer. Then, the AIM was utilized to locate the frequency peaks to estimate JCFO. Finally, the updated JCFO was obtained, and the time variations of parameters were tracked.

1.3 | Contributions

The major contributions of the proposed work are depicted below:

1. To present the pilot-based channel estimation algorithm by reducing the number of pilots and correctly estimate the channel state information.
2. To introduce the hybrid optimization deep learning model to integrate the functions like pilot pattern and CE.
3. To achieve the optimal pilot position and better CE, the model hyper convolutional neural network (CNN)-tunicate search optimization (TSO) and long short-term memory (LSTM) is introduced.
4. The simulation outcomes of the proposed model are compared with the metrics like BER and MSE with CP, without CP, high mobility environment, number of paths and different combinations of transmitters and receivers.

The remaining sections are organized in the upcoming section: Section 2 demonstrates the system model of a proposed MIMO-OFDM model. Section 3 presents the developed model. The results and the discussion are described in Section 4. The conclusion of the proposed study is contemplated in Section 5.

2 | MIMO-OFDM SYSTEM MODEL

Let's assume BS with N antennas has users U in a single antenna. For the modulation, OFDM with S number of sub-carriers is used. Figure 1 determines the system model of the CE-based MIMO-OFDM model.

In the transmitter section, the signal gets modulated using 64-bit QAM by introducing a modulation block into the MIMO-OFDM model. The system with QAM symbols and the particular time slots are integrated to form the data vector $y(s) \in C^M$ as,

$$y(s) = [y_1(s), y_2(s), \ldots\ldots\ y_M(s)]. \tag{1}$$

Here, M indicates the entire symbol modulation. Based on the transmitting antenna N_t, the modulated signal gets spilt into N_t vectors. It can be mathematically formulated as,

$$y_p(s) = \left[y_p(s), y_{p+N_t}(s), ..\right] \quad p = 1, 2, \ \ldots\ N_t. \tag{2}$$

The signal from every antenna is transferred from S/P, and the pilot signals are inserted with the transmitter and receiver signal for CE purposes. Let's assume $y_b(s)$ whereas, $b = 1, .. \ \ldots\ , N_t$ as the signal vector has inserted pilots into particular data $y_p(s)$. Finally, an IFFT block is given to $y_b(s)$; thus, the signal is converted from the frequency domain (FD) into the time domain (TD).

$$\widehat{y}_b(s) = IFFT(y_b(s)). \tag{3}$$

After applying FFT, CP is inserted with the length N_{cp} as a guard interval to compromise the intermediate interferences among the symbols using the CP insertion block. With the use of CP, the transmitted signal in TD is represented by,

$$\left[\widehat{y}_{cb}(s)\right]_m = \left\{ \begin{matrix} \left[\widehat{y}_b(s)\right]_{m+N_{fft}} & m = -N_C, -N_C + 1, \ldots\ldots\ -1 \\ \left[\widehat{y}_b(s)\right]_m & m = 0, 1\ldots\ldots\ N_{fft} - 1 \end{matrix} \right\}. \tag{4}$$

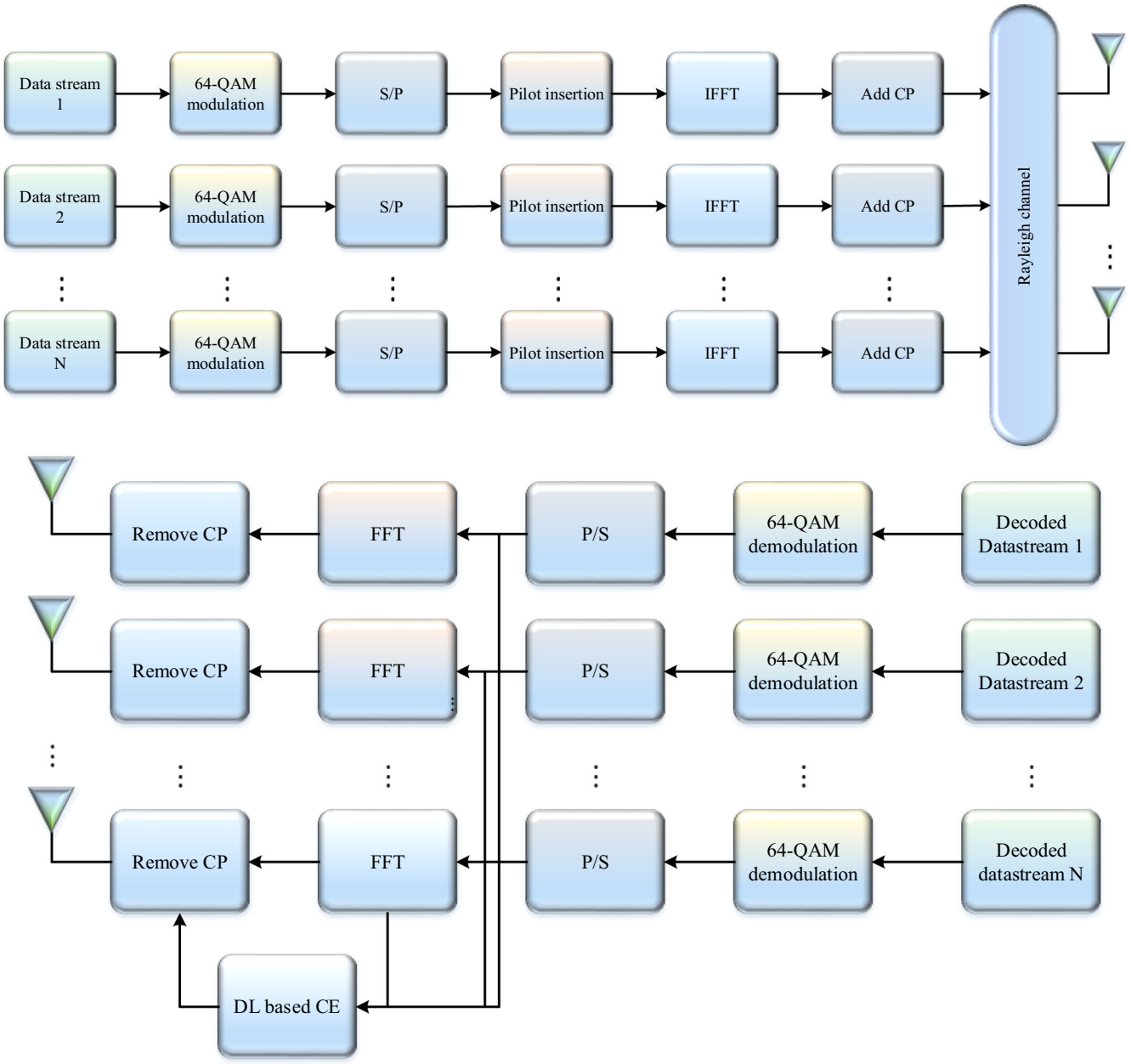

FIGURE 1 Multiple input multiple output-orthogonal frequency division multiplexing system model.

Here, N_{fft} indicates the size of a FFT. The N_C samples of $\left[\hat{y}_b(s)\right]$ are considered as the CP and inserted in the initial stage. The channel estimation accuracy degrades on the receiver side for high-mobility users. In a high-mobility environment, channels are time-variant and frequency-selective, reducing symbol transmission due to Doppler spreading. This Doppler Effect can be mathematically formulated as,

$$D(F) = \frac{1}{\pi F_d \sqrt{1 - \left(\frac{F}{F_d}\right)^2}}, \; F < F_d. \tag{5}$$

Here,F_d indicates the maximum Doppler shift, and it is represented as, $F_d = \frac{vF_c}{C}$. Here, v indicates the speed (m/s), F_c indicates the carrier frequency, C indicates the speed of light having the value 3×10^8.

The Rayleigh channel is used for signal transmission, described in the FS and time-variant channel model. This can be mathematically formulated as,

$$H_{x,y}(\alpha_l, s) = \sum_{l=0}^{L-1} H_l \mu(\alpha_l - s) * e^{j[2\pi f_{D,l}(s-\alpha_l) - 2\pi f_c \alpha_l]}. \tag{6}$$

Here, l indicates the index tap, H_l indicates the l^{th} resolved amplitude and α_l denotes the l^{th} tap delay. $f_{D,l} = v(s) f_c \cos(\theta_l)/c$ represents the Doppler frequency for both transmitter and receiver, $v(s)$ indicates the relative velocity, θ_l signifies the phase angle and c indicates the light speed.

The received signal for the multipath Rayleigh channel with the transmitted signal can be mathematically expressed as,

$$\hat{x}_{ca}(s) = \sum_{b=1}^{N_t} \hat{H}_{x,y}(\alpha, s) \otimes y_{cb}(s) + \hat{m}_a(s). \tag{7}$$

Here, $\hat{H}_{x,y}(\alpha, s) = \left[H_{x,y}(\alpha_1 s), \ldots H_{x,y}(\alpha_L s)\right]$, $\hat{m}_a(s)$ indicates the additive noise vector.

In the receiver side, CP is eliminated from the received signal $\hat{x}_{ca}(s)$ in every antenna using the CP elimination block. The signal is then transformed into an analogous form and is given to the FFT block. The FFT block converts the TD signal into an FD signal, and it is expressed as,

$$x_a(s) = FFT\left(\hat{x}_a(s)\right). \tag{8}$$

The pilots are extracted from the FD signal to estimate the channel accurately. After CE, the signal is demodulated in the demodulation block. The final signal from the MIMO-OFDM system is generated as a data sequence in binary form.

3 | PROPOSED METHODOLOGY

The hybrid processing model is utilized for the MIMO-OFDM system to reduce the cost and complexity. CE plays a major part in the network's performance of the system. But, the accuracy of CE is affected due to the interference, complexity and pilot resources. LS (least squares) estimation is generally used to achieve CE, and this model has a large estimation error. Recently, Deep learning (DL) based models can increase the performance of CE, specifically in extreme environments.

Further, deep learning models have significantly improved communication reliability and minimized the system's complexity in the MIMO-OFDM model. This work proposes a DL model for pilot design and CE to enhance the accuracy of CE with low pilot overhead. Here, a downlink CE for the MIMO-OFDM model is considered. The main goal of this work is to use the better feature extraction capacity of optimized Hyper-CNN to train the LSTM network for accurate channel estimation. This work introduces optimized hyper CNN-LSTM comprised of hyper CNN-Tunicate search optimization (TSO) and LSTM. Initially, optimized hyper CNN is used to find and select the optimized pilot design during the training process, and LSTM is to estimate the approximate CE.

3.1 | Optimal pilot design

In this model, the MIMO with channel impulse response (CIR) model is introduced with transmittern_T and receiver antennas n_r.[35] OFDM is used to modulate the signal at the transmitter side (i_T) and CP is introduced to protect the input signal. Then, the CP is removed at the receiver end (i_R) which is defined in Equation (9),

$$\vec{Y}_{i_R} = \sum_{i_T=1}^{nT} \vec{X}_{iT} \otimes \vec{H}_{i_T, i_R} + \vec{W}_{i_R}, \tag{9}$$

where, $\vec{X}_{i_T}$ indicates the transmit signal and $\otimes$ indicates transmit signal for the circular convolution of CIR $\vec{H}_{i_T,i_R}$. Then, the CIR for i_T and i_R is represented in Equation (10),

$$\vec{H}_{i_T,i_R} \equiv \sqrt{diag\left(\vec{P}_{i_T i_R}\right)\vec{G}_{i_T i_R}}. \tag{10}$$

Between i_T and i_R, the available normalized power delay is indicated as $\left(\vec{P}_{i_T i_R}\right) \in \Re^{L\times 1}$, the random Gaussian variable containing unit vector is indicated as $\left(\vec{G}_{i_T i_R}\right) \in \subset^{L\times 1}$.

Due to circular convolution, each channels in the frequency domain are in diagonal position, and the obtained signal at the i_R (receiver antenna) is shown in Equation (11),

$$\vec{Y}'_{i_R} = \sum_{i_T=1}^{n_T}\left(X'_{P,i_T} + X'_{d,i_T}\right)\vec{H}'_{i_T,i_R} + \vec{W}'_{i_R}, \tag{11}$$

where, the diagonal matrix corresponding to pilot or data sequence d ($i.e.s \in \{p,d\}$) is indicated as, $X'_{s,i_T} = diag\left(X'_{s,i_T}\right)$. The vector form of AWGN channel at the receiver antenna $i_R$having variance σ_W^2 is indicated as $\vec{W}'_{i_R}$. The total pilot subcarriers at the frequency domain is found lesser than the data subcarrier, that is $\Delta k \succ 1$. However, only a few ubset of observations having samples $N_P = [N/\Delta k]$ will be used for pilot-aided CE, as the selected subset carries required information about pilot. The signal received by subcarriers suitable for pilot design is represented as,

$$\vec{Y}_{i_R} = \sum_{iT=1}^{nT}\left(X_{P,i_T} + X_{d,i_T}\right)\vec{H}_{i_T,i_R} + \vec{W}_{i_R}. \tag{12}$$

3.1.1 | TSO based pilot design

Recently, many techniques have been introduced to select the pilot position accurately. However, the conventional approaches are highly suffered because of time complexity and errors during data transmission. To overcome this, a stable optimization technique named tunicate swarm optimization (TSO)[36] is introduced to select the pilot position with low time complexity accurately. The tunicate is one of the marine invertebrates which emits a brighter bio-luminescence. The unknown behavior of ocean tunicates highly inspires it due to their foraging process. Based on the three constraints, the mathematical formulation is manipulated. The three constraints are utilized: avoiding disputes between the exploration agents, succeeding in a position of the most intelligent agents and finally, near the optimal agents.

3.1.2 | Avoiding disputes among the pilot exploration

By avoiding inter-pilot disputes, a better position can be chosen. The new pilot position is mathematically formulated as,

$$\vec{B} = \frac{\vec{H}}{\vec{N}}, \tag{13}$$

$$\vec{H} = r_2 + r_3 - \vec{D}, \tag{14}$$

$$\vec{D} = r_1 * 2. \tag{15}$$

Here, $\vec{B}$ represents the new pilot position vector, $\vec{H}$ indicates the false pilot position, $\vec{D}$ indicates the pilot interference, r_1, r_2, r_3 manipulates the random numbers. The public forces between the pilots are gathered in a new vector, and it is mathematically formulated as,

$$\vec{N} = [Q_{\min} + r_1 * Q_{\max} - Q_{\min}], \tag{16}$$

whereas $Q_{\min}$ and $Q_{\max}$ have the values 1 and 4, and it is represented as first and second subordinates, respectively.

3.1.3 | Succeeding the position of most intelligent pilots

Based on the present best pilots, the optimal position is chosen accurately. The best position of the intelligent swarm can be chosen after confirming that no disputes arises between the agents. It is mathematically computed as follows:

$$\vec{QF} = \left|Y_{best} - c_{rand} * Q_p(y)\right|. \tag{17}$$

Here, $\vec{QF}$ indicates the dimension between the optimal pilots and the pilot placement, Y_{best}indicates the best position, c_{rand} indicates the random value with the range [0, 1] and $Q_p(y)$ signifies the pilot position during the y^{th} recursion.

3.1.4 | Getting near to the optimal pilots

The position is updated to ensure the search pilots from the adjacent best pilot. It can be mathematically expressed as,

$$\vec{Q}_p(y) = \left\{ \begin{matrix} Y_{best} + B \times \vec{QF}, & if\, c_{rand} \geq 0.5 \\ Y_{best} - B \times \vec{QF}, & if\, c_{rand} < 0.5 \end{matrix} \right\}. \tag{18}$$

Here, $\vec{Q}_p(y)$ indicates the position updation of each pilot at recursion y to the best pilot positionY_{best}.

3.1.5 | Swarming behavior

The swarming behavior of tunicates can be manipulated by updating the position of a present pilot, and it can be formulated as,

$$\vec{Q}_p(\vec{y}+1) = \frac{\vec{Q}_p(y) + \vec{Q}_p(\vec{y}+1)}{2 + r_1}. \tag{19}$$

3.1.6 | Fitness function

Finally, the best pilot position can be updated by minimizing the MSE during data transmission. The minimization of MSE is considered a fitness function. It is represented as,

$$Fitness\ function = \min(MSE). \tag{20}$$

Algorithm 1 shows the pseudocode of TSO, and to clarify the TSO, the steps followed in the TSO are given shortly below:

Step 1: Start with the initial tunicate population $\vec{Q}_p$.
Step 2: Set the parameters to the original value and the highest number of recursion.
Step 3: Calculate the fitness of each pilot in the exploration stage.
Step 4: After evaluating fitness value, the best pilot in the search space is examined.
Step 5: Positions of each exploration pilot are updated using Equation (19).
Step6: Return to the newly updated pilot within the limit.
Step 7: Calculate the fitness cost of an updated search pilot $\vec{Q}_p$ and gather the best solution as Y_{best}.
Step 8: If the termination condition is satisfied, stop the operation. Otherwise, repeat the process from steps 5–8.
Step 9: Finalize the best optimal solution Y_{best}.

Algorithm 1. : Pseudocode of TSO

Input: Initialize tunicate population $\vec{Q}_p$.
Output: The best optimal solution Y_{best}.
Initialize the parameters like $\vec{B}$, $\vec{H}$, $\vec{D}$ and max_*iter*
while $y <$ max_*iter*) **do**
  for $j \leftarrow 1$ to 2 **do**
Calculate the fitness of each pilot using Equation (20)
Avoiding disputes among the pilot exploration using Equations (13)–(15)
Getting near to the optimal pilots
if $c_{rand} \geq 0.5$ then
Update $Y_{best} + B \times \vec{QF}$,
  else
    Update $Y_{best} - B \times \vec{QF}$
end if
end for
Compute swarming behavior using Equation (19)
Update the parameters like $\vec{B}$, $\vec{H}$, and $\vec{D}$
$y = y + 1$
end while
Return best optimal solutionY_{best}.
end

3.2 | Channel estimation using the LSTM model

The sparse massive MIMO system is assessed by suggesting a deep CNN-LSTM model channel estimator after developing the optimal pilot training data selection. A channel is referred to as sparse if its length is greater than the number of multipath channels.[37]

$$C(N) = V_N + \gamma \ln(\ln(L))(N+1), \tag{21}$$

where, Cindicates the generalized Akaike information, γ is used to represent path losses exponent of user, N indicates signal length, and error reflected by channel to particular length is represented as V_N.

CE is one of the key roles in WC since the speed and precision of channel estimation greatly influence the performance and capacity of communication. Hence, this research uses a DL-based LSTM model to estimate the channel accurately. The network structure of the LSTM model is separated into three parts, namely the input layer (IL), hidden layer (HL) and output layer (OL). The input data is transmitted through the IL, and the LSTM cells are in the HL and the OL evaluates the estimated channel. The LSTM cell consists of an input gate (IG), forget gate (FG) and an output gate (OG). The optimized pilot and transmitting signals are given as input to the input layer. Then, the input gate analyses the input data and eliminates the unwanted information that enters the memory unit. The FG f_g determines to eliminate the unwanted previous state information c_{g-1} from the cell. The combination of both FG and IG updates the memory cell state c_g. The OG determines the outcome to be outputted. The following formulation determines the LSTM network model:

$$f_g = \alpha\left(w_f[H_{s-1}, y_s] + B_f\right), \tag{22}$$

$$g_s = \alpha\left(w_g[H_{s-1}, y_s] + B_g\right), \tag{23}$$

$$c'_s = \tanh(w_c[H_{s-1}, y_s] + B_c), \tag{24}$$

$$c_s = f_s * c_{s-1} + g_s \times c'_s, \tag{25}$$

$$O_s = \sigma(w_0[H_{s-1}, y_s] + B_0), \tag{26}$$

$$H_s = O_s \times \tanh(c_s). \tag{27}$$

Here, c_s and H_s signifies the cell state (CS) and output, respectively, c_{s-1} indicates the CS, H_{s-1}represents the previous cell output. w_f, w_g, w_0 and w_c indicates the weight matrix of IG, FG, OG and CS, respectively. b_f, b_g, b_0 and b_c signifies the bias vector of IG, FG, OG and cell vectors, respectively. The LSTM-based channel estimator uses circular orthogonal sequences, which is much better than the LS estimator. Based on the hyper-network training, the number of pilot blocks converges. Similar rules are fed to the pilot sequence, and each transmitter must be orthogonal. Figure 2 illustrates the schematic diagram of the LSTM model.

As the number of pilot sequences increases, pilot overhead and contamination may occur automatically. But its orthogonal property improves the estimation accuracy of different antennas via a correlation operation.

Let's assume the number of transmitters as M_t with the pilot blocks P_1 and P_2 are introduced based on the Zadoff-chu series. With the length L, the zadoff-chu sequence can be formulated and is mathematically formulated as,

$$Z(l) = \begin{Bmatrix} \exp(j\pi l^2 i/L) & if\ L\ is\ even \\ \exp(j\pi l^2 (i+1)/L) & if\ L\ is\ odd \end{Bmatrix}. \tag{28}$$

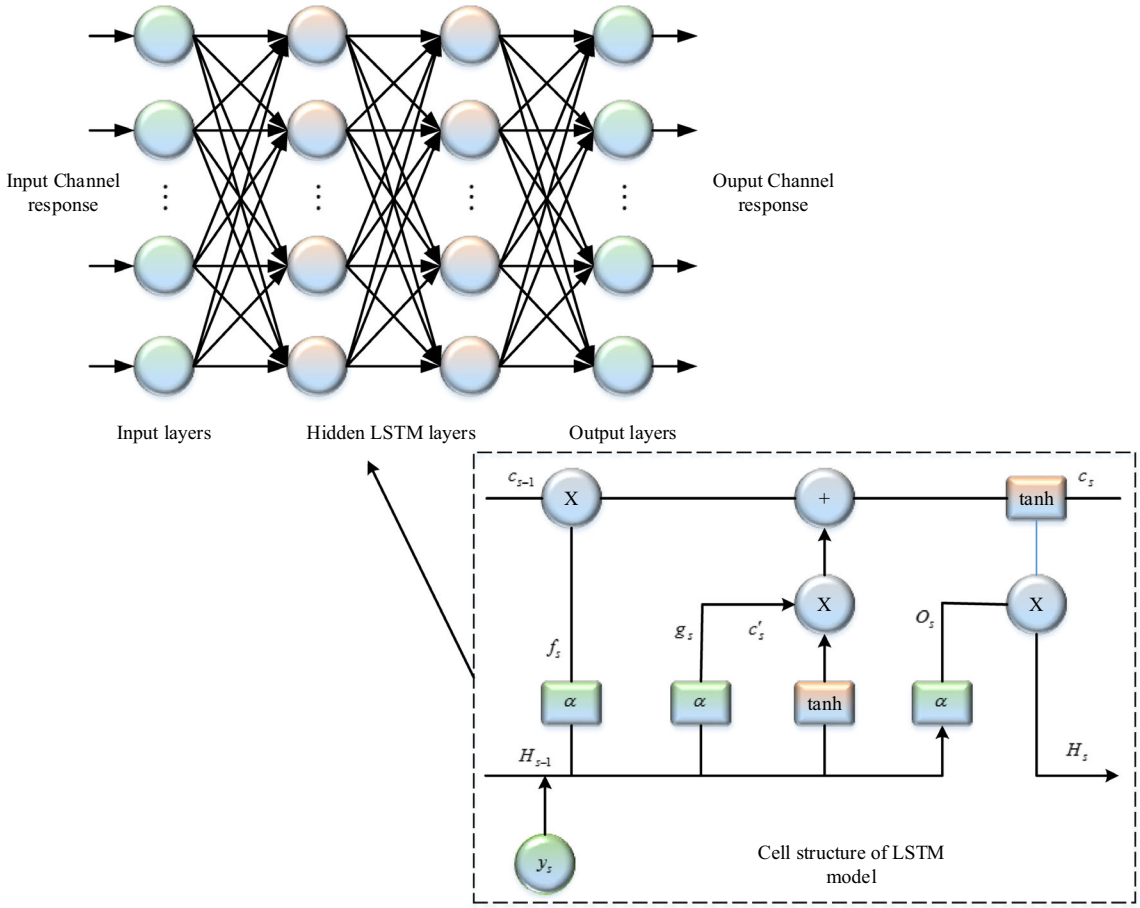

FIGURE 2 Schematic diagram of the long short-term memory model.

Here, l indicates the l^{th} symbol sequence, i indicates the co-prime integer. The m^{th} transmitting antenna with the zadoff-chu sequence can be generated using the following expression:

$$z_r = z(l + (r-1) \in L) \qquad r = 1, 2, \ldots \ldots, M_t. \tag{29}$$

Here, $\in$ indicates the shift step. By adding CP, the blocks are constructed. The initial P number of symbols is replicated, and cyclic postfix is integrated. To eliminate the ISI, the channel delay must be less than the length of a CP.

After removing the CP from the receiver side, the signal S for the n^{th} receiving antenna and the channel response from the transmit antenna to the receiving antenna is represented as,

$$\widetilde{H}_{mn}(q) = \left(\frac{1}{L}\right)\sum_{l=0}^{L-1} z_r(l)S(l+q), \quad q = 1, 2, \ldots, Q. \tag{30}$$

The hyper-CNN network starts training after predicting the channel from the LSTM model. Let's consider the time series channel response as, $(H(1)H(2)\cdots H(m-1)H(m))$, whereas $H(m)$ represents the channel response at a random time. The input sequences for the set of channel responses are fed to the LSTM network, and the final data indicates the output label.

Let's consider the sliding window with the dimension d, and training begins with the first s vector of a time series channel response. The existing $d-1$ vector is considered input to the LSTM network. When the new channel response has arrived, the vector window moves one step forward. This process is done continuously until the channel accurately converges. After training the pilot block in an LSTM network, it is implemented to analyze the probable channel response in the data block.

To update the LSTM network, the channel responses are resent for training. A probable training loop is generated and trained until all communication in the data block completes. The main advantage of this LSTM model is that it can predict channel estimation according to the recursive loop.

3.3 | Training process using hyper-CNN model

Generally, the MIMO-OFDM system is highly affected due to pilot contamination. This contamination occurs during uplink transmission when the receiver uses similar orthogonal signals. Recently, many techniques have been contemplated to design optimal pilot sequences in the MIMO system. But, those techniques are highly suffered due to computation complexity and increased pilot overhead. Hence, an optimal DL-based hyper-CNN technique is introduced to overcome the pilot contamination effect in the MIMO system. The hyper-CNN model consists of multiple trainable parameters which can adapt well under high noise power and increased channel statistical state.

Figure 3 demonstrates the hyper-CNN architecture. To achieve more adaptability, an extra spatial frequency convolutional neural network (SFCNN) is used, which can be under noisy input. In addition, the adaptive dilation convolutional neural network (ADCNN) approach is introduced to reduce the noisy input without affecting the system channel. In contrast, the noise and the channels are integrated into the SF domain, and channel correlation is considered the key SFCNN denoiser. The predicted channel from the LSTM model is given as input to the hyper-CNN. The input is convoluted using 3×3 filters and increases twice by 3D and the ReLU activation function. The down-sampling process degrades the initial 2D in half. The down-sampling and convolution process is repeated double times. The FC layer is utilized to produce the output of trainable parameters. The sigmoid function allows the output from 0 to 1, and triple trainable parameters are multiplied with SFCNN's output and control the limitations in SFCNN. The operation occurs in hyper-CNN, and trainable parameters are represented as $p_{hyper}(.)$ and $\beta_{hyper}(.)$, respectively. There are three resultant parameters, and it is represented as, $\phi = [\phi_1, \phi_2, \phi_3]$. It is mathematically formulated as,

$$\phi = p_{hyper}\left(\widetilde{H}_{LSTM}; \beta_{hyper}\right). \tag{31}$$

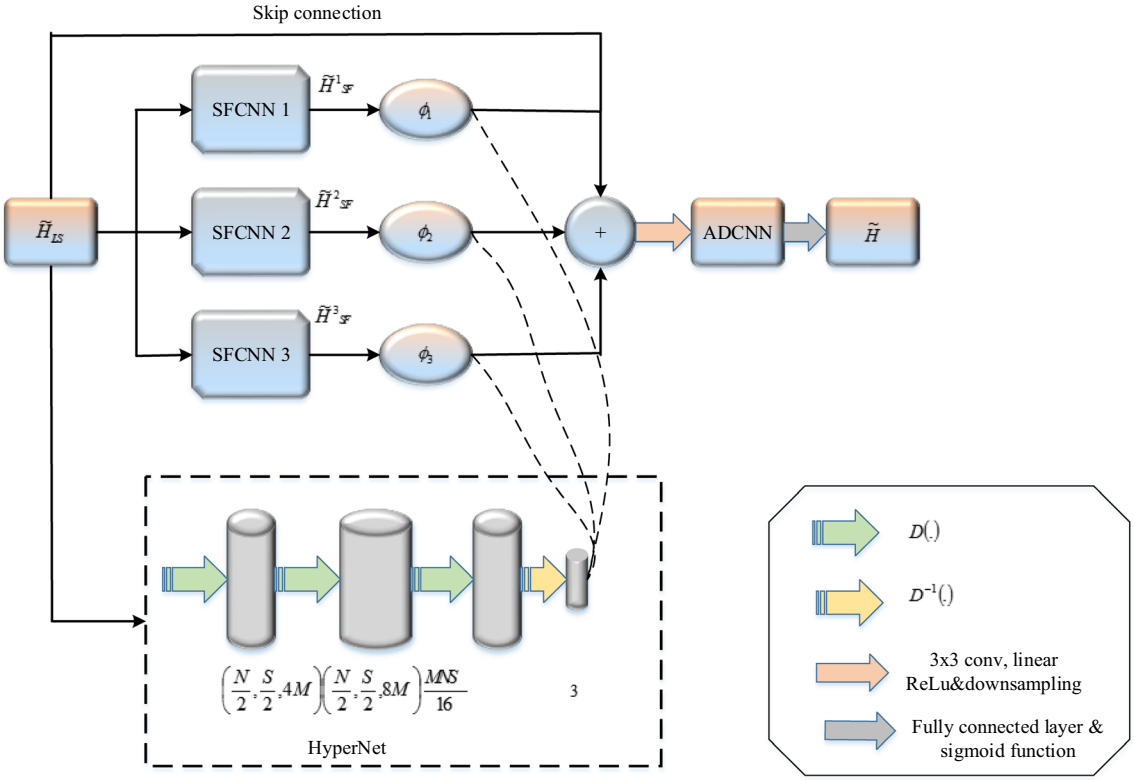

FIGURE 3 Architecture of hyper convolutional neural network model.

Here, β_{hyper}indicates the training parameters in hyper-CNN and $\widetilde{H}_{LSTM}$ is LSTM trainable parameters. The outcome of hyper-CNN can be mathematically formulated as,

$$\widetilde{H} = D^{-1}\left(f_{CNN}\left(D\left(\sum_{m=1}^{3}\phi_m\widetilde{H}_{SF}\right);\beta_{AD}\right) + D\left(\sum_{m=1}^{3}\phi_m\widetilde{H}_{SF}\right)\right) \tag{32}$$

Here, the output of m^{th} SFCNN is represented as, $\widetilde{H}_{SF}$, the domain transform (time) process and its reverse process (frequency) is denoted as $D(.)$ and $D^{-1}(.)$.

To improve the training process, three major steps are undertaken.

Initially, ADCNN is trained throughout the training process without using SFCNN. This training process can be mathematically formulated as,

$$\widetilde{\beta}_{AD} = \arg\min_{\beta_{AD}}\left\|D^{-1}\left(f_{CNN}\left(D\left(\widetilde{H}_{LSTM}\right);\beta_{AD}\right) + D\left(\widetilde{H}_{LSTM}\right) - H\right)\right\|_2^2. \tag{33}$$

The ADCNN uses end-to-end processes to improve performance during training and testing. The CNN training mainly relays on ADCNN for SF domain pre-denoising. Hence, the ADCNN is trained separately to enhance the performance under the untrained part in which the SFCNN are not applicable.

Three main categories are required for better training of SFCNN. But, the outcome of SFCNN is less, and the whole training process can be mathematically formulated as,

$$\widetilde{\beta}_{SF} = \arg\min_{\beta_{SF}}\left\|D^{-1}\left(f_{CNN}\left(D\left(\widetilde{H}_{SF}\right);\widetilde{\beta}_{AD}\right) + D\left(\widetilde{H}_{SF}\right) - H\right)\right\|_2^2. \tag{34}$$

The proposed method shows better performance compared with the traditional CNN technique. Due to hybrid SFCNN and ADCNN approaches, it can work well even in untrained categories. Table 1 tabulates the hyperparameters of a proposed system.

TABLE 1 Hyperparameters of the proposed system.

Hyperparameters	Proposed hyper CNN-TSO-LSTM model
Learning rate	0.01
Batch size	1000
Epoch	100
Drop factor	0.1
Optimizer	TSO
Hidden units	16
Filter size of hidden layers	100
Loss	MSE
Input dimension	256
Output dimension	256
LSTM neurons	256
Training samples	128 000
Testing samples	10 000
Validation samples	32 000
Number of paths	20
Doppler frequency	36–200 Hz

4 | RESULTS AND DISCUSSION

The proposed method is implemented in the MATLAB tool. The performance of a proposed model is compared with different metrics like BER and MSE with CP, without CP, different channels and different pilot numbers. For experimentation, 50 transmitting and 50 receiving antennas are used with the 128 subcarriers. In the OFDM model, 2500 symbols are used along with CP. The total BW of 3 kHz with a carrier frequency of 2.6GHz and CP length of 16 is considered. The maximum delay spread is about 6 μs. In addition, the Rayleigh channel model is used with 64 QAM modulation for effective modulation of the transmitting and receiving signal. Table 2 shows the simulation parameters of a proposed system.

TABLE 2 Simulation parameters.

Parameters	Values
Number of paths	20
Doppler frequency	36–200 Hz
Fading channel	Rayleigh
Modulation	64 QAM
Transmitting antennas	50
Receiving antennas	50
Subcarriers	128
BW	3 kHz
Carrier frequency	2.6 GHz
CP length	16
Delay spread	6 μs

4.1 | Performance metrics

The BER is the ratio of number of errors on the receiver side to the total number of bits at a transmitter side.

$$\text{BER} = \frac{\text{number of error bits at the receiver}}{\text{total number of bits at the transmitter}}. \tag{35}$$

MSE is the average of square of the difference between original and calculated value.

$$X = \frac{1}{m}\sum_{u=1}^{m}\left(G_u - \widetilde{G}_u\right)^2. \tag{36}$$

Here, X signifies the MSE, m denotes the number of data used, G_u signifies the actual value and $\widetilde{G}_u$ signifies the calculated value.

The NMSE is the normalization of a MSE under varying signal power. It can be mathematically formulated as,

$$NMSE = e\frac{\|h - \widetilde{h}\|^2_2}{\|h\|^2_2}, \tag{37}$$

Here, e indicates the error signal, h and $\widetilde{h}$ indicates the real and predicted channels, respectively.

4.2 | Analysis of channel estimation using the LSTM model

This section illustrates the analysis of CE using the LSTM model. Figure 4A,B determines the accuracy and loss curve under training, testing and validation. For analyzing the superiority of the proposed method, a total of 100 epochs are considered. From the graphical illustration, it is clear that the proposed model obtains high training accuracy of 95%. For the validation process, the accuracy attained about 94.9%, and for the testing process, the accuracy attained about 93%. As the number of epochs increases, the accuracy automatically enhances. In addition, an error of 0.01 is obtained for the training process. For the testing and validation, there is an error of 0.1 and 0.5.

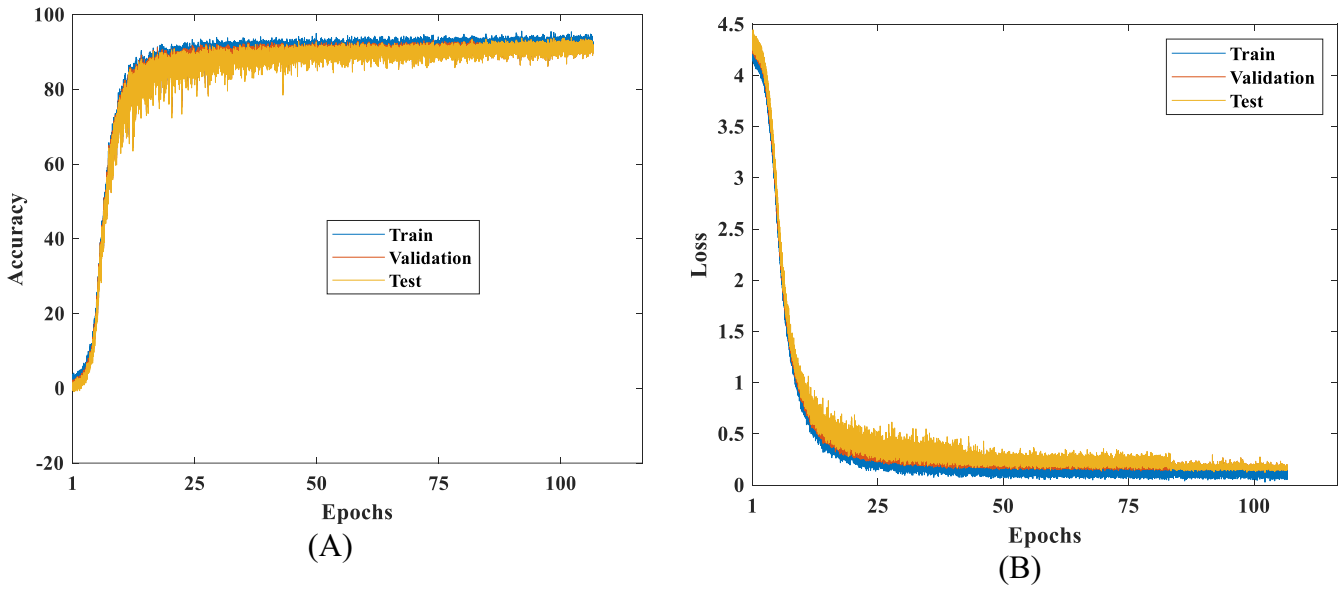

FIGURE 4 Accuracy and loss curve for training and testing samples, (A) accuracy curve and (B) loss curve.

4.3 | Performance comparison of the developed model with conventional models

In this section, the performance of a developed method is compared with various optimization and DL techniques. Some performance measures, such as BER and MSE, are analyzed under CP, channels, and pilot numbers. The graphical illustration proves that the proposed system works better than conventional techniques.

Figure 5 illustrates BER performance under various existing methods for 64 pilots. Figure 5A demonstrates the BER performance with the presence of CP. From the graphical illustration, it is clear that the developed technique shows outstanding performance compared to other techniques. The error is reduced by automatically increasing the signal-to-noise ratio (SNR). The CP overcomes the ISI by converting a linear convolution into a circular convolution. Therefore, it enhances the spectral efficiency performance. The existing LS technique shows the worst performance for higher pilots. For SNR = 20 dB, the existing MMSE, LS and the proposed LSTM attain the BER of 0.77, 0.74, and 0.0011, respectively. Figure 5B demonstrates the BER performance without the presence of CP. The graphical illustration clearly shows that the BER can be reduced without using CP in the MIMO-OFDM model. The absence of CP degrades the spectral efficiency gradually. For SNR = 20 dB, the existing MMSE, LS,[24] and the proposed LSTM attains the BER of 0.23, 0.14, and 0.0016, respectively.

Figure 6 illustrates the BER performance under various existing methods with 8 pilots. Figure 6A signifies the BER performance with CP under fewer pilot numbers. The graphical analysis show that, the BER is reduced when the SNR gets increases. When fewer pilots are used, the training time increases, resulting in time consumption and computational complexity. Thus, the performance also decreases when fewer pilots are utilized. For the SNR = 20 dB, the existing MMSE, LS,[24] and the proposed LSTM attain the BER of 0.20, 0.13, and 0.10, respectively. Figure 6B signifies the BER performance without CP under fewer pilots. The graphical illustration shows that the error increases without CP, and the performance of a network model decreases during channel estimation. For the SNR = 20 dB, the existing MMSE, LS,[24] and the proposed LSTM attain the BER of 0.70, 0.81, and 0.05, respectively.

Figure 7 illustrates the BER performance under various optimization techniques. Figure 7A indicates the BER performance with eight pilot numbers. The graphical illustration clearly shows that the existing optimization techniques have a higher error than the proposed TSO technique. In general, the smaller pilot numbers cannot improve system efficiency. However, the proposed method gradually improves spectral efficiency more than ADAM and SGDM techniques. For SNR = 20 dB, the existing ADAM, SGDM, and the proposed TSO techniques attain the BER of 0.26, 0.052, and 0.003, respectively. Figure 7B indicates the BER performance under 64 pilot numbers. The proposed MIMO-OFDM model shows its enhanced performance under more pilot numbers. A system with high pilot numbers can efficiently reduce the error

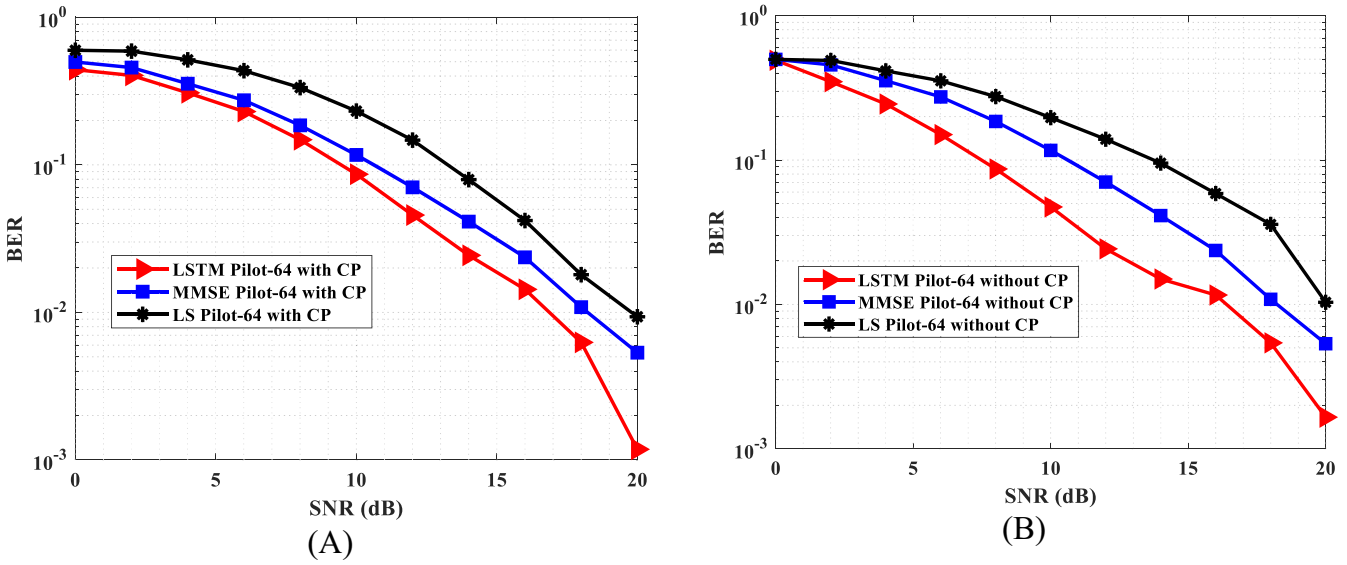

FIGURE 5 Bit error rate (BER) performance under various existing methods for 64 pilots, (A) BER results with channel estimation (CP) and (B) BER results without CP.

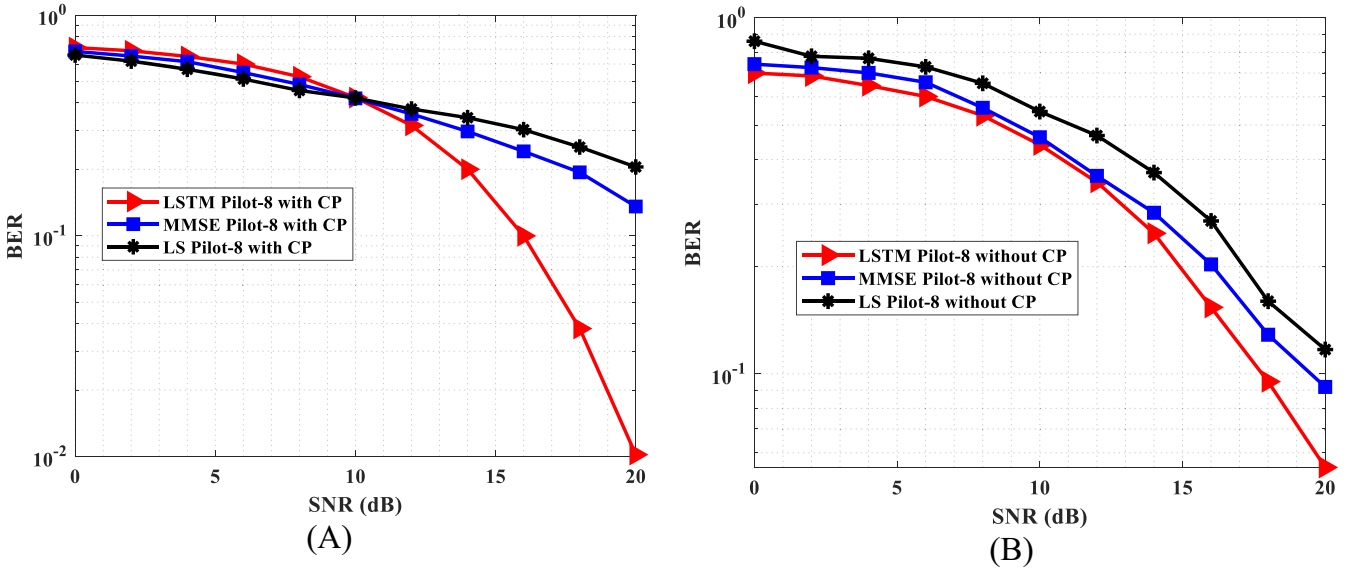

FIGURE 6 Bit error rate (BER) performance under various existing methods with eight pilots, (A) BER results with channel estimation (CP) and (B) BER results without CP.

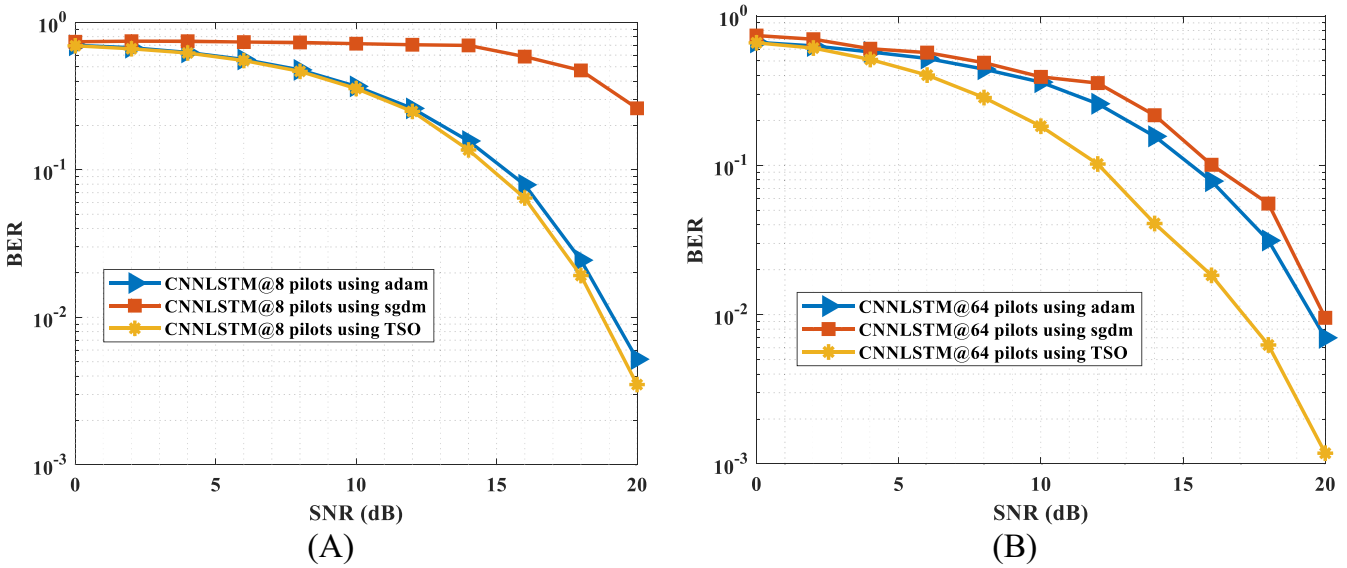

FIGURE 7 Bit error rate (BER) performance under various optimization techniques, (A) BER results with 8 pilots and (B) BER results with 64 pilots.

on the receiving side. For SNR = 20 dB, the existing ADAM, SGDM and proposed TSO techniques attain the BER of 0.007, 0.52, and 0.0011, respectively.

Figure 8 shows the MSE analysis under varying SNR. As the SNR is automatically increased, the MSE is reduced. Compared with the Rician and AWGN channels, the Rayleigh channel performs better during data transmission. The channel from fading and unwanted data loss is eliminated by using the Rayleigh channel. The Rician channel shows the worst performance of all channels due to high data loss and poor capability. For the SNR = 20 dB, the AWGN, Rician and Rayleigh channel attains the MSE of 0.37, 0.59 and 0.40, respectively.

Figure 9 illustrates the MSE analysis under pilots 8, 16, 48 and 64, respectively. From the graph, it is clear that the proposed method improves performance when more pilots are used. With the increase in pilot numbers, the estimation accuracy remains the same for 16 and 48 pilot numbers. It can be seen that the proposed CNN-LSTM technique has better

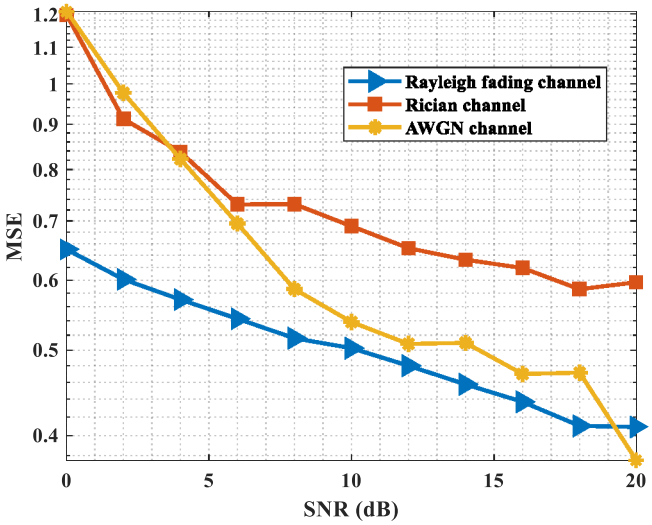

FIGURE 8 Mean square error versus signal-to-noise ratio.

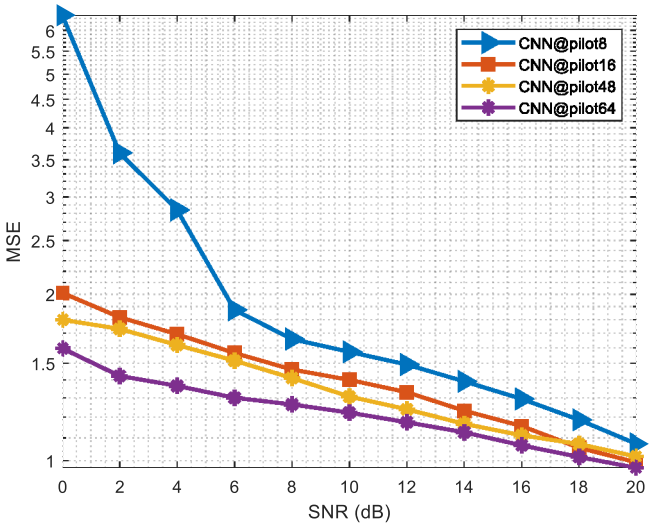

FIGURE 9 Mean square error analysis under pilots.

robustness to varying pilots for the lesser SNR values. Further, it can also enhance the SE of the MIMO-OFDM system to ensure CE accuracy. For the SNR = 20 dB, the MSE attained the values of 1.14, 0.99, 1.08 and 0.97 for 8, 16, 48 and 64 pilot numbers, respectively.

Figure 10 illustrates the comparative analysis of BER under various existing methods. From the graphical illustration, the proposed system CNN-LSTM technique attains a low error compared to conventional approaches. The traditional methods attain high error due to high channel fading and data loss. The proposed MIMO-OFDM system uses a highly enhanced channel to prevent the channel from fading and estimate the channel accurately. But, the existing RNN-LSTM and ELM-LSTM failed to mitigate the interference and update the channel state information. For SNR = 20 dB, the existing RNN-LSTM,[27] ELM-LSTM,[38] and the proposed CNN-LSTM attains the BER of 0.0015, 0.007, and 0.0010, respectively.

Figure 11 manipulates the NMSE performance under different approaches. The graphical illustration shows that the developed model attains a low NMSE metric compared to existing approaches. Even though the refined LMMSE obtains

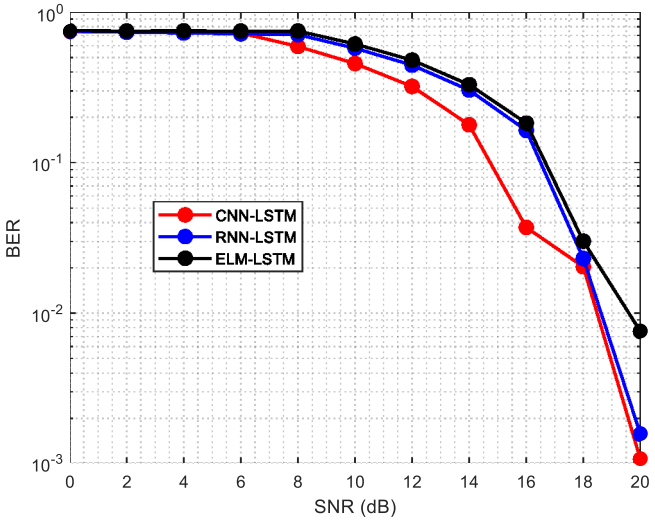

FIGURE 10 Comparative analysis of bit error rate under various existing methods.

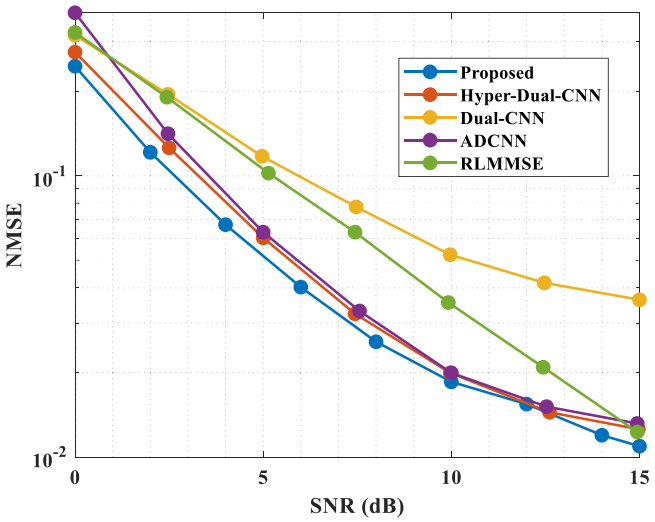

FIGURE 11 Normalized mean square error performance under different approaches.

better performance, it suffers greatly from the time complexity than the proposed method. When $SNR \leq 15dB$, the proposed model works without the use of online training and performs better than the existing RLMMS approach. The existing SFCNN approach cannot operate well under untrained processes, so the ADCNN approach is used throughout training the samples. However, channel power is one of the important parameters, and it is also required even under the offline process. The proposed system works better when $SNR \leq 5dB$, and it is hybridized with SFCNN, which increases the gain even if the noise power is high. When SNR = 15 dB, existing dual CNN, ADCNN, hyper dual CNN, RLMMSE[27] and proposed method attains the NMSE of 0.24, 0.12, 0.06, 0.02, and 0.01, respectively.

Figure 12 depicts the MSE performance under 8 and 16 pilots. The graphical representation shows a lower MSE for the proposed method than existing approaches. when the number of pilots, the performance is increased. The existing CA-GAN, channelNet, interpolated MMSE, and LS[24] cannot perform well even if the number of pilots increases and it is due to the low capability of a system. The proposed system uses a highly enhanced strategy to increase estimation accuracy. When SNR = 20 dB, the proposed method attains the MSE of 0.013 and 0.025 for 16 and 8 pilots, respectively.

Figure 13 illustrates the BER performance under different Doppler frequencyF_d. The existing MMSE and LS[24] approaches show the worst performance at SNR = 20 dB. But the proposed method uses the TSO optimizer to minimize

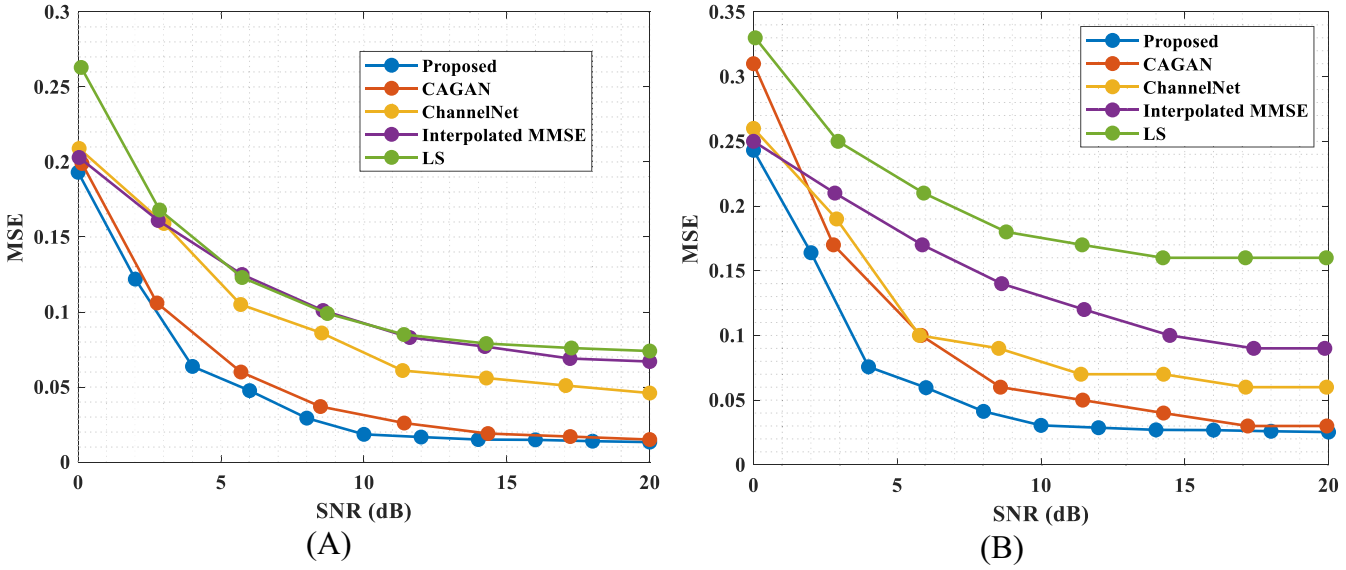

FIGURE 12 Mean square error performance under 8 and 16 pilots, (A) channel estimation under 8 pilots and (B) channel estimation under 16 pilots.

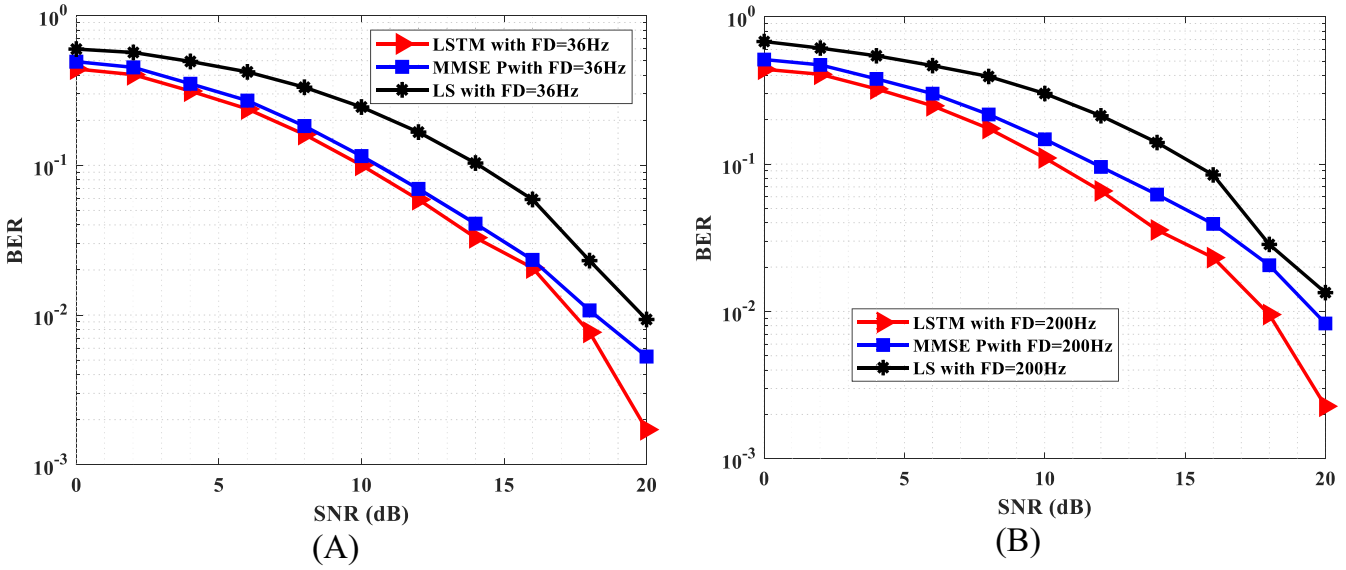

FIGURE 13 Bit error rate (BER) performance under different Doppler frequency F_d, (A) BER performance under $F_d = 36Hz$ and (B) BER performance under $F_d = 200Hz$.

the loss in a MIMO-OFDM model. The existing method consumes high loss due to a lack of optimizing capability during channel estimation. At SNR = 20 dB and $F_d = 36Hz$, the existing MMSE, LS,[39] and the proposed LSTM attains 0.005, 0.009 and 0.001, respectively. At SNR = 20 dB and $F_d = 200Hz$, the existing MMSE, LS,[24] and the proposed LSTM attain 0.008, 0.01, and 0.002, respectively.

Figure 14 presents the MSE performance under $F_d = 36Hz$ and $F_d = 200Hz$ by considering the 64 QAM modulation. The methods like LS, LMMSE, FDNN, CNN, and LSTM are compared with the proposed model. The graph shows that the MSE vales decreases as the SNR value increases. Similarly, Figure 15 presents the BER performance under $F_d = 36Hz$ and $F_d = 200Hz$ by considering the 64 QAM modulation. Also in this comparison, the proposed model attained better BER performance.

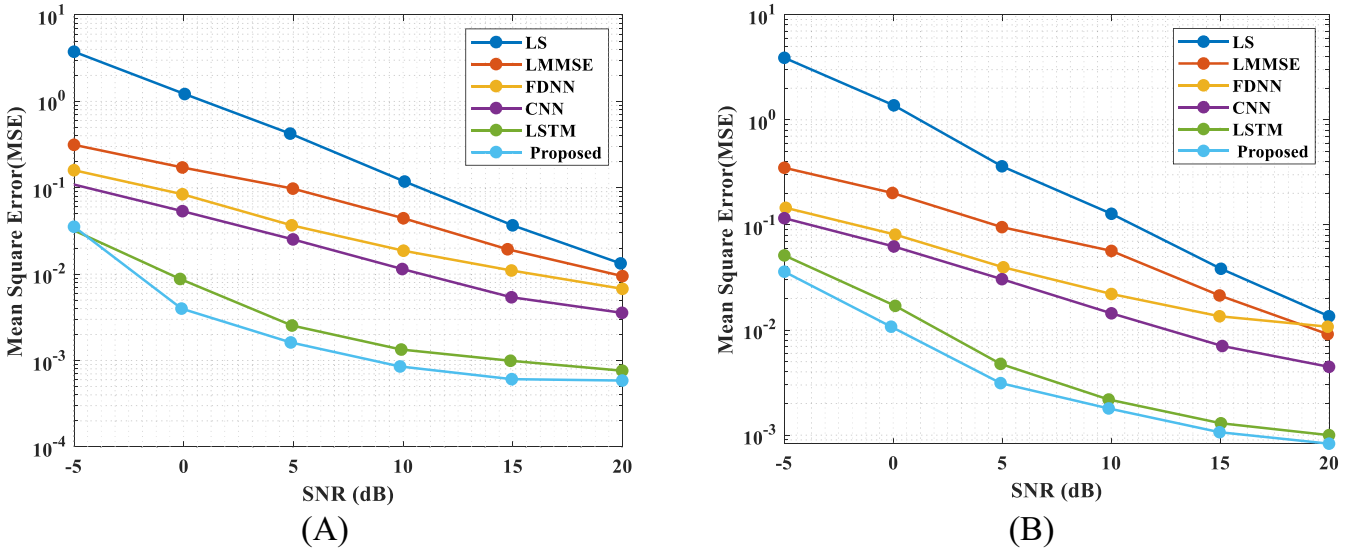

FIGURE 14 Mean square error performance under (A) $F_d = 36Hz$ and (B) $F_d = 200Hz$.

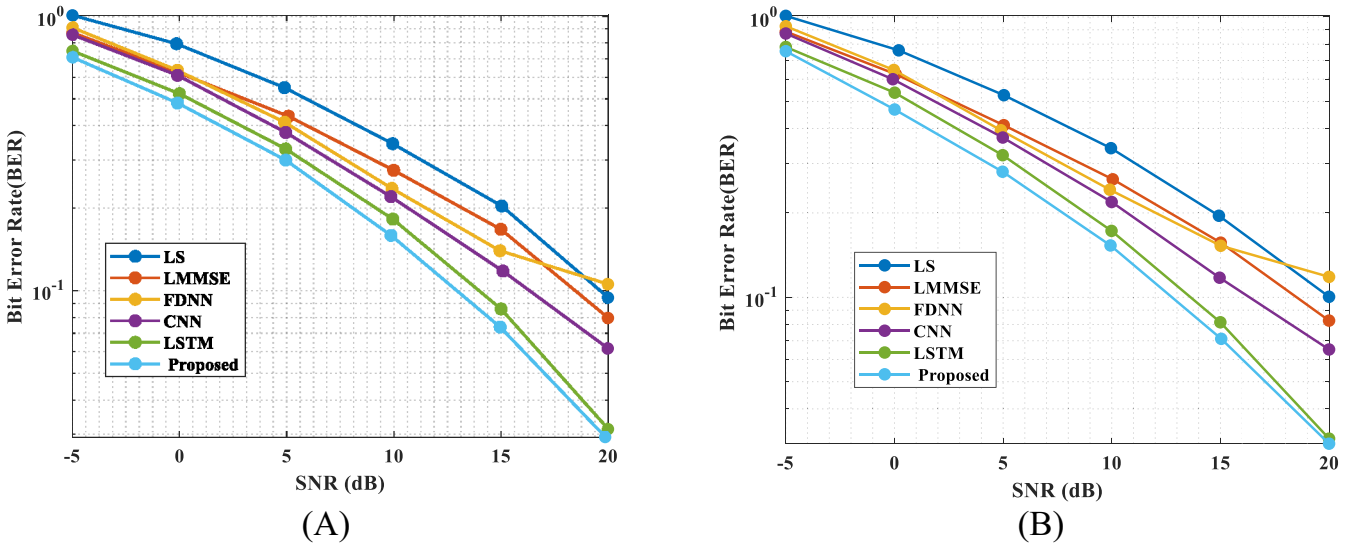

FIGURE 15 Comparison of Bit error rate performance under (A) $F_d = 36Hz$ and (B) $F_d = 200Hz$.

5 | CONCLUSION

This research proposes an optimal hyper DL-based CNN-TSO-LSTM for pilot design and CE in the downlink MIMO-OFDM system. At the initial stage, an optimized hyper CNN-based training model is proposed to find and select the pilot design accurately. Then, a DL-based LSTM model is introduced for pilot-based CE in the MIMO-OFDM model. Finally, the developed system is implemented in the MATLAB tool and analyses the performance based on BER and MSE under various existing methods. In the experimental scenario, the proposed method attains the BER of 0.20 and 0.03 for CP and without CP, respectively. In addition, the proposed method attains the MSE of 1.14, 0.99, 1.08, and 0.97 for 8, 16, 48, and 64 pilots, respectively. The proposed MIMO-OFDM model is highly used in many applications like multi-cast video, local area networks (LAN), and metropolitan area networks (MAN). In future studies, the researchers must focus on developing a deep channel estimator for higher mobility channels. Furthermore, improving the interference mitigation ability will be a better future research direction. In particular, how the proposed channel estimator outperforms interfering users with Eigen covariance matrices is not observed, and this will be addressed in the upcoming studies.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

1. Unnisa N, Tatineni M. Adaptive deep learning strategy with red deer algorithm for sparse channel estimation and hybrid pre-coding in millimeter wave massive MIMO-OFDM systems. *Wirel Pers Commun*. 2022;122(4):3019-3051.
2. Mashhadi MB, Gündüz D. Pruning the pilots: deep learning-based pilot design and channel estimation for MIMO-OFDM systems. *IEEE Trans Wireless Commun*. 2021;20(10):6315-6328.
3. Bi Y, Zhang J, Zeng M, Xu X. Channel modeling and estimation for OFDM systems in high-speed trains scenarios, in Proc. *IEEE 83*rd *Veh. Technol. Conf*. (VTC Spring), Nanjing, China, *May*, 1–6. 2016.
4. Ghazal A, Yuan Y, Wang C-X, et al. A non-stationary IMT-advanced MIMO channel model for high-mobility wireless communication systems. *IEEE Trans Wireless Commun*. 2017;16(4):2057-2068.
5. Elnakeeb A, Mitra U. Bilinear channel estimation for MIMO OFDM: lower bounds and training sequence optimization. *IEEE Trans Signal Process*. 2021;69:1317-1331.
6. Ragunathan S, Perumal D. Enhancement of energy efficiency in massive MIMO network using superimposed pilots. *J Ambient Intell Human Comput*. 2021;12(7):7391-7398.
7. Na Z, Pan Z, Xiong M, Xia J, Lu W. Soft decision control iterative channel estimation for the internet of things in 5G networks. *IEEE Internet Things J*. 2018;6(4):5990-5998.
8. Ijiga OE, Ogundile OO, Familua AD, Versfeld DJ. Review of channel estimation for candidate waveforms of next generation networks. *Electronics*. 2019;8(9):956.
9. Nie Y, Yu X, Yang Z. Deterministic pilot pattern allocation optimization for sparse channel estimation based on CS theory in OFDM system. *EURASIP J Wireless Commun Netw*. 2019;2019(1):1-8.
10. Riadi A, Boulouird M, Hassani MMR. Least squares channel estimation of an OFDM massive MIMO system for 5G wireless communications. In International Conference on the Sciences of Electronics, *Technologies of Information and Telecommunications Springer, Cham*, 440–450. 2018.
11. Ranjan A, Singh AK, Sahana BC. A review on deep learning-based channel estimation scheme. In: Pant M, Kumar Sharma T, Arya R, Sahana B, Zolfagharinia H, eds. *Soft Computing: Theories and Applications. Advances in Intelligent Systems and Computing*, Springer: 2020;1154:1007-1016. https://doi.org/10.1007/978-981-15-4032-5_90
12. Cai J, He X, Wang H, Song R. Deterministic pilot design for structured sparse channel estimation in MISO systems. *Wireless Netw*. 2020;26(4):2609-2621.
13. Trinh QK, Tran THT, Dang TB, Nguyen VT. An optimized pilot-Assisted Channel estimation method for low-dispersive channels. *International Conference on Industrial Networks and Intelligent Systems*. Springer; 2021:48-56.
14. Albataineh Z, Hayajneh K, Salameh HB, Dang C, Dagmseh A. Robust massive MIMO channel estimation for 5G networks using compressive sensing technique. *AEU Int J Electron Commun*. 2020;120:153197.
15. Li T, Noels N, Steendam H. Efficient pilot allocation for sparse channel estimation in UWB OFDM systems. *Signal Process*. 2020;175:107666.
16. Singh H, Bansal S. Channel estimation with ISFLA based pilot pattern optimization for MIMO OFDM system. *AEU Int J Electron Commun*. 2017;81:143-149.
17. Raslan WA, Mohamed MA, Abdel-Atty HM. Deep-BiGRU based channel estimation scheme for MIMO–FBMC systems. *Phys Commun*. 2022;51:101592.
18. Wu X, Huang Z, Ji Y. Deep neural network method for channel estimation in visible light communication. *Optics Commun*. 2020;462:125272.
19. Amirabadi MA, Kahaei MH, Nezamalhosseini SA, Vakili VT. Deep learning for channel estimation in FSO communication system. *Optics Commun*. 2020;459:124989.
20. Yu F, Song L, Lei X, Xiao Y, Jiang ZX, Jin M. Optimal power allocation for SM-OFDM systems with imperfect channel estimation. *Chaos Solitons Fractals*. 2016;89:263-269.
21. Shalavi N, Atashbar M, Feghhi MM. Downlink channel estimation of FDD based massive MIMO using spatial partial-common sparsity modeling. *Physical Commun*. 2020;42:101138.
22. Sun J, Mu X, Kong D, Wang Q, Li X, Cheng X. Channel estimation approach with low pilot overhead in FBMC/OQAM systems. *Wireless Commun Mobile Comput*. 2021;2021:1-9.
23. Nandi S, Nandi A, Pathak NN. Channel estimation of massive MIMO-OFDM system using Elman recurrent neural network. *Arab J Sci Eng*. 2022;47(8):9755-9765.
24. Le HA, Chien TV, Nguyen TH, Choo H, Nguyen VD. Machine learning-based 5G-and-beyond channel estimation for MIMO-OFDM communication systems. *Sensors*. 2021;21(14):4861.
25. Ponnaluru S, Penke S. Deep learning for estimating the channel in orthogonal frequency division multiplexing systems. *J Amb Intell Human Comput*. 2021;12(5):5325-5336.

26. Liao Y, Hua Y, Cai Y. Deep learning based channel estimation algorithm for fast time-varying MIMO-OFDM systems. *IEEE Commun Lett*. 2019;24(3):572-576.
27. Jiang P, Wen CK, Jin S, Li GY. Dual CNN-based channel estimation for MIMO-OFDM systems. *IEEE Trans Commun*. 2021;69(9):5859-5872.
28. Vidhya K, Shankar Kumar KR. Channel estimation of MIMO–OFDM system using PSO and GA. *Arab J Sci Eng*. 2014;39:4047-4056.
29. Seyman MN, Taspinar N. MIMO-OFDM channel estimation using ANFIS. *Elektron Elektrotech*. 2012;120(4):75-78.
30. Seyman MN, Taşpinar N. Particle swarm optimization for pilot tones design in MIMO-OFDM systems. *EURASIP J Adv Signal Process*. 2011;2011(1):1.
31. Seyman MN, Taşpinar N. Optimization of pilot tones using differential evolution algorithm in MIMO-OFDM systems. *Turk J Elec Eng Comp Sci*. 2012;20(1):15-23.
32. Seyman MN, Taşpınar N. Pilot tones optimization using artificial bee colony algorithm for MIMO–OFDM systems. *Wireless Personal Commun*. 2013;71:151-163.
33. Seyman MN, Taspinar N. Radial basis function neural networks for channel estimation in MIMO-OFDM systems. *Arab J Sci Eng*. 2013 Aug;38:2173-2178.
34. Cheng NH, Huang KC, Chen YF, Tseng SM. Maximum likelihood-based adaptive iteration algorithm design for joint CFO and channel estimation in MIMO-OFDM systems. *EURASIP J Adv Signal Process*. 2021;2021:1-21.
35. Ge X, Sun Y, Gharavi H, Thompson J. Joint optimization of computation and communication power in multi-user massive MIMO systems. *IEEE Trans Wireless Commun*. 2018;17(6):4051-4063.
36. Houssein EH, El-Din Helmy B, Elngar AA, Abdelminaam DS, Shaban H. An improved tunicate swarm algorithm for global optimization and image segmentation. *IEEE Access*. 2021;9:56066-56092.
37. Farzamnia A, Hlaing NW, Haldar MK, Rahebi J. Channel estimation for sparse channel OFDM systems using least square and minimum mean square error techniques. *In 2017 International Conference on Engineering and Technology (ICET)*. IEEE; 2017:1-5.
38. Dong F, LiuJ HL, Hu X, Liu H. Channel estimation based on extreme learning machine for high speed environments. *Proceedings of ELM-2015*. Springer; 2016:159-167.
39. Kang X-F, Liu Z-H, Yao M. Deep learning for joint pilot design and channel estimation in MIMO-OFDM systems. *Sensors*. 2022;22(11):4188.

How to cite this article: Silpa C, Vani A, Naidu KR. Optimized deep learning based hypernet convolution neural network and long short term memory for joint pilot design and channel estimation in MIMO-OFDM model. *Trans Emerging Tel Tech*. 2023;e4925. doi: 10.1002/ett.4925