RESEARCH ARTICLE

Deep Learning-Based Channel Estimation Method for MIMO Systems in Spatially Correlated Channels

SANGGEUN LEE[ID]**1 AND DONGKYU SIM**[ID]**2, (Member, IEEE)**

1 Samsung Electronics Company Ltd., Suwon, Gyeonggi 16677, South Korea
2 School of Information and Communication Engineering, Chungbuk National University, Cheongju, Chungbuk 28644, South Korea

Corresponding author: Dongkyu Sim (dongkyu.sim@chungbuk.ac.kr)

ABSTRACT This paper proposes a deep learning-based channel estimation method for multiple-input multiple-output (MIMO) systems in spatially correlated channels. To reduce the pilot overhead of pilot symbol-assisted channel estimation, the proposed method utilizes fewer pilot symbols than the number of transmit antennas. Firstly, based on pilot symbols, the estimated partial MIMO channel matrix, consisting of the partial coefficients of the MIMO channel matrix, is obtained by the linear minimum mean square error algorithm. After that, a deep neural network uses the estimated partial MIMO channel matrix as an input and we have the predicted MIMO channel matrix, that corresponds to the channel state information not transmitting pilot symbols. Finally, by aggregating the estimated partial MIMO channel matrix and the predicted MIMO channel matrix, the proposed method can acquire the reconstructed MIMO channel matrix. In simulation results, to show the validity of the proposed method, various performances of the proposed and conventional channel estimation methods were evaluated. Simulation results show that even though the proposed method does not send the pilot symbols for all transmit antennas, it can achieve almost the same bit error rate and improved throughput performances compared with the conventional channel estimation method.

INDEX TERMS Channel estimation method, deep learning, MIMO system, spatially correlated channel.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) systems can provide improved data rate and reliability performances by utilizing diversity in MIMO channels [1], [2], [3]. Especially, ultra-massive MIMO systems, which have emerged in 6G wireless systems, are considered a key technology to achieve more than terabits/second peak data rate [4], [5], [6].

To exploit the advantages of MIMO systems, channel state information (CSI) is essential. Many wireless communication standards adopt pilot symbol-assisted channel estimation to acquire CSI because it can guarantee high estimation accuracy with simple implementation [7], [8], [9], [10], [11]. However, the channel estimation performance

The associate editor coordinating the review of this manuscript and approving it for publication was Parul Garg.

of massive MIMO systems highly depends on the number of pilot symbols and is limited due to pilot contamination. Furthermore, the pilot symbol-assisted channel estimation based on the least square (LS) or linear minimum mean square error (LMMSE) algorithm requires at least the number of pilot symbols same as the number of transmit antennas, and it would be a big burden in ultra-massive MIMO systems [2], [3], [12], [13], [14], [15]. Meanwhile, in vehicle-to-everything (V2X) communications that support various applications based on vehicles and transportation infrastructures such as road safety, traffic efficiency, and autonomous driving, the performances of pilot symbol-assisted channel estimation are degraded due to the high mobility of vehicles [16], [17]. As a result, to maintain the accuracy of pilot symbol-assisted channel estimation, more pilot symbols or advanced techniques such as data-aided schemes and

orthogonal time frequency space (OTFS) modulations are required and these lead to burdensome in terms of resource utilization and computational complexity [18], [19], [20].

In recent years, deep learning has been researched for various applications of communication systems such as detection, localization, and sensing [21], [22], [23], [24], [25]. Specifically, many studies on deep learning-based channel estimation have been carried out because the advantages of deep learning can be used to overcome the problems of the conventional channel estimation method [13], [14], [15], [26], [27], [28], [29], [30], [31], [32], [33], [34]. In [13], a deep learning-based two-stage channel estimation scheme consisting of pilot-aided channel estimation and data-aided iterative channel estimation was proposed for massive MIMO systems. The authors of [14] proposed a deep learning-based channel estimation method that does not require any training and is robust to pilot contamination by utilizing a deep image prior network. The end-to-end deep neural network (DNN) architecture, which jointly designs the pilot signals and channel estimator, was introduced in [15]. In [26] and [27], deep learning-based channel estimation algorithms were proposed for doubly selective fading channels and vehicular communications which are challenging environments for channel estimation. Moreover, the channel prediction methods for massive MIMO systems and deep learning-based algorithms for OTFS modulation were investigated in high mobility scenarios [28], [29], [30]. To reduce the overwhelming overheads required for downlink training and uplink feedback, the channel prediction algorithm based on the sparse complex-valued neural networks was introduced in [31]. The authors of [32] and [33] analyzed the performances of deep learning-based channel estimation. In [34], the deep channel prediction using recurrent neural networks, which can reduce the number of pilot symbols by learning channel variations in time-varying fading channels, was presented. Various works have been conducted for deep learning-based channel estimation so far, but studies to reduce pilot overhead in MIMO systems have not been sufficient. Especially, considering that next-generation communication systems will utilize a very large number of antennas, which causes spatial correlation, research on reducing resource overhead for MIMO systems in spatially correlated channels is essential [35], [36].

In this paper, we propose a deep learning-based channel estimation method to reduce the pilot overhead of pilot symbol-assisted channel estimation in spatially correlated MIMO channels. To do this, the proposed method sends the pilot symbols for the predetermined transmit antenna set, whose cardinality is smaller than the number of transmit antennas. Firstly, the estimated partial MIMO channel matrix, which corresponds to the CSI of the predetermined transmit antenna set, is acquired by pilot symbols and the LMMSE algorithm. The CSI of transmit antennas not sending pilot symbols, called the predicted MIMO channel matrix, is obtained by a DNN whose input is the estimated partial MIMO channel matrix. Finally, aggregating two results, the estimated partial MIMO channel matrix and the predicted

MIMO channel matrix, we have the reconstructed MIMO channel matrix. Various performances of the conventional and proposed channel estimation methods are compared in simulation results. Since the proposed method uses fewer pilot symbols than the conventional methods, it shows the degraded normalized mean square error (NMSE) performances. However, the throughput performance of the proposed method outperforms that of the conventional method by maintaining the bit error rate (BER) performance.

II. SYSTEM MODEL

In this paper, we consider massive MIMO systems that have N_T transmit and N_R receive antennas, and a uniform planar array (UPA) is adopted to deploy antennas, efficiently [37]. To overcome multipath fading channels, orthogonal frequency division multiplexing with cyclic prefix (CP-OFDM) is applied. Assuming the perfect synchronization at the receiver side, the received signal vector $y(k, n) = \left[y_1(k, n) \cdots y_{n_r}(k, n) \cdots y_{N_R}(k, n)\right]^T$, associated with kth subcarrier and nth symbol interval, can be expressed as

$$\mathbf{y}(k, n) = \sqrt{\gamma}\mathbf{H}(k, n) s(k, n) + v(k, n), \quad (1)$$

where γ is the transmit power, $\mathbf{H}(k, n) = [\mathbf{h}_1(k, n) \cdots \mathbf{h}_{n_t}(k, n) \cdots \mathbf{h}_{N_T}(k, n)]$ denotes an $N_R \times N_T$ MIMO channel matrix in the frequency domain, $\mathbf{h}_{n_t}(k, n) = \left[h_{1,n_t}(k, n) \cdots h_{n_r,n_t}(k, n) \cdots h_{N_R,n_t}(k, n)\right]^T$ represents an $N_R \times 1$ channel vector for the n_tth transmit antenna, $h_{n_r,n_t}(k, n)$ is a channel frequency response (CFR) between the n_tth transmit antenna and n_rth receive antenna, $s(k, n) = \left[s_1(k, n) \cdots s_{n_t}(k, n) \cdots s_{N_T}(k, n)\right]^T$ is an $N_T \times 1$ transmitted signal vector with $\mathbb{E}\left[s(k, n) s^H(k, n)\right] = \mathbf{I}_{N_T}$, $\mathbf{v}(k, n) = \left[v_1(k, n) \cdots v_{n_r}(k, n) \cdots v_{N_R}(k, n)\right]^T$ is an $N_R \times 1$ additive white Gaussian noise vector with $\mathbb{E}\left[v(k, n) \mathbf{v}^H(k, n)\right] = \sigma_v^2 \mathbf{I}_{N_R}$, $\mathbb{E}[\cdot]$ stand for expectation, $(\cdot)^H$ represents Hermitian operator, and $\mathbf{I}_K$ is the identity matrix of size K [1], [8]. In spatially correlated channels, $\mathbf{H}(k, n)$ can be modeled by the one-ring model, which is given by

$$\mathbf{H}(k, n) = \mathbf{H}_w(k, n) \mathbf{R}_t^{\frac{1}{2}}, \quad (2)$$

where $\mathbf{H}_w(k, n)$ is the spatially white MIMO channel matrix whose entries are independent and identically distributed zero-mean circularly symmetric complex Gaussian random variables with unit variance and $\mathbf{R}_t$ is a transmit correlation matrix of UPA [1], [3], [38]. Defining the number of antennas for UPA in the horizontal and vertical direction as $N_{T,H}$ and $N_{T,V}$, $\mathbf{R}_t$ can be represented as

$$\mathbf{R}_t = \mathbf{R}\left(N_{T,H}\right) \otimes \mathbf{R}\left(N_{T,V}\right), \quad (3)$$

where $\mathbf{R}(N)$ is the $N \times N$ transmit correlation matrix of a uniform linear array, the (u, v) entry of $\mathbf{R}(N)$ is given by

$$\{\mathbf{R}(N)\}_{u,v} = \begin{cases} \rho^{|u-v|} & for\ u \leq v, \\ (\rho^*)^{|u-v|} & \text{otherwise,} \end{cases} \quad (4)$$

$\rho(|\rho| \leq 1)$ is the correlation coefficient between adjacent antennas, and $\otimes$ denotes the Kronecker product [39], [40].

To obtain the CSI, we adopt pilot symbol-assisted channel estimation. In MIMO-OFDM systems, N_T pilot symbols are transmitted through different subcarriers and symbol intervals to estimate MIMO channels. We define N_T pilot symbols as a pilot group and it is transmitted N_L times, repeatedly [9], [41]. Assuming the pilot symbol associated with the n_tth transmit antenna for the lth pilot group as $s_{n_t}^l$, the received signal vector for $s_{n_t}^l$ can be expressed as

$$\mathbf{y}\left(k_{n_t}^l, n_{n_t}^l\right) = \sqrt{\gamma}\mathbf{h}_{n_t}\left(k_{n_t}^l, n_{n_t}^l\right)s_{n_t}^l + \mathbf{v}(k_{n_t}^l, n_{n_t}^l), \quad (5)$$

where l $(1 \le l \le N_L)$ is an integer number and $\left(k_{n_t}^l, n_{n_t}^l\right)$ is the position of the pilot symbol associated with the n_tth transmit antenna for the lth pilot group. To estimate CFR between the n_tth transmit antenna and n_rth receive antenna, $h_{n_r,n_t}\left(k_{n_t}^l, n_{n_t}^l\right)$, for all pilot groups, we define $\mathcal{Y}_{n_r,n_t} = \left[y_{n_r}\left(k_{n_t}^1, n_{n_t}^1\right) \cdots y_{n_r}\left(k_{n_t}^{N_L}, n_{n_t}^{N_L}\right)\right]^T$, which is given by

$$\mathcal{Y}_{n_r,n_t} = \sqrt{\gamma}\mathbf{S}_{n_t}\mathcal{H}_{n_r,n_t} + \mathcal{V}_{n_t,n_r}, \quad (6)$$

where

$$\mathbf{S}_{n_t} = \text{diag}\left(s_{n_t}^1, \cdots, s_{n_t}^{N_L}\right), \quad (7)$$

$$\mathcal{H}_{n_r,n_t} = \left[h_{n_r,n_t}\left(k_{n_t}^1, n_{n_t}^1\right) \cdots h_{n_r,n_t}\left(k_{n_t}^{N_L}, n_{n_t}^{N_L}\right)\right]^T, \quad (8)$$

$$\mathcal{V}_{n_r,n_t} = \left[v_{n_r}\left(k_{n_t}^1, n_{n_t}^1\right) \cdots v_{n_r}\left(k_{n_t}^{N_L}, n_{n_t}^{N_L}\right)\right]^T. \quad (9)$$

Note that $\mathcal{Y}_{n_r,n_t}$ is obtained by stacking $y_{n_r}\left(k_{n_t}^1, n_{n_t}^1\right)$ to $y_{n_r}\left(k_{n_t}^{N_L}, n_{n_t}^{N_L}\right)$ and defined for all transmit and receive antenna pair [42]. For estimations, LS or LMMSE algorithm can be applied and each estimation result is given by

$$\hat{\mathcal{H}}_{n_r,n_t}^{\text{LS}} = \frac{1}{\sqrt{\gamma}}\mathbf{S}_{n_t}^{-1}\mathcal{Y}_{n_r,n_t}, \quad (10)$$

$$\hat{\mathcal{H}}_{n_r,n_t}^{\text{MMSE}} = \Xi_{n_r,n_t}\left(\Xi_{n_r,n_t} + \frac{1}{\gamma}\mathbf{I}_{N_L}\right)^{-1}\hat{\mathcal{H}}_{n_r,n_t}^{\text{LS}}, \quad (11)$$

where $\Xi_{n_r,n_t} = \mathbb{E}\left[\mathcal{H}_{n_r,n_t}\mathcal{H}_{n_r,n_t}^H\right]$ and $(\cdot)^{-1}$ denotes matrix inversion [41], [43]. Finally, by utilizing an interpolation method, we can obtain an estimated MIMO channel matrix $\hat{\mathbf{H}}(k, n)$ for all subcarriers and symbol intervals, and the (n_r, n_t) entry of $\hat{\mathbf{H}}(k, n)$ is $\hat{h}_{n_r,n_t}(k, n)$ [44]. Based on $\hat{\mathbf{H}}(k, n)$, the decoder matrix $\mathbf{G}(k, n)$ for symbol detection can be designed and the post-processing signal vector $\mathbf{z}(k, n)$ is represented by

$$\mathbf{z}(k, n) = \mathbf{G}(k, n)\mathbf{y}(k, n). \quad (12)$$

To evaluate the performance of systems, we consider the throughput performance T, the amount of successfully transmitted data per unit time [45]. Defining the number of pilot symbols as $N_{\text{ps}} = N_L N_T$, data symbols as $N_{\text{ds}} = N_{\text{ts}} - N_{\text{ps}}$, and N_{ts} as total symbols, the throughput performance T during a specific time period τ can be expressed as

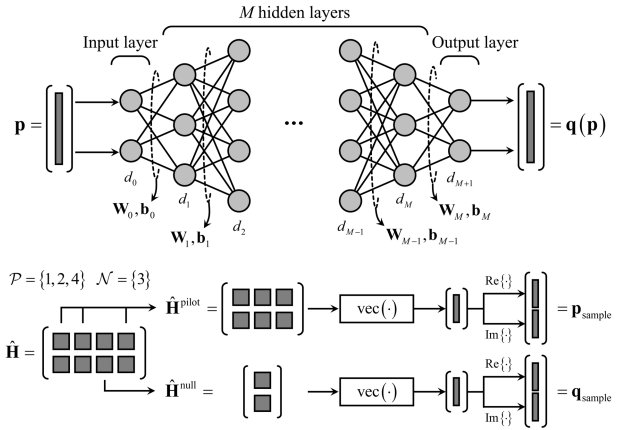$$T = \frac{\log_2 M}{\tau}N_{\text{ds}}\left(1 - P_{B,\text{avg}}\right)$$

FIGURE 1. An architecture and input-output relationship of the proposed DNN-based channel predictor $\mathcal{D}$.

$$= \frac{\log_2 M}{\tau}\left(N_{\text{ts}} - N_{\text{ps}}\right)\left(1 - P_{B,\text{avg}}\right)$$
$$= \left(T_0 - T_{\text{ps}}\right)\left(1 - P_{B,\text{avg}}\right), \quad (13)$$

where M is the modulation order, $P_{B,\text{avg}}$ denotes the average bit error probability, $T_0\left(= \frac{\log_2 M}{\tau}N_{\text{ts}}\right)$ is the total number of bits that can be transmitted during τ, and $T_{\text{ps}}\left(= \frac{\log_2 M}{\tau}N_{\text{ps}}\right)$ is the number of bits lost due to the transmission of the pilot symbols during τ [46], [47].

III. PROPOSED DEEP LEARNING-BASED CHANNEL ESTIMATION METHOD

In this section, we present the deep learning-based channel estimation method to reduce the overhead of pilot symbol-assisted channel estimation in spatially correlated MIMO channels. To do this, the proposed method utilizes fewer pilot symbols than the number of transmit antennas, i.e., $N_T^{\text{pilot}} < N_T$. We define $\mathcal{P}$ and $\mathcal{N}$ as a set containing indices of transmit antennas that send pilot symbols and not, respectively. Then, $\mathcal{U} = \mathcal{P} \cup \mathcal{N}$ is a set collecting all indices of transmit antennas. Note that the cardinalities of $\mathcal{P}$ and $\mathcal{N}$ are $|\mathcal{P}| = N_T^{\text{pilot}}$ and $|\mathcal{N}| = N_T - N_T^{\text{pilot}} = N_T^{\text{null}}$. From now on, we drop (k, n), the index of the subcarrier and symbol interval, for notation simplicity.

The proposed method first performs pilot symbol-assisted channel estimation for the predetermined transmit antenna set, $\mathcal{P}$, and we can obtain an $N_R \times N_T^{\text{pilot}}$ estimated partial MIMO channel matrix $\hat{\mathbf{H}}^{\text{pilot}}$ whose ith column vector $\left[\hat{\mathbf{H}}^{\text{pilot}}\right]_i$ is given by

$$\left[\hat{\mathbf{H}}^{\text{pilot}}\right]_i = \hat{\mathbf{h}}_\mathcal{P}(i), \quad (14)$$

where $i = 1, 2, \cdots, |\mathcal{P}|\left(= N_T^{\text{pilot}}\right)$ and $\mathcal{A}(i)$ represents the ith element of a set $\mathcal{A}$. Note that we have $\hat{\mathbf{H}}^{\text{pilot}}$ for all subcarriers and symbol intervals by applying an interpolation method. However, defining $N_R \times N_T^{\text{null}}$ matrix $\hat{\mathbf{H}}^{\text{null}}$ as

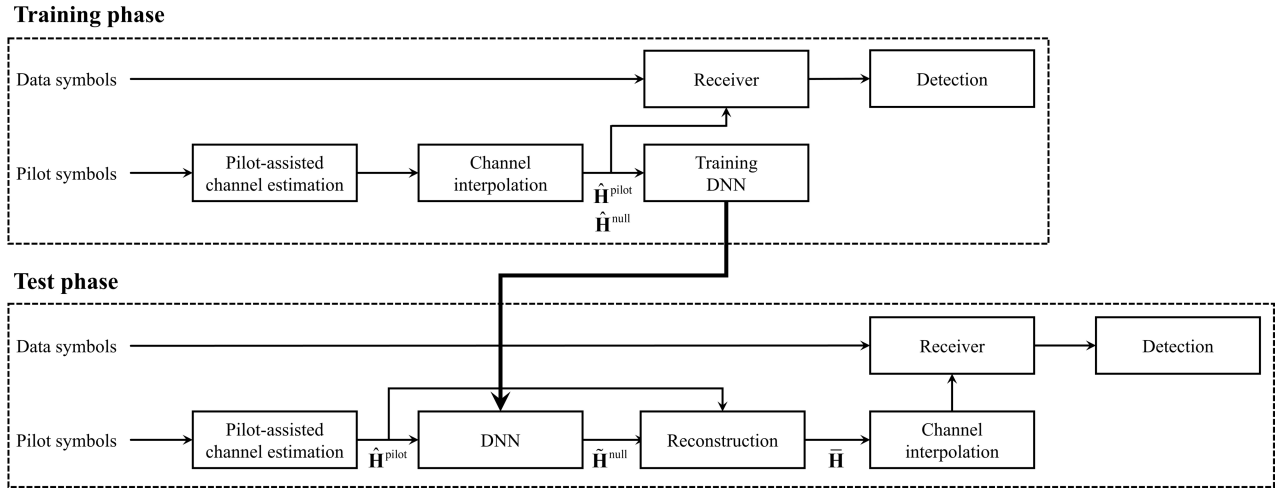$$\left[\hat{\mathbf{H}}^{\text{null}}\right]_i = \hat{\mathbf{h}}_\mathcal{N}(i), \quad (15)$$

FIGURE 2. The block diagram of the proposed channel estimation method.

where $i = 1, 2, \cdots, |\mathcal{N}| \left(= N_T^{\text{null}}\right)$, we cannot acquire the estimated information of $\hat{\mathbf{H}}^{\text{null}}$ because transmit antennas corresponding $\mathcal{N}$ do not send pilot symbols. Considering the spatial correlation of MIMO channel matrix $\mathbf{H}$, a mathematical relationship between $\mathbf{H}^{\text{pilot}} = \left[\mathbf{h}_{\mathcal{P}(1)} \cdots \mathbf{h}_{\mathcal{P}\left(N_T^{\text{pilot}}\right)}\right]$ and $\mathbf{H}^{\text{null}} = \left[\mathbf{h}_{\mathcal{N}(1)} \cdots \mathbf{h}_{\mathcal{N}\left(N_T^{\text{null}}\right)}\right]$ exists. To figure out this relationship, we utilize a DNN for data modeling which is widely used as a universal feedforward approximator [48]. The objective of a DNN is to predict $\mathbf{H}^{\text{null}}$ without pilot symbols by minimizing $\mathbb{E}\left\{\left\|\mathbf{H}^{\text{null}} - f\left(\hat{\mathbf{H}}^{\text{pilot}}\right)\right\|_F^2\right\}$ where $\|\cdot\|_F^2$ is the Frobenius norm.

Based on the above discussions, we propose the DNN-based channel predictor $\mathcal{D}$. As depicted in Fig. 1, it consists of M hidden layers and we use the rectified linear unit (ReLU) activation function $\phi(x) = \max\{0, x\}$ for each hidden layer. In Fig. 1, $d_0, \cdots, d_{M+1}$ represent the number of nodes of each layer. Then, the depth, width, and size of the DNN are $M + 1$, $\max\{d_1, \cdots, d_M\}$, and $\sum_{m=1}^{M} d_m$, respectively. Note that d_0 and d_{M+1} are the number of nodes of the input and output layer, respectively. Defining $\mathbf{W}_m \in \mathbb{R}^{d_{m+1} \times d_m}$ and $b_m \in \mathbb{R}^{d_{m+1}}$ for $m = 0, 1, \cdots, M$ as the weight matrix and bias vector for the DNN, the output of the mth layer is represented by

$$\mathbf{x}_m = \phi\left(\mathbf{W}_{m-1}\mathbf{x}_{m-1} + \mathbf{b}_{m-1}\right). \quad (16)$$

Note that (16) only holds for $m = 1, \cdots, M$, $\mathbf{x}_0 = \mathbf{p} \in \mathbb{R}^{d_0}$ is an input vector, and $\mathbf{W}_M \mathbf{x}_M + \mathbf{b}_M = \mathbf{q}(\mathbf{p}) \in \mathbb{R}^{d_{M+1}}$ is a hypothesis for input vector $\mathbf{p}$. For a fully-connected layer, the number of floating-point operations (FLOPs) for mth layer is given by $2d_{m-1}d_m$ [49]. Therefore, by summing the number of FLOPs for all layers, the computational complexity of the proposed DNN can be calculated as

$$2\sum_{m=1}^{M+1} d_{m-1}d_m. \quad (17)$$

Denoting the input and output data sample vectors for training the DNN as $\mathbf{p}_{\text{sample}}$ and $\mathbf{q}_{\text{sample}}$, respectively, weight matrices and bias vectors are updated to decrease the loss function for $\mathbf{q}_{\text{sample}}$ and $\mathbf{q}\left(\mathbf{p}_{\text{sample}}\right)$. In the proposed method, we use MSE for the loss function, which is commonly used for regression, and it is given by [50]

$$\mathcal{L} = \mathbb{E}\left\{\left\|\mathbf{q}_{\text{sample}} - \mathbf{q}\left(\mathbf{p}_{\text{sample}}\right)\right\|_F^2\right\}. \quad (18)$$

Considering the objective of a DNN-based channel predictor $\mathcal{D}$, $\mathbf{p}_{\text{sample}}$ and $\mathbf{q}_{\text{sample}}$ are given by

$$\mathbf{p}_{\text{sample}} = \left[\text{Re}\left\{\left(\text{vec}\left(\hat{\mathbf{H}}^{\text{pilot}}\right)\right)^T\right\}, \text{Im}\left\{\left(\text{vec}\left(\hat{\mathbf{H}}^{\text{pilot}}\right)\right)^T\right\}\right]^T, \quad (19)$$

$$\mathbf{q}_{\text{sample}} = \left[\text{Re}\left\{\left(\text{vec}\left(\mathbf{H}^{\text{null}}\right)\right)^T\right\}, \text{Im}\left\{\left(\text{vec}\left(\mathbf{H}^{\text{null}}\right)\right)^T\right\}\right]^T, \quad (20)$$

where $d_0 = 2N_R N_T^{\text{pilot}}$ and $d_{M+1} = 2N_R N_T^{\text{null}}$. Note that the input and output vectors of a DNN are made by vectorizing and stacking the real and imaginary parts of a MIMO channel matrix.

With the architecture and input-output relationship of the DNN-based channel predictor $\mathcal{D}$, the proposed channel estimation operates in two phases, the training phase and the test phase. As described in Fig. 2, in the training phase, both $\hat{\mathbf{H}}^{\text{pilot}}$ and $\hat{\mathbf{H}}^{\text{null}}$ are required to train the DNN. Therefore, the transmitter has to send N_T pilot symbols, corresponding to all transmit antennas. At the receiver side, by utilizing the LS or LMMSE algorithm and interpolation method, $\bar{\mathbf{H}}$ associated with all subcarriers and symbol intervals can be obtained and used as the input and output data sample vectors, p_{sample} and q_{sample}, for the DNN training. Note that $\mathbf{H}^{\text{null}}$ in (20) is not available because it is the real MIMO channel

TABLE 1. System parameters for simulations.

Parameters	Values
Antenna configuration, (N_T, N_R)	(64, 2), (128, 2)
Number of pilot symbols, N_T^{pilot}	$r = 0.875$: 56, 112, for $N_T = 64, 128$ $r = 0.938$: 60, 120, for $N_T = 64, 128$ $r = 0.969$: 62, 124, for $N_T = 64, 128$
Modulation scheme	QAM
Modulation order	4
Correlation coefficient, ρ	0.7, 0.8, 0.9
Subcarrier spacing	15kHz
Number of subcarriers	1024
Pilot spacing in the frequency domain	12
Pilot spacing in the time domain	16
Channel model	3GPP pedestrian A
Interpolation method	Linear [44]
Precoder	DFT precoding [55]

TABLE 2. DNN parameters for simulations.

Parameters	Values
Number of nodes of input, output layers, $\{d_0, d_{M+1}\}$	$r = 0.875$: {224, 256}, {448, 512} for $N_T = 64, 128$ $r = 0.938$: {240, 256}, {480, 512} for $N_T = 64, 128$ $r = 0.969$: {248, 256}, {496, 512} for $N_T = 64, 128$
Number of nodes of hidden layers, $\{d_1, \cdots, d_M\}$	{1024, 1024, 1024}
Activation function	ReLU
Learning rate	1×10^{-5}
Number of epochs	500
Batch size	32
Optimizer	Adam [56]

matrix, unlike $\hat{\mathbf{H}}^{\text{pilot}}$ in (19). Therefore, $\hat{\mathbf{H}}^{\text{null}}$ is used as q_{sample} instead of $\mathbf{H}^{\text{null}}$. However, considering $\mathbf{H}^{\text{null}} \approx \hat{\mathbf{H}}^{\text{null}}$ with high signal-to-noise ratio (SNR) assumption, there is no effect on the training of DNN. In the test phase, after the DNN is well-trained, the transmitter exploits N_T^{pilot} pilot symbols, which correspond to the predetermined transmit antenna set, $\mathcal{P}$. Therefore, the proposed method can transmit additional data symbols at the subcarriers and symbol intervals corresponding to pilot symbol positions of $\mathcal{N}$ in the training phase. Instead, the receiver can acquire only $\hat{\mathbf{H}}^{\text{pilot}}$ by pilot symbol-assisted channel estimation, and the remaining part of $\mathbf{H}$, $\mathbf{H}^{\text{null}}$, is predicted as $\tilde{\mathbf{H}}^{\text{null}}$ by the proposed DNN-based channel predictor $\mathcal{D}$. Finally, by aggregating $\hat{\mathbf{H}}^{\text{pilot}}$ and $\hat{\mathbf{H}}^{\text{null}}$ which are obtained by the pilot symbols and DNN-based channel predictor $\mathcal{D}$, respectively, and assigning appropriate transmit and receive antenna indices, the reconstructed MIMO channel matrix $\bar{\mathbf{H}}$ can be obtained as follows:

$$\left[\bar{\mathbf{H}}\right]_{n_t} = \begin{cases} \left[\hat{\mathbf{H}}^{\text{pilot}}\right]_i & \text{if } \mathcal{P}(i) = n_t, \\ \left[\tilde{\mathbf{H}}^{\text{null}}\right]_i & \text{if } \mathcal{N}(i) = n_t. \end{cases} \tag{21}$$

Note that, using an interpolation method, we can acquire $\bar{\mathbf{H}}$ for all subcarriers and symbol intervals.

In the proposed channel estimation method, the number of transmitted pilot symbols is less than in the conventional channel estimation method and it leads to transmitting additional data symbols. To evaluate the impact of this advantage on throughput performance, we define $r = N_T^{\text{pilot}}/N_T$, the number of employed pilot symbols for the proposed method-to-that for the conventional method ratio. Then, from (13), we can rewrite the throughput performance of the proposed method with r as

$$T = \left(T_0 - rT_{\text{ps}}\right)\left(1 - P_{B,\text{avg}}\right). \tag{22}$$

Considering $0 < r < 1$, the effective number of bits that can be transmitted during τ for the proposed method, $T_0 - rT_{\text{ps}}$, is always higher than that for the conventional method, $T_0 - T_{\text{ps}}$. Moreover, assuming spatially correlated channels, the
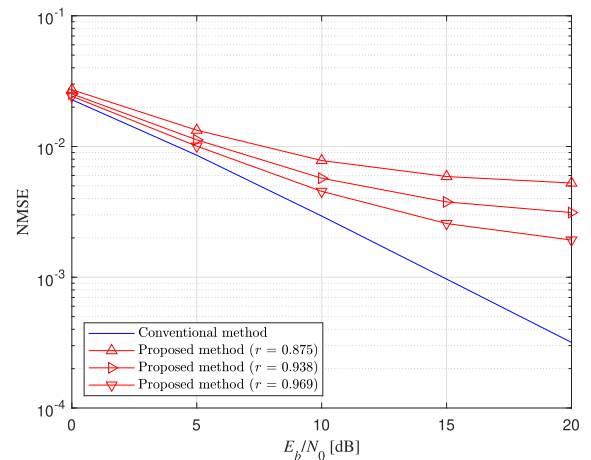

FIGURE 3. Comparison of NMSE performances ($N_T = 64$ and $N_R = 2$).

proposed method can be applied to next-generation communication systems including ultra-massive MIMO systems, high mobility, and wideband support [4], [6], [51]. However, since the DNN of the proposed method predicts the MIMO channel matrix without pilot symbols by only utilizing spatial correlation, the accuracy of channel estimation may be affected by r and ρ, which represent the number of known spatial channels and the degree of correlation between spatial channels, respectively, and the NMSE performances of the proposed method will be degraded compared to the conventional method. Therefore, performances of the proposed method according to r and ρ have to be evaluated.

IV. SIMULATION RESULTS

In this section, we evaluate the performances of the proposed channel estimation method in MIMO-OFDM systems by Monte Carlo simulations. System and DNN parameters used in simulations are described in Table 1 and 2. We use the pedestrian channel models in [52] and adopt the LMMSE algorithm and zero-forcing decoder for channel estimation and symbol detection, respectively [40], [53], [54]. For comparison, we also evaluate the performances of the conventional channel estimation method in [43]. In a conventional method, each transmit antenna sends a pilot symbol which means N_T pilot symbols are utilized. At the receiver side, an LMMSE algorithm is applied and N_T estimated
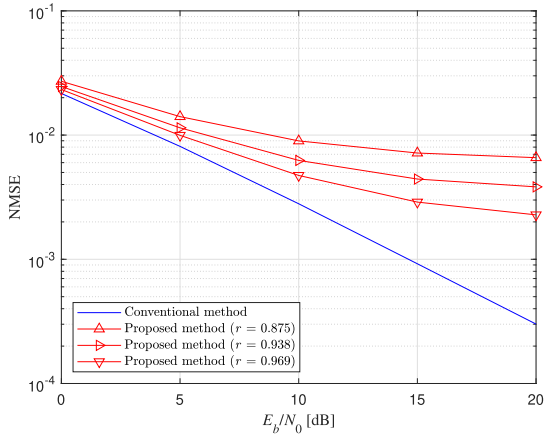
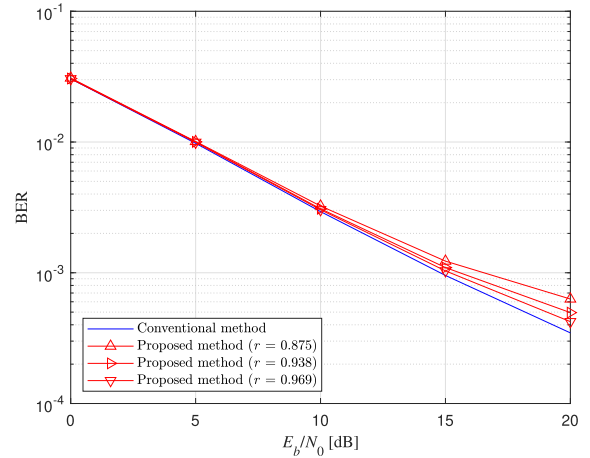**FIGURE 4.** Comparison of NMSE performances ($N_T = 128$ and $N_R = 2$).
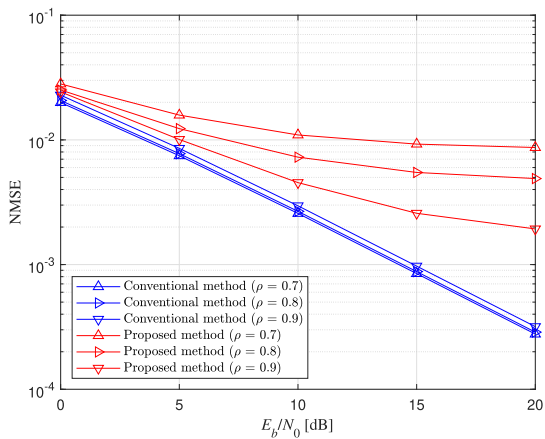

FIGURE 5. Comparison of NMSE performances according to the correlation coefficient, ρ ($N_T = 64$, $N_R = 2$ and $r = 0.969$).
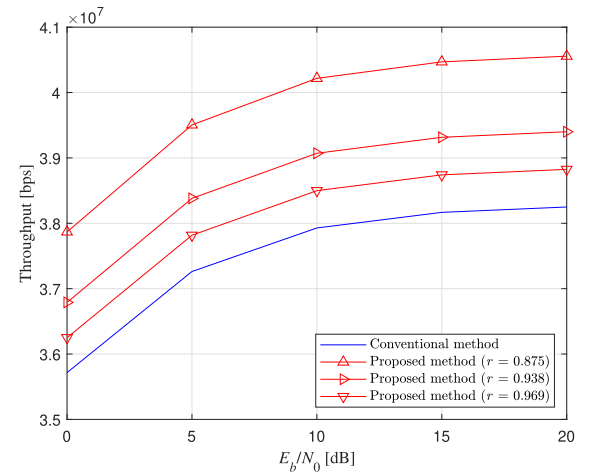

FIGURE 6. Comparison of BER performances ($N_T = 64$ and $N_R = 2$).
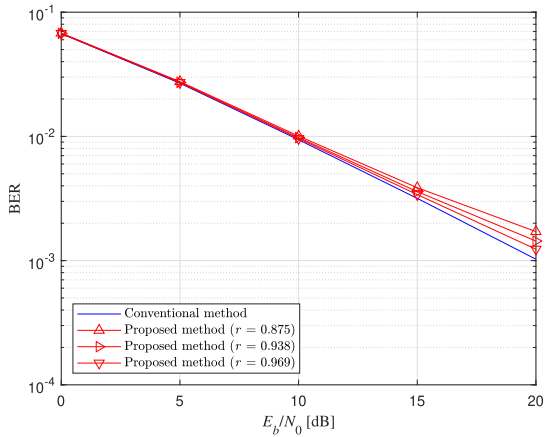

FIGURE 7. Comparison of BER performances ($N_T = 128$ and $N_R = 2$).

FIGURE 8. Comparison of throughput performances ($N_T = 64$ and $N_R = 2$).
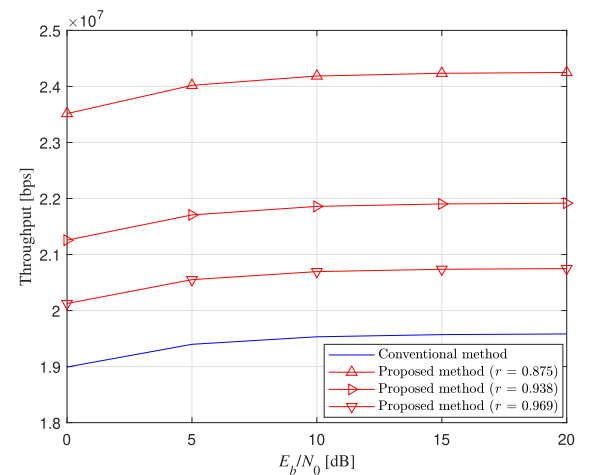

FIGURE 9. Comparison of throughput performances ($N_T = 128$ and $N_R = 2$).

channel vectors $\hat{\mathbf{h}}_{n_t}$ are obtained by N_T pilot symbols. Finally, with linear interpolation, we can acquire an estimated MIMO channel matrix $\hat{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{h}}_1 \cdots \hat{\mathbf{h}}_{n_t} \cdots \hat{\mathbf{h}}_{N_T} \end{bmatrix}$ for all subcarriers and symbol intervals.

First, the NMSE performance, defined as [57]

$$\text{NMSE} = \mathbb{E}\left[\frac{\|\mathbf{H} - \bar{\mathbf{H}}\|_F^2}{\|\mathbf{H}\|_F^2} \right], \tag{23}$$

is evaluated according to various r and ρ. As shown in Fig. 3 and 4, since the conventional method sends pilot symbols for all transmit antennas and applies the LMMSE algorithm, the channel estimation accuracy also improves as the SNR
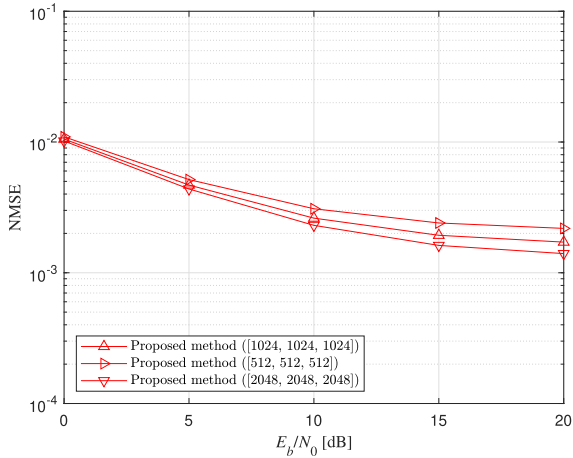
FIGURE 10. Comparison of NMSE performances according to DNN parameters ($N_T = 64$, $N_R = 2$, and $r = 0.969$).
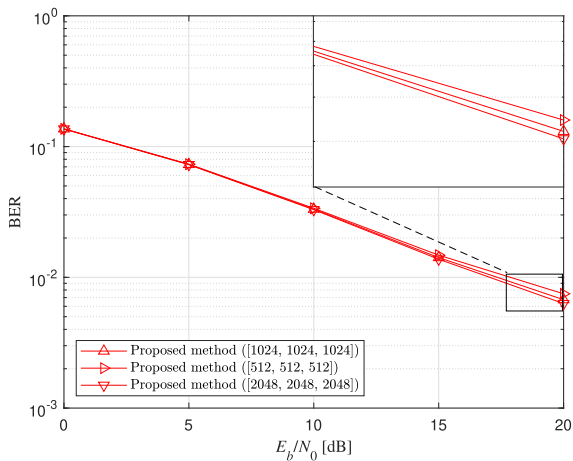

FIGURE 11. Comparison of BER performances according to DNN parameters ($N_T = 64$, $N_R = 2$, $r = 0.969$, and 64-QAM).

TABLE 3. Computational complexity for various DNN parameters.

Number of nodes of hidden layers	Computational complexity
$\{1024, 1024, 1024\}$	6.26MFLOPs
$\{512, 512, 512\}$	2.08MFLOPs
$\{2048, 2048, 2048\}$	20.91MFLOPs

spatial correlation, this phenomenon is inevitable. From now on, to show the performance gain of the proposed method, effectively, we only evaluate the $\rho = 0.9$ case.

Fig. 6 and 7 depict the BER performances of the proposed and conventional methods. Regardless of antenna configurations, the proposed method with $r = 0.969$ exhibits almost the same BER performance compared to the conventional method, unlike the NMSE performance. These results imply that the proposed method with $r = 0.969$ satisfies the NMSE performance to maintain the BER performance of the conventional method. Furthermore, we can see that the BER performance degradation is negligible, unlike the NMSE performance, which suffers severe degradation as r decreases.

In Fig. 8 and 9, the throughput performances of the proposed and conventional methods are compared. Since the proposed method utilizes fewer pilot symbols than the conventional method while maintaining BER performances, it shows better throughput performances. Moreover, even though the proposed method with $r = 0.875$ exhibits relatively large BER performance degradation, it outperforms the throughput performances of other cases. This implies that the effect of reducing pilot overhead dominates the throughput performance. However, if the number of pilot symbols utilized by the proposed method is significantly reduced, the target BER performance cannot be satisfied. Therefore, depending on the scenarios in which the proposed method is applied, r should be carefully chosen to compromise the BER and throughput performance.

Comparing the performances of the proposed method according to the antenna configurations, the NMSE performances of the $N_T = 128$ case are slightly worse than the $N_T = 64$ case. However, the $N_T = 128$ case shows better BER performance than the $N_T = 64$ case because it can achieve higher diversity gain. Note that since both cases adopt the same modulation order and the $N_T = 128$ case requires more pilot overhead for channel estimation than the $N_T = 64$ case, improvement of BER performance cannot lead to better throughput performance compared to $N_T = 64$. In practical systems, the $N_T = 128$ case adopts a higher modulation order based on the diversity gain and adaptive modulation, and it can show improved throughput performances [59]. Considering these points, we can see that reducing the pilot overhead is very important when applying pilot symbol-assisted channel estimation in massive MIMO systems.

Fig. 10 and 11 depict the NMSE and BER performances according to the DNN structures. To focus the performance tendency according to the number of nodes of hidden layers, we fixed the number of hidden layers of the DNN

increases [43]. However, the NMSE performances of the proposed method are degraded when compared with the conventional channel estimation method and this phenomenon is more severe as r decreases. This is because the proposed method is to predict the CSI of transmit antennas not sending pilot symbols by utilizing a DNN. As a result, there is a limitation in the NMSE performance of the proposed method despite the increase in SNR values and this implies that r should be carefully chosen to guarantee high link reliability of the communication system. In this paper, we will use $r = 0.875, 0.938$, and 0.969 which can obtain almost the same BER performances of the conventional method. Fig. 5 shows the NMSE performances of the conventional and proposed method according to the correlation coefficient, ρ, when $N_T = 64$, $N_R = 2$, and $r = 0.969$. In Fig. 5, the conventional method exhibits low NMSE performances for highly correlated channels [58]. Contrary to the conventional method, we can see that the NMSE performances of the proposed method are degraded as the value of ρ decreases. Considering that the DNN of the proposed method has to find the unknown variables without pilot symbols by only utilizing

structure and only varied the number of nodes of hidden layers. Note that other hyperparameters, e.g., the number of hidden layers, etc., must be optimized according to the system parameters [60], [61]. In Fig. 10 and 11, we can see that the performance of the proposed method can be improved as the dimension of the DNN structure is larger, and the BER performance of the {1024, 1024, 1024} case is closer to the {2048, 2048, 2048} case than the {512, 512, 512} case. However, to extend the proposed method to the next-generation communication systems, which guarantee a low latency, the computational complexity should be also considered [4], [6]. Therefore, in Table 3, we also compare the computational complexity of each DNN structure. Table 3 depicts that the computational complexity increases as the number of nodes of hidden layers [49]. Moreover, note that the computational complexity of the {1024, 1024, 1024} case is closer to the {512, 512, 512} case than the {2048, 2048, 2048} case. Therefore, based on the above two results, we chose the number of nodes of hidden layers as 1024 for the DNN structure by compromising the BER performances and computational complexity.

In summary, the proposed method can improve the throughput performances by reducing the pilot overhead. However, the degradation of the channel estimation accuracy is inevitable. Therefore, the proposed method has to be appropriately adopted according to the various target performances that the communication systems define for each use case.

V. CONCLUSION

We propose a deep learning-based channel estimation method for MIMO systems in spatially correlated channels to reduce the number of employed pilot symbols. Utilizing a DNN, the proposed method predicts the CSI of transmit antennas not sending pilot symbols and reconstructs the estimated MIMO channel matrix. Simulation results show the validity of the proposed method by comparing various performances with the conventional channel estimation method.

In future works, we will extend the proposed deep learning-based channel estimation method to a data-aided channel estimation method in time-varying channels to improve channel estimation accuracy and throughput performance.

REFERENCES

[1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[3] H. Noh, Y. Kim, J. Lee, and C. Lee, "Codebook design of generalized space shift keying for FDD massive MIMO systems in spatially correlated channels," *IEEE Trans. Veh. Technol.*, vol. 64, no. 2, pp. 513–523, Feb. 2015.

[4] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 118–125, Aug. 2020.

[5] C. Han, L. Yan, and J. Yuan, "Hybrid beamforming for terahertz wireless communications: Challenges, architectures, and open problems," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 198–204, Aug. 2021.

[6] H. Tataria, M. Shafi, M. Dohler, and S. Sun, "Six critical challenges for 6G wireless systems: A summary and some solutions," *IEEE Veh. Technol. Mag.*, vol. 17, no. 1, pp. 16–26, Mar. 2022.

[7] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: General model, design criteria, and signal processing," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 12–25, Nov. 2004.

[8] T. Hwang, C. Yang, G. Wu, S. Li, and G. Ye Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[9] *Technical Specification Group Radio Access Network; NR; Physical channels and modulation (Release 17)*, document TS 38.211 V17.1.0, 3GPP, Apr. 2022.

[10] *IEEE Approved Draft Standard for Information Technology—Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks—Specific Requirements: Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard 802.11-2020, 2020.

[11] *IEEE Standard for Low-Rate Wireless Network, Amendment 7: Defining Enhancements to the Smart Utility Network (SUN) Physcial Layer (PHYs) Supporting Up to 2.4 Mb/s Data Rates*, IEEE Standard 802.15.4-2015, Apr. 2019.

[12] Y. Gu and Y. D. Zhang, "Information-theoretic pilot design for downlink channel estimation in FDD massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2334–2346, May 2019.

[13] C.-J. Chun, J.-M. Kang, and I.-M. Kim, "Deep learning-based channel estimation for massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1228–1231, Aug. 2019.

[14] E. Balevi, A. Doshi, and J. G. Andrews, "Massive MIMO channel estimation with an untrained deep neural network," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2079–2090, Mar. 2020.

[15] X. Ma and Z. Gao, "Data-driven deep learning to design pilot and channel estimator for massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5677–5682, May 2020.

[16] L. Hobert, A. Festag, I. Llatser, L. Altomare, F. Visintainer, and A. Kovacs, "Enhancements of V2X communication in support of cooperative autonomous driving," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 64–70, Dec. 2015.

[17] S. Gyawali, S. Xu, Y. Qian, and R. Q. Hu, "Challenges and solutions for cellular based V2X communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 222–255, 1st Quart., 2021.

[18] Y. Yang, S. Dang, Y. He, and M. Guizani, "Markov decision-based pilot optimization for 5G V2X vehicular communications," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 1090–1103, Feb. 2019.

[19] S. Baek, I. Lee, and C. Song, "A new data pilot-aided channel estimation scheme for fast time-varying channels in IEEE 802.11p systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 5169–5172, May 2019.

[20] S. He, Q. Zhang, and J. Qin, "Pilot pattern design for two-dimensional OFDM modulations in time-varying frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1335–1346, Feb. 2022.

[21] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Model-driven deep learning for physical layer communications," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 77–83, Oct. 2019.

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

[23] F. Mahdavi, H. Zayyani, and R. Rajabi, "RSS localization using an optimized fusion of two deep neural networks," *IEEE Sensors Lett.*, vol. 5, no. 12, pp. 1–4, Dec. 2021.

[24] C. Lin and Y. Shin, "Deep learning-based multifloor indoor tracking scheme using smartphone sensors," *IEEE Access*, vol. 10, pp. 63049–63062, 2022.

[25] U. Demirhan and A. Alkhateeb, "Integrated sensing and communication for 6G: Ten key machine learning roles," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 113–119, May 2023.

[26] Y. Yang, F. Gao, X. Ma, and S. Zhang, "Deep learning-based channel estimation for doubly selective fading channels," *IEEE Access*, vol. 7, pp. 36579–36589, 2019.

[27] A. K. Gizzini, M. Chafii, A. Nimr, and G. Fettweis, "Deep learning based channel estimation schemes for IEEE 802.11p standard," *IEEE Access*, vol. 8, pp. 113751–113765, 2020.

[28] C. Wu, X. Yi, X. Zhu, W. Wang, L. You, and X. Gao, "Channel prediction in high-mobility massive MIMO: From spatio-temporal autoregression to deep learning," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1915–1930, Jul. 2021.

[29] Y. Yue, J. Shi, Z. Li, J. Hu, and Z. Tie, "Model-driven deep learning assisted detector for OTFS with channel estimation error," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 842–846, Apr. 2024.

[30] L. Guo, P. Gu, J. Zou, G. Liu, and F. Shu, "DNN-based fractional Doppler channel estimation for OTFS modulation," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 15062–15067, Nov. 2023.

[31] Y. Yang, F. Gao, G. Y. Li, and M. Jian, "Deep learning-based downlink channel prediction for FDD massive MIMO system," *IEEE Commun. Lett.*, vol. 23, no. 11, pp. 1994–1998, Nov. 2019.

[32] Q. Hu, F. Gao, H. Zhang, S. Jin, and G. Y. Li, "Deep learning for channel estimation: Interpretation, performance, and comparison," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2398–2412, Apr. 2021.

[33] A. Melgar, A. de la Fuente, L. Carro-Calvo, Ó. Barquero-Pérez, and E. Morgado, "Deep neural network: An alternative to traditional channel estimators in massive MIMO systems," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 2, pp. 657–671, Jun. 2022.

[34] S. R. Mattu, L. N. Theagarajan, and A. Chockalingam, "Deep channel prediction: A DNN framework for receiver design in time-varying fading channels," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6439–6453, Jun. 2022.

[35] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 232–257, Jan. 2020.

[36] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[37] Y.-H. Nam, B. L. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim, and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 172–179, Jun. 2013.

[38] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[39] D. Ying, F. W. Vook, T. A. Thomas, D. J. Love, and A. Ghosh, "Kronecker product correlation model and limited feedback codebook design in a 3D channel model," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 5865–5870.

[40] J. Choi and D. J. Love, "Bounds on eigenvalues of a spatial correlation matrix," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1391–1394, Aug. 2014.

[41] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," in *Proc. Veh. Technol. Conf.*, 1996, pp. 923–927.

[42] Y. Liu, Z. Tan, H. Hu, L. J. Cimini, and G. Y. Li, "Channel estimation for OFDM," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1891–1908, 4th Quart., 2014.

[43] J.-J. Van De Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE 45th Veh. Technol. Conf.*, May 1995, pp. 815–819.

[44] X. Dong, W.-S. Lu, and A. Soong, "Linear interpolation in pilot symbol assisted channel estimation for OFDM," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1910–1920, May 2007.

[45] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1971–1988, Jul. 2001.

[46] M. Simko, P. S. R. Diniz, Q. Wang, and M. Rupp, "Adaptive pilot-symbol patterns for MIMO OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4705–4715, Sep. 2013.

[47] R. M. Rao, V. Marojevic, and J. H. Reed, "Adaptive pilot patterns for CA-OFDM systems in nonstationary wireless channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1231–1244, Feb. 2018.

[48] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, Jan. 1989.

[49] X. Wang, X. Wang, T. Lv, L. Jin, and M. He, "HARNAS: Human activity recognition based on automatic neural architecture search using evolutionary algorithms," *Sensors*, vol. 21, no. 20, p. 6927, Oct. 2021.

[50] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[51] A. Faisal, H. Sarieddeen, H. Dahrouj, T. Y. Al-Naffouri, and M.-S. Alouini, "Ultramassive MIMO systems at terahertz bands: Prospects and challenges," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 33–42, Dec. 2020.

[52] *Technical Specification Group Radio Access Network; Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations (Release 17)*, document TR 25.996 V17.0.0, 3GPP, 2022.

[53] H. Lim, Y. Jang, and D. Yoon, "Bounds for eigenvalues of spatial correlation matrices with the exponential model in MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1196–1204, Feb. 2017.

[54] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[55] D. J. Love and R. W. Heath, "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1102–1110, Jul. 2003.

[56] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[57] X. Wei, C. Hu, and L. Dai, "Deep learning for beamspace channel estimation in millimeter-wave massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 182–193, Jan. 2021.

[58] M. Hajjaj, W. Chainbi, and R. Bouallegue, "Low-rank channel estimation for MIMO MB-OFDM UWB system over spatially correlated channel," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 48–51, Feb. 2016.

[59] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[60] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020.

[61] A. L. Ha, T. Van Chien, T. H. Nguyen, W. Choi, and V. D. Nguyen, "Deep learning-aided 5G channel estimation," in *Proc. 15th Int. Conf. Ubiquitous Inf. Manage. Commun. (IMCOM)*, Jan. 2021, pp. 1–7.

SANGGEUN LEE received the B.S. and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, South Korea, in 2013 and 2019, respectively.

In 2019, he joined Samsung Electronics, and has been working on 5G communication systems. His current research interests include beam management for mmWave systems and AI-based communication systems.

DONGKYU SIM (Member, IEEE) received the B.S. and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, South Korea, in 2011 and 2017, respectively.

From 2017 to 2018, he was a Postdoctoral Fellow with the Department of Electrical and Electronic Engineering, Yonsei University. During most of that time, he has carried out research on the next-generation communication systems, especially for the multi-antenna techniques of filter-bank multicarrier systems. From 2018 to 2020, he was a Senior Researcher with KT Corporation, Seoul, where he was involved in an AI-based network optimization research project. Since 2020, he has been an Assistant Professor with the School of Information and Communication Engineering, Chungbuk National University, Cheongju, Chungbuk, South Korea. His current research interests include AI-based communications and next-generation communication systems.

• • •