IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 14, 2021, accepted August 17, 2021, date of publication August 23, 2021, date of current version August 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3106739

Hybrid Neural Network-Based Fading Channel Prediction for Link Adaptation

CHAHYEON EOM AND CHUNGYONG LEE, (Member, IEEE)
School of Electrical and Electronic Engineering, Yonsei University, Seoul 03722, South Korea
Corresponding author: Chungyong Lee (cylee@yonsei.ac.kr)

This work was supported in part by Samsung Research in Samsung Electronics.

ABSTRACT We propose an artificial intelligence-based channel prediction scheme that can potentially facilitate link adaptation in customized communication systems. Link adaptation is a key process for wireless communication that requires accurate channel state information (CSI). However, the CSI may be outdated because of computational and propagation delays. In addition, a subframe with no CSI reference signal cannot provide CSI feedback. The proposed scheme solves these problems by predicting future channels. Although traditional stochastic methods suffer from marginal prediction accuracy or unacceptable computational complexity, neural networks allow time series prediction for channels even considering constraints for practical application. We introduce a hybrid architecture for improving the prediction accuracy of the neural network when extracting meaningful features. The proposed scheme uses a single hybrid network that can predict channels in different environments. Simulations were performed using a spatial channel model to evaluate the performance at the system-level, and the results indicated that the proposed scheme effectively increases the prediction accuracy for the channel quality indicator and spectral efficiency.

INDEX TERMS Adaptive modulation and coding, artificial intelligence, channel prediction, customized communication system, link adaptation, neural network, sixth generation.

I. INTRODUCTION

The fifth-generation (5G) technology standard for wireless communication systems is becoming commercially available worldwide [1]. However, despite the 5G era still being in its infancy, academia and industry are beginning to focus on the sixth-generation (6G) and beyond. Unlike previous generations, 6G will have revolutionary and stricter requirements. A key concept for 6G innovation is expected to be the evolution of wireless communication systems from *connected things* to *connected intelligence* [2]. Intelligence, autonomy, and context awareness are also expected to increase in importance [2]–[7]. Thus, 6G is expected to facilitate intelligent user-centered communication services.

In this paper, we define a user-centered communication system as a *customized communication system* for which services are optimized according to user characteristics. This allows the support method to be changed flexibly according to the requirements of different users. A customized communication system can be optimized according to user's life-style and experience by using big data. Adaptive services can be provided by considering the physical environment context (PEC) such as the surrounding location, mobility, and information from other users. One application of customized communication systems is channel prediction based on the user's channel experience and PEC. When the location, environment, and mobility are the same, artificial intelligence (AI) can be used to learn the functions of previous and future channels and to predict the channels of various scenarios for improved link adaptation.

Link adaptation is a key process for satisfying the various requirements of 5G new radio (NR) access technology and involves the dynamic adjustment of the transmitted information data rate (i.e., modulation scheme and coding rate) to match the radio channel capacity to each user. Adaptive modulation and coding (AMC) is a technique for determining the proper modulation and coding scheme (MCS) depending on the channel quality and guarantees the throughput for time-varying channels of mobile wireless communication systems. According to AMC, the user equipment (UE) periodically measures the channel quality via a reference signal (RS) and a signal-to-noise ratio (SNR) is utilized to map the channel quality to the channel quality index (CQI). The base station (BS) then selects an MCS by considering

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

the CQI reported by the UE. The BS periodically transmits a channel state information reference signal (CSI-RS), which the UE used to estimate the channel quality, and the BS acquires the CSI based on feedback from the UE. In practical applications, however, the CSI feedback can be outdated because of computational and propagation delays at both the BS and UE [8], [9]. Moreover, a subframe without the CSI-RS cannot provide CSI feedback. For 5G NR, CSI-RS transmissions are supported at 5, 10, 20, 40, 80, 160, 320, and 640 slots [10].

To address the performance degradation due to CSI feedback delays, researchers have focused on developing methods for predicting future channels and delivering the predicted CSI to the BS [11]–[13]. Channel prediction has attracted increased attention as an efficient technology that can directly improve the CSI quality without requiring additional resources. There are two representative channel prediction approaches as follows: autoregressive (AR) and sum of sinusoids (SOS) [14]. AR models approximate future channels based on past channels [15], [16]. SOS models represent a channel as a superposition of a finite number of sinusoids and assume that their parameters (e.g., the amplitude, Doppler shift, and scatter) change much more slowly than those of the channel [17]. Because AR models directly estimate the model coefficients even when the channel is time-variant, they are more suitable than SOS models for realistic channel simulations [14]. However, AR models are susceptible to damage such as additional noise [18].

AI can be used to overcome the limitations of traditional stochastic methods and obtain the advantages of channel prediction. AI has attracted attention as a core technology for 6G, and it is widely used in communication systems [2]–[7]. Among AI applications, the recurrent neural network (RNN) and long short-term memory (LSTM) are effective for time series prediction and are being actively used for channel prediction. Liu *et al.* [19] applied an RNN to predicting single- and multiple-antenna channels. Jiang and Schotten [20], [21] proposed a real-value RNN for implementation in a multistep predictor. Neural networks with structures other than the RNN have also been used for channel prediction [22], [23]. Neural network-based channel prediction is a data-driven approach that can flexibly adapt to various scenarios.

Although existing prediction schemes that utilize not only LSTM but also other neural network structures predict multi-path channels as a whole, most focus on modeling flat-fading multiple-input multiple-output (MIMO) channels. Some previous studies have focused on frequency-selective channels [17], [24]–[30]. Semmelrodt and Kattenbach [17] and Jiang and Schotten [24]–[26] predicted channels by converting the frequency-selective channel into independent frequency-flat channels considering orthogonal frequency-division multiplexing (OFDM). Lv *et al.* [27] proposed a channel prediction scheme for a millimeter wave (mmWave) MIMO-OFDM system that considers the four following domains: array-frequency, array-time, angle-frequency, and angle-time. Ahrens *et al.* [28] insisted that a shift-invariant prediction model could be applied to all subcarriers for channel prediction in the frequency domain because the utilized bandwidth is always significantly smaller than the carrier frequency. However, their assumption may not hold depending on the environment. Finally, Liu *et al.* [29] and Sui *et al.* [30] proposed channel prediction in the time-delay domain. Liu *et al.* [29] were the first to propose a MIMO-OFDM channel prediction model in time domain, and Sui *et al.* [30] introduced a channel prediction framework based on the temporal and spatial correlations in MIMO-OFDM systems to prevent overfitting and ensure efficient learning. Sui *et al.* [30] also argued that correlations exist only for the same transmitting path in the channel impulse response based on the wide-sense stationary uncorrelated scattering (WSSUS). They proposed a broad echo state network (BESN) to predict the channel by path for every antenna pair.

We propose a frequency-selective channel prediction scheme based on a neural network with a hybrid architecture (i.e., hybrid network). The proposed scheme can improve the channel prediction framework based on temporal and spatial correlations in MIMO-OFDM systems by deploying an LSTM for time series prediction and convolutional neural networks (CNN) for capturing spatial information. The proposed scheme uses a single hybrid network to predict channels in different environments. We also propose a system operation method for link adaptation. The performance of the proposed scheme was verified through a system-level simulation (SLS) with a standardized spatial channel model (SCM) developed by 3GPP and 3GPP2 [31]–[33].

Notation: $\|\cdot\|_2$ is the Euclidean norm, $|\cdot|$ returns its absolute value, $(\cdot)^*$ denotes the complex conjugate, $\otimes$ represents the Kronecker product, Re$(\cdot)$ refers to the real part, Im$(\cdot)$ denotes the imaginary part, $\mathbb{C}^{N\times M}$ stands for a $N \times M$ size complex-valued matrix, and E$[\cdot]$ represents the expectation operation.

II. SYSTEM MODEL AND CONVENTIONAL SCHEMES

A. SYSTEM MODEL

We considered a MIMO-OFDM system where the BS and UE have N_T and N_R antennas, respectively. The transmitted signal passes through the SCM channel comprising N clusters, which causes the multipath delay. Each cluster consists of M rays. Fig. 1 shows the simplified channel model. For the t^{th} time slot, the channel impulse response between the s^{th} antenna element of the BS and u^{th} antenna element of the UE is given by

$$h_{u,s}(t,\tau) = \sum_{n=1}^{N} h_{u,s,n}(t)\delta(\tau - \tau_n), \tag{1}$$

where n is the cluster index, $u = 1, 2, \ldots, N_R$ is the antenna index for the UE, $s = 1, 2, \ldots, N_T$ is the antenna index for the BS, $\tau = 1, 2, \ldots, \tau_{max}$ is the delay index in the discrete time domain, τ_n is the multipath delay of the n^{th} path in the discrete time domain, and $\delta(\cdot)$ represents the impulse function.

According to the power delay profile of the channel model, the complex path coefficient $h_{u,s,n}$ is given by

$$h_{u,s,n}(t) = \sqrt{P_n} g_{u,s,n}(t), \tag{2}$$

where P_n and $g_{u,s,n}$ are the gain and channel coefficient, respectively, of the n^{th} cluster. The channel coefficient of the n^{th} cluster is then expressed as

$$g_{u,s,n}(t) = \sum_{m=1}^{M} \frac{A_m}{\sqrt{M}} \exp(j\phi_m) \exp(j2\pi f_m t) \times \exp(j\frac{2\pi}{\lambda} d_T \cos\alpha_{n,m}) \exp(j\frac{2\pi}{\lambda} d_R \cos\beta_{n,m}), \tag{3}$$

where A_m, ϕ_m and f_m are the attenuation, phase shift, and Doppler frequency, respectively, of the m^{th} ray, λ indicates the carrier wavelength in meters, d_T and d_R denote the distances between the BS antenna elements and between the UE antenna elements, respectively, in meters, and $\alpha_{n,m}$ and $\beta_{n,m}$ are the angle of departure and angle of arrival, respectively, for the m^{th} ray of the n^{th} cluster. These parameters vary over time, and the process for generating each parameter is described in the literature [31]–[33].

Let $\mathbf{h}_{n,t} \in \mathbb{C}^{N_R \times N_T}$ denote a MIMO channel matrix, observed at the t^{th} time slot of the n^{th} cluster, whose elements are given by $[\mathbf{h}_{n,t}]_{u,s} = h_{u,s,n}(t)$. The fast Fourier transform (FFT) can be used with (2) to derive the channel frequency response for the t^{th} time slot and k^{th} subcarrier:

$$H_{u,s}(t,k) = \sum_{n=1}^{N} \sqrt{P_n} g_{u,s,n}(t) \exp(-j\frac{2\pi\tau_n k}{N_{FFT}}), \tag{4}$$

where N_{FFT} is the FFT size. FFTs can be replaced by the discrete Fourier transforms. Let $\mathbf{H}(t,k)$ denote a MIMO channel matrix observed at the k^{th} subcarrier in the frequency domain whose elements are given by $[\mathbf{H}(t,k)]_{u,s} = H_{u,s}(t,k)$. At the receiver side, the received signal in the time domain is converted into the frequency domain after the cyclic prefix is removed. The input-output relationship is then represented as

$$\mathbf{y}(t,k) = \mathbf{H}(t,k)\mathbf{x}(t,k) + \mathbf{n}(t,k), \tag{5}$$

where $\mathbf{y}$, $\mathbf{H}$ and $\mathbf{n}$ are the received signal, transmitted OFDM symbol and white Gaussian noise, respectively.

The least squares method can be used to estimate the channel at the pilot position (i.e., CSI-RS). The channel at the pilot position is given by

$$\hat{\mathbf{H}}_p(t,k) = \arg\min_{\mathbf{H}_p} \|\mathbf{y}_p(t,k) - \mathbf{H}_p(t,k)\mathbf{x}_p(t,k)\|_2^2, \tag{6}$$

where $\mathbf{y}_p$, $\mathbf{H}_p$ and $\mathbf{x}_p$ represents the received signal, channel and transmit signal, respectively, at the pilot position. To estimate the channel at resource elements other than the pilot positions, we can exploit a two-dimensional interpolation method. Link adaptation can be performed by using the estimated channels in the time domain, $\hat{\mathbf{h}}_{n,t}$, $n = 1, 2, \ldots, N$, computed with an inverse FFT.

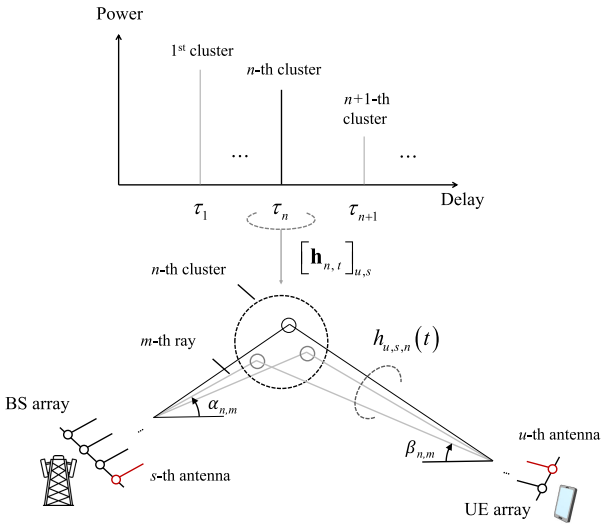


FIGURE 1. Simplified SCM channel model and the power delay profile.

B. CONVENTIONAL CHANNEL PREDICTION ALGORITHMS

In AR model-based algorithms, the predicted complex path coefficient is given by

$$\tilde{h}(t+1) = \sum_{r=1}^{R_a} d(r) h(t-r), \tag{7}$$

where $\tilde{h}$ and h are the predicted complex path coefficient and the past complex path coefficient used as the input, respectively, R_a denotes the AR model order, and $d(r)$ indicates the time-variant AR coefficient that is computed by using the minimum mean squared error (MMSE) criterion [34]. For the sake of simplicity but without loss of generality, we excluded the indices of the antenna elements.

Sui *et al.* [30] proposed a channel prediction framework called group forward variable selection-based BESN (GFVS-BESN). It considers a broad learning system and reduces the time consumed during the training stage compared to prediction schemes using other neural networks.

Jiang and Schotten [24] proposed a predictor based on deep recurrent neural networks including LSTM. Although they designed their predictor for flat fading channels, it can be applied to frequency-selective channels by converting them into a set of N narrow-band subcarriers. In other words, it predicts channels in the frequency domain. Therefore, neural networks are likely to have vertical stacks of vast layers, which increase the risk of overfitting.

Hochreiter and Schmidhuber [35] proposed the LSTM network, which is a kind of RNN that can be exploited for channel prediction. The LSTM network is good at handling sequence dependencies and time series prediction. The LSTM utilizes a cell to store the long-term state, which is mainly controlled by the three following gates: the forget gate, input gate, and output gate. The LSTM network comprises multiple LSTM cells that are connected in cascade. Let L denote the number of LSTM cells or inputs to reference. The weight matrices and biases for the LSTM cells are

denoted by $\mathbf{W}_f$, $\mathbf{b}_f$, $\mathbf{W}_i$, $\mathbf{b}_i$, $\mathbf{W}_c$, $\mathbf{b}_c$, $\mathbf{W}_o$ and $\mathbf{b}_o$. Fig. 2 shows the structure of the LSTM network.

The forget gate determines whether the information from the previous state contributes to the next cell state. For the v^{th} LSTM cell, this is expressed as

$$\mathbf{f}_v = \sigma(\mathbf{W}_f \cdot [\mathbf{s}_{v-1}, \mathbf{x}_v] + \mathbf{b}_f), \tag{8}$$

where $\mathbf{x}_v$ is the current input, $\mathbf{s}_{v-1}$ is the previous state and $\sigma(\cdot)$ represents the sigmoid function. The input gate determines how much of the input of the current network is saved to the cell state. The related equation is as follows:

$$\begin{aligned} \mathbf{i}_v &= \sigma(\mathbf{W}_i \cdot [\mathbf{s}_{v-1}, \mathbf{x}_v] + \mathbf{b}_i), \\ \tilde{\mathbf{c}}_v &= \tanh(\mathbf{W}_c \cdot [\mathbf{s}_{v-1}, \mathbf{x}_v] + \mathbf{b}_c). \end{aligned} \tag{9}$$

The cell state is then updated as follows:

$$\mathbf{c}_v = \mathbf{f}_v \otimes \mathbf{c}_{v-1} + \mathbf{i}_v \otimes \tilde{\mathbf{c}}_v. \tag{10}$$

The outputs of the LSTM cell are calculated as follows:

$$\begin{aligned} \mathbf{o}_v &= \sigma(\mathbf{W}_o \cdot [\mathbf{s}_{v-1}, \mathbf{x}_v] + \mathbf{b}_o), \\ \mathbf{s}_v &= \mathbf{o}_v \otimes \mathrm{ReLU}(\mathbf{c}_v). \end{aligned} \tag{11}$$

Finally, the output of the last LSTM cell goes through the fully connected (FC) layers. Each hidden layer consists of a linear operation and nonlinear activation function. In this case, the mean-squared error (MSE) is used to train the LSTM. Jiang and Schotten [24] utilized a deep LSTM network built by stacking multiple LSTM layers, which is also described in Fig. 2.

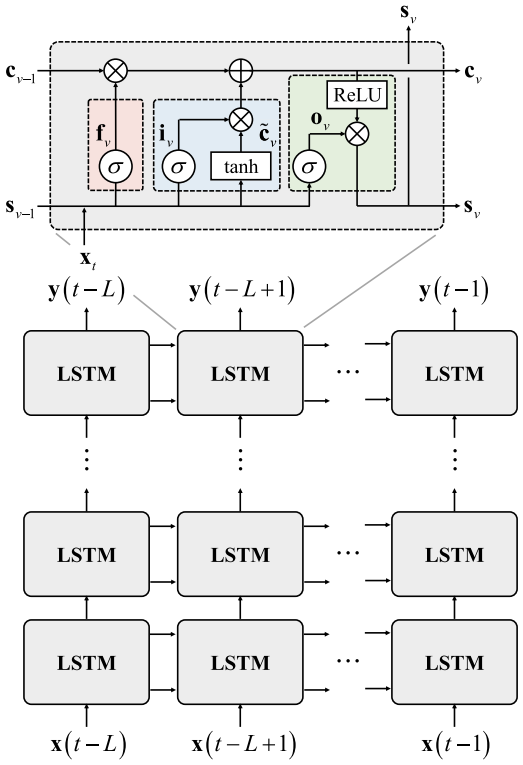


FIGURE 2. Architecture of the LSTM cell and deep LSTM.

III. PROPOSED CHANNEL PREDICTION SCHEME FOR THE AMC SOLUTION

A. NETWORK ARCHITECTURE FOR CHANNEL PREDICTION

The hybrid network in the proposed scheme connects an LSTM, CNN and FC layer in series and repeatedly. The hybrid network overcomes the drawbacks of conventional deep learning methods for channel prediction, which have to be retrained with a new network whenever the channel environment changes. The hybrid network can predict channels in various environments with one integrated network because of the FC layer, which receives the channel environment as an additional input. In addition, it can capture temporal and spatial correlations. The channel prediction accuracy is improved by the CNN, which can consider spatial correlations to extract meaningful features.

The composition and role of the hybrid network are as follows. Features are extracted by considering the spatial correlation between antenna elements. The features are then transformed according to the environment by using the PEC as an additional network input. The LSTM then uses these features for time series prediction, and the CNN regenerates channels with the predicted features. In summary, the hybrid network consists of the following four parts: CNN-based feature extraction, FC layer-based reflection of the PEC, LSTM-based feature prediction, and CNN-based channel reproduction. In particular, the hybrid network is similar to the structure of a U-net as proposed by Ronnerberger *et al.* [36]. U-net consists of a contracting path for downsampling, an expanding path for upsampling, and a bottleneck. In contrast, the hybrid network exploits the FC layers and LSTM instead of the bottle neck to enable channel prediction. Fig. 3 shows the structure of the hybrid network.

According to the WSSUS, correlations exist only for same transmitting path [30]. Therefore, the hybrid network predicts the channel for the n^{th} path by using the same path for every antenna pair. In this study, we considered only N paths with the largest channel gain. We also assumed that the delay of each tap does not change.

1) CNN-BASED FEATURE EXTRACTION

Unlike conventional LSTMs, this partial network extracts meaningful features by considering the spatial correlation between antenna elements for effective learning. A CNN can capture spatial information by using convolutional filters [37]. Thus, it is used here to easily extract and handle features for obtaining the spatial correlation between antenna elements.

This partial network is similar to the contracting path of U-net. Because convolution is performed without padding, the size of the feature map is reduced. Although max-pooling was not considered in this study, it can be applied depending on the input dimensions, and parameters such as the filter size can be adjusted. The following parameters were used in this study. The network input is a $N_R \times N_T \times 2N$ tensor that contains channel matrices for each cluster, which are

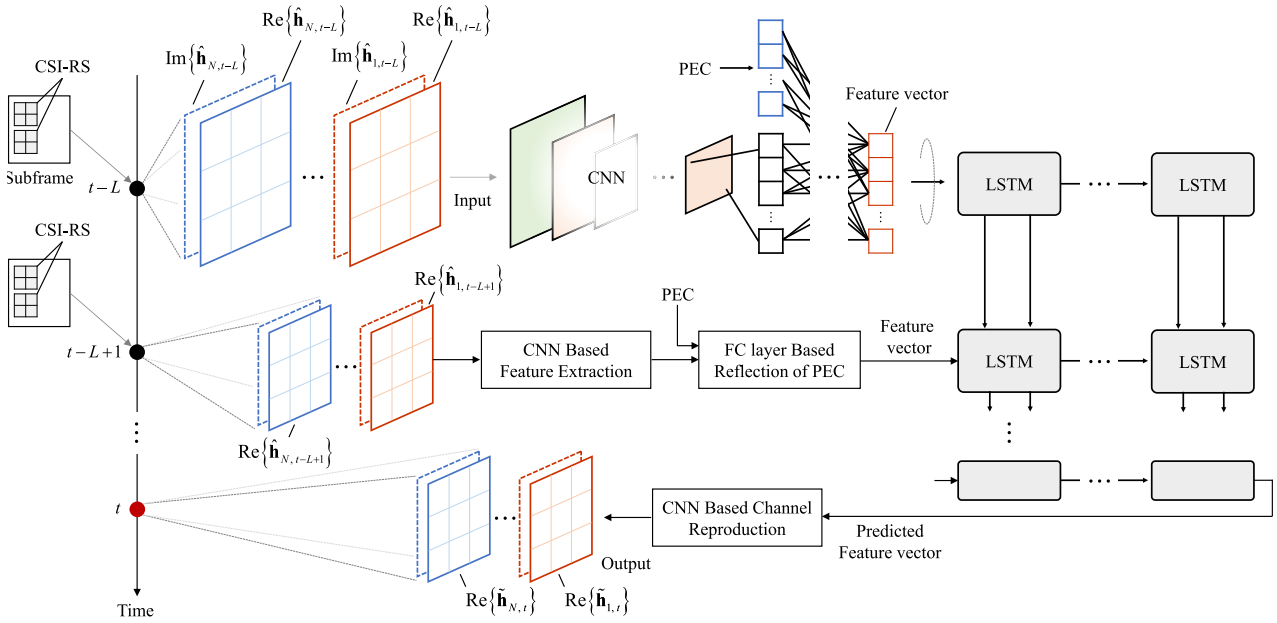


FIGURE 3. Architecture of the hybrid network for the proposed scheme.

divided into real and imaginary matrices (i.e., $\mathrm{Re}\{\hat{\mathbf{h}}_{n,t}\}$ and $\mathrm{Im}\{\hat{\mathbf{h}}_{n,t}\}$, respectively, for all n). The first convolution layer consists of $2N$ filters of size 2×2 (stride 1), the second layer is composed of $2N$ filters of size 3×1 (stride 1), and the third convolution layer is made up $2N$ filters of size 3×1 (stride 1). Finally, the CNN is connected to an FC layer with 64 units. Each layer has a rectifier linear unit nonlinear activation function. The CNN extracts $n_e \times 1$ feature vectors from channels of N clusters as inputs and finally outputs the $n_eN \times 1$ feature vector. It can compress the amount of information needed to improve the prediction accuracy for effective training of the LSTM.

2) FC LAYER-BASED REFLECTION OF PEC

Conventional LSTMs are limited because the prediction accuracy decreases when the channel environment varies. Here, the channel environment is input to the FC layer after the CNN so that channels can be predicted for various environments with only an integrated network. In other words, the FC layer adds the PEC to the $n_eN \times 1$ feature vector extracted from the CNN. The feature vector is transformed to increase the prediction accuracy by reflecting the PEC, which is represented by the 3D location of the UE, indoor/outdoor condition, and mobility. Thus, the PEC is given as a 7×1 vector containing this information. The 3D location is a real number; the coordinates are given in a Cartesian coordinates where the BS is the origin. The indoor/outdoor condition is indicated by 1 bit. The mobility is the velocity (km/h) of the UE relative to the BS in Cartesian coordinates. The number of neurons is given by (32, 32). The transformed feature vector output of the FC layer is expressed as an $n_fN \times 1$ vector.

3) LSTM-BASED FEATURE PREDICTION

In this partial network, the features extracted in the previous process are predicted in time series. The dimensions of the LSTM input are reduced by the previous processes to increase the prediction accuracy and decrease the computational complexity. The structure of the network is the same as that shown in Fig. 2 except that the input is the extracted features. For a deep LSTM with 12 memory cells, L features are utilized to predict the future.

4) CNN-BASED CHANNEL REPRODUCTION

This partial network reproduces channels from the predicted feature vectors and is similar to the expanding path of U-net. The structure is the inverse of that used for the CNN-based feature extraction. The deconvolution layer can be utilized if max-pooling was used in the previous step. The output is an $N_R \times N_T \times 2N$ tensor that contains channels for each cluster in the same form as the input. In other words, the output comprises $\mathrm{Re}\{\tilde{\mathbf{h}}_{n,t}\}$, $\mathrm{Im}\{\tilde{\mathbf{h}}_{n,t}\}$ for all n. Eventually, the predicted channels is obtained. The hybrid network is trained according to the MSE.

B. AMC SOLUTION-BASED CHANNEL PREDICTION

The proposed scheme utilizes the hybrid network in the time domain for channel prediction. The AMC solution is obtained in the following two phases: training and prediction. Fig. 4 illustrates two phases and multistep prediction.

1) TRAINING PHASE

In the training phase, the UE receives a series of subframes containing the CSI-RS. The UE estimates the channel coefficients and PEC in advance. Such data are then used to train

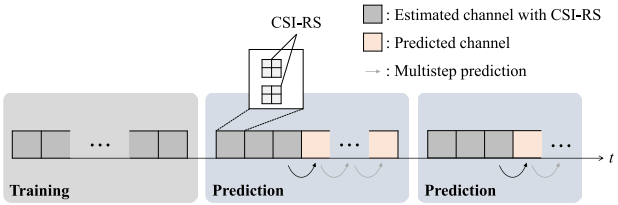


FIGURE 4. Illustration of training phase and prediction phase.

the hybrid network. Thus, the training phase requires not only large amounts of data to be collected in various environments but also more RS transmissions. The previous channel coefficients and PEC serve as the network input, and the output is a complex path coefficient that has passed through different clusters. The objective of the hybrid network is to minimize the difference between the predicted and labeled data. Training continues until the prediction accuracy is less than the target accuracy. Algorithm 1 summarizes the training phase for obtaining the AMC solution.

2) PREDICTION PHASE

After training, the hybrid network predict channels that cannot be estimated because of the lack of CSI-RS in the subframe. The BS transmits L consecutive RSs, and the UE estimates the channel according to the PEC. The UE then predicts the channel for the next frame by using the trained hybrid network. The SNR and CQI for the next frame can also be predicted according to the predicted channel. The UE reports the predicted CQI to the BS, which selects the MCS based on the conventional outer loop link adaptation technique of using a lookup table to map the SNR to the CQI [38], [39]. The prediction can be repeated by using the predicted channel as the input and corrected according to the estimated channel with the RS that is received later. Algorithm 2 summarizes the prediction phase for obtaining the AMC solution.

C. COMPUTATIONAL COMPLEXITY

We analyzed the computational complexity of the hybrid network. The hybrid network is divided into four networks comprising the following three types of neural networks: CNN, LSTM, and FC layer (i.e., deep-NN). The computational complexities of these neural networks have previously been calculated in the literature [24], [40]. Thus, the computational complexity per time step of the hybrid network is given by $\mathcal{O}(2N_C + N_L + N_F)$, where N_C, N_L, and N_F are the numbers of parameters (i.e., weights and biases) of the CNN, LSTM, and FC layer, respectively. The computational complexity of each partial network is calculated as follows. $\mathcal{O}(N_C)$ can be calculated simply as $\mathcal{O}(W_F H_F L_C)$, where the CNN performs convolution by using L_C filters with a width W_F and height H_F [40]. If we assume that a single-input single-output (SISO) channel is predicted, then N_L is given by $N_L = n_e N N_p$. N_p is the number of parameters of LSTM for the SISO channel predictor and is calculated as

Algorithm 1 Training Phase

1: Initialize weights of the hybrid network with random numbers.
2: The BS transmits a series of subframes containing CSI-RS.
3: The UE estimates the channel coefficients and PEC using the received RSs.
4: **repeat**
5: Get channel coefficients for T_s time.
6: **for** $t = L$ to T_s **do**
7: **while** $n \leq N$ **do**
8: Stack $\mathrm{Re}\{\hat{\mathbf{h}}_{n,t-L+1}\}$, $\mathrm{Im}\{\hat{\mathbf{h}}_{n,t-L+1}\}, \ldots,$ $\mathrm{Re}\{\hat{\mathbf{h}}_{n,t-1}\}$ and $\mathrm{Im}\{\hat{\mathbf{h}}_{n,t-1}\}$.
9: **end while**
10: Train the hybrid network: use stacked data as the input and use $\mathrm{Re}\{\hat{\mathbf{h}}_{n,t}\}$ and $\mathrm{Im}\{\hat{\mathbf{h}}_{n,t}\}$ as the label.
11: **end for**
12: **until** The prediction accuracy is less than the target accuracy.

Algorithm 2 Prediction Phase

1: **repeat**
2: The BS transmits L consecutive RSs and the UE receives them.
3: Prediction for time $t+1$: input $\mathrm{Re}\{\hat{\mathbf{h}}_{n,t-L+1}\}$, $\mathrm{Im}\{\hat{\mathbf{h}}_{n,t-L+1}\}, \ldots, \mathrm{Re}\{\hat{\mathbf{h}}_{n,t}\}$ and $\mathrm{Im}\{\hat{\mathbf{h}}_{n,t}\}$ for $n \in \{1, 2, \ldots, N\}$ to the trained network.
4: The output of the network consisting of $\mathrm{Re}\{\tilde{\mathbf{h}}_{n,t}\}$, $\mathrm{Im}\{\tilde{\mathbf{h}}_{n,t}\}$ for $n \in \{1, 2, \ldots, N\}$, is the predicted channel coefficients.
5: The UE gets the CQI based on the predicted channel and feeds it back to the BS.
6: The BS takes the MCS.
7: L channels including the predicted channel are again input to the hybrid network to predict the next channel.
8: **until** The BS transmits L consecutive RSs again.

$$N_p = \sum_{l=1}^{H_L} 4\left(n_L^{(l-1)} n_L^{(l)} + n_L^{(l)} n_L^{(l)} + n_L^{(l)}\right) + n_L^{(H_L)} n_L^{(H_L+1)} + n_L^{(H_L+1)}, \quad (12)$$

where $n_L^{(l)}, l = 1, \cdots, H_L$, represents the number of memory cells in the l^{th} vertical layer, $n_L^{(0)}$ and $n_L^{(H_L+1)}$ are the number of input units and output units, respectively [24]. Suppose that the FC layer consists of an input layer, L_F hidden layers in which the number of nodes in the l^{th} layer is $n_F^{(l)}, l = 1, \cdots, L_F$, and an output layer. Then, the number of parameters for the FC layer can be computed as

$$N_F = \sum_{l=0}^{L_F} n_F^{(l)} \times n_F^{(l+1)}, \quad (13)$$

where $n_F^{(0)}$ and $n_F^{(L+1)}$ are the numbers of input nodes and output nodes, respectively. Finally, the total computational complexity of the hybrid network is given by $\mathcal{O}\left(n_t\left(2N_C + n_e N N_p + N_F\right)\right)$, where n_t is the time step.

IV. SIMULATION RESULTS

A. SIMULATION PARAMETERS

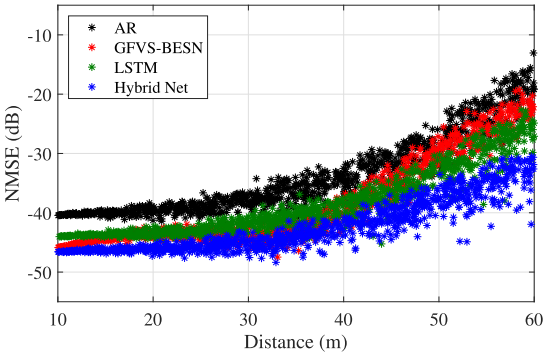
We evaluated the performance of the proposed prediction scheme by utilizing an SLS based on the SCM of 3GPP. The system had a bandwidth of 10 MHz and a carrier frequency of 2.1 GHz. Each resource block had 12 subcarriers in total and a subcarrier spacing of 15 kHz. A 3D urban micro-cell was considered. We used mutual information effective signal-to-noise-ratio mapping to calculate the effective SNR. Table 1 presents the main parameters (refer to [31]–[33] for the other simulation parameters).

To train the neural network, 50,000 input–label pairs were generated. Specifically, the previous and current channel coefficients after channel estimation were used as the input and label, respectively. For the hybrid network, PEC was used as an additional input. We set the batch size to 200, n_f to 16, and number of epochs to 10,000 based on trial and error. Batch normalization and an ADAM optimizer with an initial learning rate of 10^{-6} were used. The filters and neurons were weighted according to He initialization.

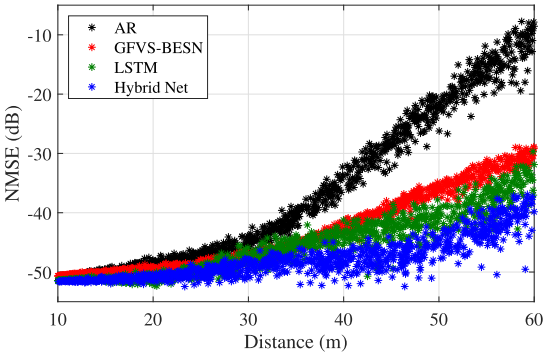
We compared four prediction schemes as follows: the AR model, GFVS-BESN [30], conventional LSTM network [24] (i.e., deep-LSTM), and the proposed scheme (i.e., hybrid network). Because the PEC cannot be considered with existing prediction schemes other than the proposed scheme, we generated multiple models or networks according to the PEC and trained each. In particular, prediction schemes were created by dividing each into five models according to the radial distance from the BS. For example, a particular LSTM network was trained to predict the channel of a UE that is outdoors at a radial distance of 10–20 m from the BS and speed of 3 km/h or less. Thus, the performances of the conventional schemes could be compared with that of the proposed scheme. The AR model was extracted by utilizing the MMSE of 500 samples. For channel prediction with the GFVS-BESN, we assumed that the spatial correlation matrix is known exactly. We also exploited parameters introduced by Jiang and Schotten [24] and Sui *et al.* [30] to evaluate the GFVS-BESN and LSTM except for the number of cells of the LSTM. An LSTM network with 60 memory cells was trained to predict channels for 600 subcarriers instead of channels of 16 paths in the time domain. We used a computer with an Intel(R) Core(TM) i7-10700k CPU operating at 3.8 GHz and 16 GB of RAM. We also utilized Python 3.7 and MATLAB (2018a). Table 2 presents the average time consumed during the training stage per input.

TABLE 1. Main simulation parameters.

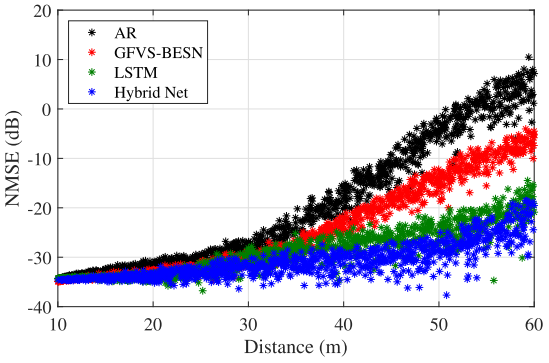
Parameter	Value
Frame length	10 ms
Subframe length	1 ms
Total number of resource block	50
Number of antennas of BS, N_T	8
Number of antennas of UE, N_R	2
Number of clusters, N	16
Height of BS	10 m
Height of building for indoor UE	8 m
Antenna array structure	UPA
UE speed	$\leq$ 3 km/h or 30 km/h
BS transmit power	41 dBm



(a) Indoor and UE speed $\leq$ 3 km/h



(b) Outdoor and UE speed $\leq$ 3 km/h



(c) Outdoor and UE speed = 30 km/h

FIGURE 5. NMSE performances of the various schemes according to distance.

B. PREDICTION ACCURACY

The normalized MSE (NMSE) was used to measure the prediction accuracy:

$$\mathrm{NMSE} = \mathrm{E}\left[\frac{\sum_u \sum_s \sum_n \left(h_{u,s,n} - \tilde{h}_{u,s,n}\right)^2}{\sum_u \sum_s \sum_n {h_{u,s,n}}^2}\right], \qquad (14)$$

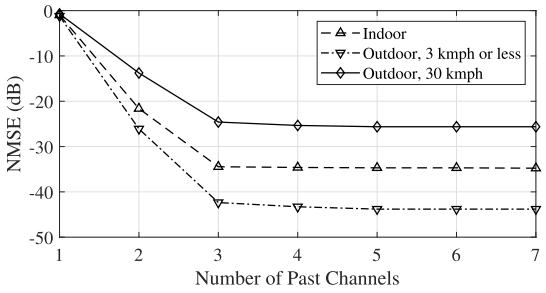


FIGURE 6. NMSE performances of the proposed scheme according to the number of past channels.

TABLE 2. Average time consumed during the training stage per input.

Method	Average time consumed (ms)
GFVS-BESN [30]	6.172
LSTM [24]	683.755
Hybrid Net	709.148

where $h_{u,s,n}$ and $\tilde{h}_{u,s,n}$ denote the correct and predicted complex path coefficients, respectively.

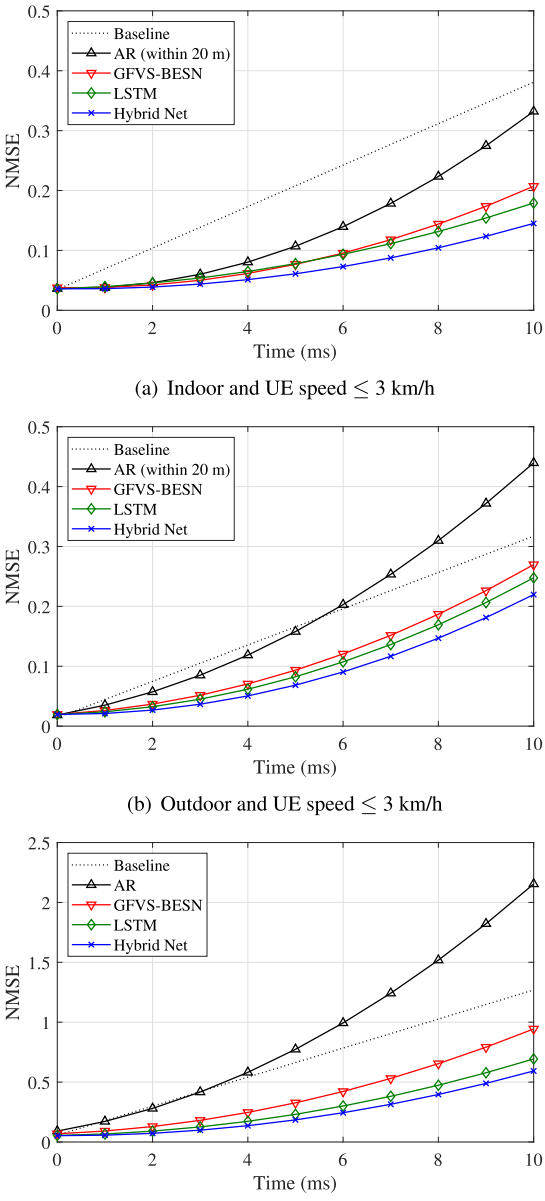
Fig. 5 depicts the NMSE performances of the various schemes according to the distance between the BS and UE for $L = 3$. Fig. 5(a), (b), and (c) present the following different scenarios: indoor or outdoor and pedestrian (UE speed $\leq$ 3 km/h) or vehicular (UE speed $=$ 30 km/h). We tested 2000 independent UEs for each case. The channel prediction error increased with the radial distance. Therefore, the NMSE indicates the prediction error according to the initial input of L consecutive channels with estimation error. The NMSEs decreased with increasing distance from the BS for all schemes and scenarios. However, the proposed scheme improved the prediction accuracy by extracting significant features. For all scenarios, the proposed scheme performed better than the other schemes. In addition, the proposed scheme was the most robust against channel estimation error. Note that the proposed scheme was able to achieve this result with a single unified network.

Fig. 6 shows the prediction accuracy according to the number of past channels (i.e., L) used for prediction. For all schemes, the NMSE increased with up to three past channels, and then it became saturated. The appropriate number of past channels needed to be determined because the CSI and feedback overheads increase with the number of past channels.

C. LINK ADAPTATION PERFORMANCE

Next, the link adaptation performance was evaluated. We assumed that 1) the neural networks learned in the training phase, 2) the CSI-RS had a period of 10 ms, 3) three consecutive subframes including the CSI-RS were initially received for the prediction (i.e., $R_a = L = 3$), and 4) the PEC could be estimated correctly every time.

We tested 2000 independent UEs for 20 transmission time intervals in each case. The previous channel prediction was reused as the input for predicting the next channel



(a) Indoor and UE speed $\leq$ 3 km/h

(b) Outdoor and UE speed $\leq$ 3 km/h

(c) Outdoor and UE speed $=$ 30 km/h

FIGURE 7. Multistep prediction performances in the prediction phase.

(i.e., multistep prediction introduced by Sui *et al.* [30]). We compared the following five schemes: the baseline scheme, AR model, GFVS-BESN, LSTM, and proposed scheme. Each scheme used channel prediction for link adaptation. For the baseline scheme, when the subframe did not have the CSI-RS, the CSI of the previous subframe was used as-is.

Fig. 7 shows the multistep prediction performances of various schemes in the prediction phase. Fig. 7(a), (b), and (c) present three scenarios. For the AR model, the results extracted for UEs within 20 m from the BS are shown. Fig. 7(c) shows only the results for UEs within 30 m from the BS. The error propagation due to recursive processes was lowest with the proposed scheme, even though the conventional schemes used multiple models or networks.

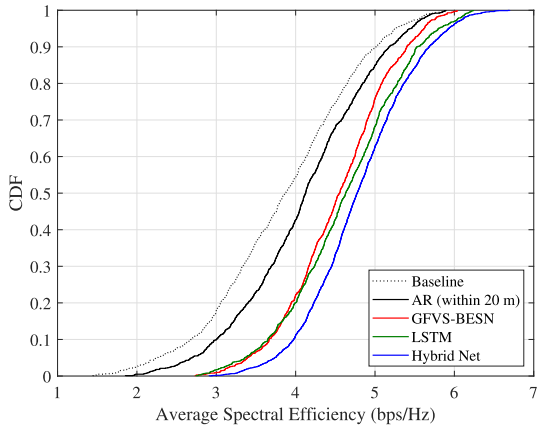


FIGURE 8. CDF of the average spectral efficiency in the pedestrian scenario.

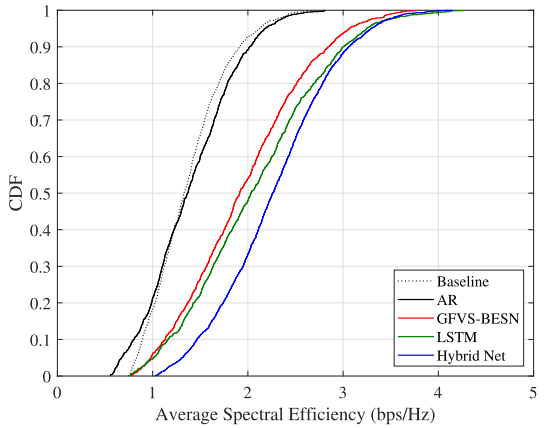


FIGURE 9. CDF of the average spectral efficiency in the vehicular scenario.

The performance was analyzed according to the root mean square error (RMSE) of the signal-to-interference-plus-noise ratio (SINR), which is used for CQI mapping and the false alarm probability of the CQI. Table 3 presents SINR and CQI obtained from the predicted channel given in Fig. 7. The proposed scheme had the highest prediction accuracy for both the SINR and CQI.

Figs. 8 and 9 show the cumulative distribution functions (CDF) of the average spectral efficiency for each run of the Monte Carlo simulation. The average spectral efficiency reflects the correctly detected data bits, as introduced by Seo *et al.* [41]. The correctly detected data bits were used as a criterion to demonstrate that the proposed scheme not only provides accurate CSI feedback compared to conventional schemes but is also more efficient in terms of resource usage. The average spectral efficiencies derived from correctly detected data bits were experimentally obtained for the two following cases: pedestrians including indoor and outdoor, and vehicles. Fig. 9 shows the results for UEs within 30 m of the BS. Although the RS overhead increased for exploiting the channel prediction scheme, the proposed scheme showed a higher spectral efficiency than the baseline.

TABLE 3. Performance according to SINR and CQI predictions.

		RMSE of SINR	False alarm prob. of CQI
Indoor, Ped.	Baseline	0.374	0.245
	AR	0.162	0.0225
	GFVS.	0.114	0.0155
	LSTM	0.0924	0.091
	Hybrid Net	0.0831	0.0047
Outdoor, Ped.	Baseline	0.856	0.365
	AR	0.952	0.377
	GFVS.	0.528	0.216
	LSTM	0.492	0.195
	Hybrid Net	0.328	0.109
Outdoor, Veh.	Baseline	1.724	0.496
	AR	3.527	0.512
	GFVS.	1.406	0.361
	LSTM	1.054	0.308
	Hybrid Net	0.912	0.272

Thus, the results showed that the proposed scheme for channel prediction improved the link adaptation performance.

V. CONCLUSION

We proposed a neural network-based channel prediction scheme and system operation algorithm for AMC that can be applied to customized communication systems. The proposed scheme utilizes a hybrid network to extract meaningful features for improving the prediction accuracy. The scheme also allows the channel prediction in different environments with a single network. Simulation results revealed that the proposed scheme improves the prediction accuracy for the CQI and correctly detected data bits per frame compared to existing schemes. Therefore, the proposed scheme can solve the problems caused by delayed CSI feedback and enable CSI feedback even for subframes without a CSI-RS. This can potentially be combined with the reinforcement learning-based method proposed by Mota *et al.* [42] to improve the link adaptation performance even further.

REFERENCES

[1] A. Dogra, R. K. Jha, and S. Jain, "A survey on beyond 5G network with the advent of 6G: Architecture and emerging technologies," *IEEE Access*, vol. 9, pp. 67512–67547, 2021.

[2] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J.-A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

[3] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, 2021.

[4] L. Xu, H. Wang, and T. A. Gulliver, "Outage probability performance analysis and prediction for mobile IoV networks based on ICS-BP neural network," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3524–3533, Mar. 2021.

[5] P. Yang, Y. Xiao, M. Xiao, and S. Li, "6G wireless communications: Vision and potential techniques," *IEEE Netw.*, vol. 33, no. 4, pp. 70–75, Jul./Aug. 2019.

[6] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 28–41, Sep. 2019.

[7] E. C. Strinati, S. Barbarossa, J. L. Gonzalez-Jimenez, D. Ktenas, N. Cassiau, L. Maret, and C. Dehos, "6G: The next frontier: From holographic messaging to artificial intelligence using subterahertz and visible light communication," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 42–50, Sep. 2019.

[8] D. L. Goeckel, "Robust adaptive coded modulation for time-varying channels with delayed feedback," in *Proc. 35th Annu. Allerton Conf. Commun., Control, Comput.*, Oct. 1997, pp. 370–379.
[9] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 844–855, Jun. 1999.
[10] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 173–196, 1st Quart., 2019.
[11] T.-S. Yang and A. Duel-Hallen, "Adaptive modulation using outdated samples of another fading channel," in *Proc. IEEE Wireless Commun. Netw. Conf. Rec.*, vol. 35, Mar. 2002, pp. 477–481.
[12] J. Heo, Y. Wang, and K. Chang, "A novel two-step channel-prediction technique for supporting adaptive transmission in OFDM/FDD system," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 188–193, Jan. 2008.
[13] V. H. Pham, X. Wang, and J. Nadeau, "Long term cluster-based channel envelope and phase prediction for dynamic link adaptation," *IEEE Commun. Lett.*, vol. 15, no. 7, pp. 713–715, Jul. 2011.
[14] A. Duel-Hallen, "Fading channel prediction for mobile radio adaptive transmission systems," *Proc. IEEE*, vol. 95, no. 12, pp. 2299–2313, Dec. 2007.
[15] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
[16] S. M. Kay, *Modern Spectral Estimation: Theory & Application*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1988.
[17] S. Semmelrodt and R. Kattenbach, "Investigation of different fading forecast schemes for flat fading radio channels," in *Proc. IEEE 58th Veh. Technol. Conf.*, vol. 1, Oct. 2003, pp. 149–153.
[18] W. Jiang and H. D. Schotten, "Neural network-based fading channel prediction: A comprehensive overview," *IEEE Access*, vol. 7, pp. 118112–118124, 2019.
[19] W. Liu, L.-L. Yang, and L. Hanzo, "Recurrent neural network based narrowband channel prediction," in *Proc. IEEE 63rd Veh. Technol. Conf.*, vol. 1, May 2006, pp. 2173–2177.
[20] W. Jiang and H. D. Schotten, "Multi-antenna fading channel prediction empowered by artificial intelligence," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Chicago, IL, USA, Aug. 2018, pp. 1–6.
[21] W. Jiang and H. D. Schotten, "Neural network-based channel prediction and its performance in multi-antenna systems," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Chicago, IL, USA, Aug. 2018, pp. 1–6.
[22] T. Ding and A. Hirose, "Fading channel prediction based on combination of complex-valued neural networks and chirp Z-transform," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 9, pp. 1686–1695, Sep. 2014.
[23] Y. Zhao, H. Gao, N. C. Beaulieu, Z. Chen, and H. Ji, "Echo state network for fast channel prediction in Ricean fading scenarios," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 672–675, Mar. 2017.
[24] W. Jiang and H. D. Schotten, "Deep learning for fading channel prediction," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 320–332, 2020.
[25] W. Jiang and H. D. Schotten, "Recurrent neural networks with long short-term memory for fading channel prediction," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring)*, Antwerp, Belgium, May 2020, pp. 1–5.
[26] W. Jiang and H. D. Schotten, "Recurrent neural network-based frequency-domain channel prediction for wideband communications," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC-Spring)*, Kuala Lumpur, Malaysia, Apr. 2019, pp. 1–6.
[27] C. Lv, J.-C. Lin, and Z. Yang, "Channel prediction for millimeter wave MIMO-OFDM communications in rapidly time-varying frequency-selective fading channels," *IEEE Access*, vol. 7, pp. 15183–15195, 2019.
[28] L. Ahrens, J. Ahrens, and H. D. Schotten, "Convolutional-type neural networks for fading channel forecasting," *IEEE Access*, vol. 8, pp. 193075–193090, 2020.
[29] L. Liu, H. Feng, B. Hu, and J. Zhang, "MIMO-OFDM wireless channel prediction by exploiting spatial correlation," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Huangshan, China, Oct. 2012, pp. 1–6.
[30] Y. Sui, Y. He, T. Cheng, Y. Huang, and S. Ning, "Broad echo state network for channel prediction in MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13383–13399, Nov. 2020.
[31] *3rd Generation Partnership Project (3GPP); Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations (Release 15)*, document TR 25.996, 3GPP, 2018.
[32] J. C. Ikuno, "System level modeling and optimization of the LTE downlink," Ph.D. dissertation, Dept. Elect. Eng., Vienna Univ. Technol., Vienna, Austria, 2013.
[33] *3rd Generation Partnership Project (3GPP); Study on 3D Channel Model for LTE (Release 12)*, document TR 36.873, 3GPP, 2018.
[34] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2001.
[35] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[36] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.
[37] Z. Wu, C. Shen, and A. van den Hengel, "High-performance semantic segmentation using very deep fully convolutional networks," 2016, *arXiv:1604.04339*. [Online]. Available: http://arxiv.org/abs/1604.04339
[38] R. Fantacci, D. Marabissi, D. Tarchi, and I. Habib, "Adaptive modulation and coding techniques for OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4876–4883, Sep. 2009.
[39] A. Sampath, P. Sarath Kumar, and J. M. Holtzman, "On setting reverse link target SIR in a CDMA system," in *Proc. IEEE 47th Veh. Technol. Conf. Technol. Motion*, Phoenix, AZ, USA, May s1997, pp. 929–933.
[40] B. Yuen, X. Dong, and T. Lu, "Inter-patient CNN-LSTM for QRS complex detection in noisy ECG signals," *IEEE Access*, vol. 7, pp. 169359–169370, 2019.
[41] B. Seo, D. Sim, T. Lee, and C. Lee, "Efficient time synchronization method with adaptive resource configuration for FBMC systems," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5563–5574, Sep. 2020.
[42] M. P. Mota, D. C. Araujo, F. H. Costa Neto, A. L. F. de Almeida, and F. R. Cavalcanti, "Adaptive modulation and coding based on reinforcement learning for 5G networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

CHAHYEON EOM received the B.S. degree in electrical and electronic engineering from Yonsei University, Seoul, South Korea, in 2017, where he is currently pursuing the Ph.D. degree in electrical and electronic engineering. His research interests include MIMO techniques, artificial intelligence, localization techniques, and sixth communication.

CHUNGYONG LEE (Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from Yonsei University, Seoul, South Korea, in 1987 and 1989, respectively, and the Ph.D. degree in electrical and computer engineering from Georgia Institute of Technology, Atlanta, GA, USA, in 1995. From 1996 to 1997, he was a Senior Engineer with Samsung Electronics Company Ltd., and Kiheung, South Korea. Since 1997, he has been with the School of Electrical and Electronic Engineering, Yonsei University, where he is currently a Professor. His research interests include array signal processing and communication signal processing.