

Learning the Wireless V2I Channels Using Deep Neural Networks

Tian-Hao Li¹, Muhammad R. A. Khandaker¹, Faisal Tariq², Kai-Kit Wong³ and Risala T. Khan⁴

¹School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, United Kingdom

²James Watt School of Engineering, University of Glasgow, United Kingdom

³Department of Electronic and Electrical Engineering, University College London, United Kingdom

⁴Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh

Corresponding e-mail: m.khandaker@hw.ac.uk

Abstract—For high data rate wireless communication systems, developing an efficient channel estimation approach is extremely vital for channel detection and signal recovery. With the trend of high-mobility wireless communications between vehicles and vehicles-to-infrastructure (V2I), V2I communications pose additional challenges to obtaining real-time channel measurements. Deep learning (DL) techniques, in this context, offer learning ability and optimization capability that can approximate many kinds of functions. In this paper, we develop a DL-based channel prediction method to estimate channel responses for V2I communications. We have demonstrated how fast neural networks can learn V2I channel properties and the changing trend. The network is trained with a series of channel responses and known pilots, which then speculates the next channel response based on the acquired knowledge. The predicted channel is then used to evaluate the system performance.

Index Terms—Vehicle-to-infrastructure; V2X; Channel estimation; Machine learning; Deep neural network.

I. INTRODUCTION

Since the first generation (1G) wireless communication network entered the market in the 1980s, the world has been dramatically changed by the development of mobile communication technologies. It has gone through several evolutions in the past few decades, from 1G to 5G, and due to the huge potential demand all over the world, the technology will continue to upgrade rapidly [1]. While 2G, 3G and 4G were about connecting people and parts of things, 5G will connect everything and it can provide unlimited access to anywhere, anytime, anybody and anything [2].

In recent years, orthogonal frequency division multiplexing (OFDM) has become a popular choice for fast-speed and high-quality communication systems. However, channel modelling and channel estimation are two major challenges affecting the performance of OFDM systems. In order to estimate the channel response, pilot-based channel estimation is commonly adopted, in which a training sequence composed of known data symbols (pilots) is transmitted, and the channel parameters are initially estimated using the received pilot signals [3]. Minimum mean-square error (MMSE) and least squares (LS) are two traditional estimation approaches.

Vehicle-to-Infrastructure (V2I) communication is about the data transmission between vehicles and infrastructure on the roads. V2I communication system is normally wireless and

two-way: infrastructure like traffic lights can provide information to cars and vice-versa. This communication system can provide quantitative and real-time information that can be used for safety, mobility, and environmental benefits. When V2I communications are widely used, revolution of roadways will occur all over the world. For example, self-driving vehicles will become reality and it is vital to make successful and safe autonomous cars. To make this huge change happen, the basic theory and techniques must now be developed urgently.

To enhance the performance of communication systems and to solve signal processing and communications problems, deep learning (DL) has recently drawn utmost popularity [4]–[7]. A deep neural network (DNN) is an algorithm with learning ability and optimization capability that can approximate many kinds of functions. Particularly, problems without any precise numerical model can now be solved using DL methods [8]. In order to leverage the advantages of using a large group of data for communication performance improvement, several machine learning methods including supervised, unsupervised and reinforcement learning have been proposed based on the traditional approaches. The machine learning can be useful in analyzing communication environment variance, making decisions autonomously, transmission routing, network security, and system resource management [6], [8].

Recently, some works have concentrated in the area of channel estimation using DL methods [9]–[13]. The authors in [10], [11] proposed the back propagation (BP) learning algorithm to build a multilayered perceptron (MLP) neural network as an estimator for OFDM communication channels. In [11], a method of implicit channel state information (CSI) estimation and direct recovery of transmission symbols based on deep learning has been proposed. The authors in [12] proposed an approximate information passing network for millimetre-wave massive multiple-input and multiple-output (MIMO) systems based on learned denoising which is a deep learning model that can analyze channel structure and estimate channel from a big training database. In [13], the authors developed a CSI feedback model by spreading a new learned CSI perception and restoration architecture. By learning the spatial features directly and combining the correlation of samples in time-varying MIMO channels, the system greatly improves the quality of recovery and the trade-off between the compression

ratio (CR) and the recovery quality [13].

All of the above works [9]–[13] have developed DL-based channel estimation techniques in traditional quasi-stationary wireless communication systems. However, with the trend of high-mobility wireless communications between vehicles and vehicles-to-infrastructure (V2I), many rising problems appear because of the high variances in communication environments, which are fundamentally different from traditional wireless communication problems. For the communication with large Doppler-shift and fast-varying channels, the above methods may not work because in fast time-varying fading environment channel response is difficult to obtain in real time and the outdated CSI exhibits a significant negative impact on the performance. The critical point to fit high-mobility networks is to develop future informatized and intelligent vehicles combined with machine learning to benefit the communication networks [14].

In this paper, we address the adverse effects of imperfect CSI on V2I communication systems by developing a DL-based channel prediction approach. Two major challenges facing V2I channel estimation addressed in this paper are *i*) the way to find CSI in real-time with the knowledge of vehicle information including position and velocity, *ii*) the technique to estimate the fast time-variant channel properly and recover the information with minimum error. The major contributions of this paper are as follows:

- Firstly, an OFDM modulation based V2I channel estimation technique is introduced as a baseline to verify effectiveness of the DL-based approach.
- Secondly, DL-based channel prediction method is proposed to deal with the channel estimation problem in V2I communication system in high-mobility environments.
- Thirdly, the performance of the proposed DL-based approach is compared with other common algorithms through extensive numerical simulations.

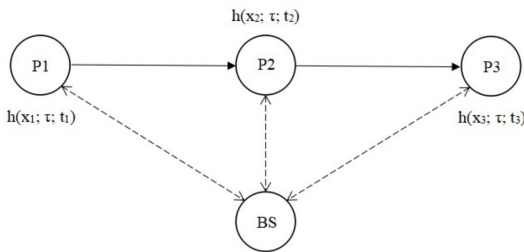


Figure 1. The proposed V2V and V2I communication system.

II. SYSTEM MODEL

We consider a high-mobility vehicular communication system (cf. Fig. 1) in which a vehicle intends to communicate with a roadside infrastructure, which can be a traditional cellular base station (BS), traffic lamppost, building or any other fixed structure. At both BS and the vehicle, only one antenna for transmitting and receiving is considered. For simplification, only 2D horizontal plane is considered where the vehicle

travels at a constant speed along a straight road. Hence, at the same position the channel is highly similar even at different time slots, and it can be considered as position related channel estimation which means that change of channel is related to the position of vehicle. The channel power gain between the vehicle at position n and the BS is defined as [15]

$$h(n) = G(n)\beta(n)Ad(n)^{-\gamma}, \quad (1)$$

where $G(n)$ is assumed to be exponentially distributed fast fading power gain, $\beta(n)$ is the log normal shadow fading component, A is the constant pathloss, $d(n)$ is the distance between the vehicle and the BS, and γ is the pathloss exponent. For a frequency-selective multipath fading channel, the channel impulse response (CIR) is given by

$$h(n) = \sum_{l=0}^{L-1} C_l e^{j\phi_l(n)} \delta(n - \tau_l) \quad (2)$$

$$= \alpha_l \delta(n - \tau_l), \quad (3)$$

where L is the total number of multipath component and ϕ_l , τ_l , and $\alpha_l = \sum_{l=0}^{L-1} C_l e^{j\phi_l(n)}$ are phase, time delay, and the time-dependent complex path gain of the l th multipath component. For the V2I channel, the time-variant phase associated with the l th path is defined as [16, Chapter 2]

$$\phi_l(n) = \phi_l - 2\pi c\tau_l/\lambda_c + 2\pi f_{D,l}n, \quad (4)$$

where $c = f_c\lambda_c$ is the speed of light, λ_c is the wavelength of the arriving plane wave (carrier signal), $\tau_l = d_l/c$, d_l is the length of the l th path, and $f_{D,l}$ is the maximum Doppler shift of the l th propagation path. For OFDM transmission, the m th modulated time domain symbols can be expressed as

$$x_m(n) = \frac{1}{\sqrt{N}} \sum_{l=1}^{N-1} X_m(l) e^{j2\pi nl/N}, \quad n = 0, \dots, N-1, \quad (5)$$

where $X_m(l)$ is the m th modulated symbol, n is the subcarrier index and N is the total number of subcarriers. Cyclic prefix is then appended to the symbols to prevent inter-symbol interference (ISI). The signals are then transmitted through the wireless channel and the signal received at the BS can be expressed as

$$y_m(n) = h_m(n) \otimes x_m(n) + w_m(n), \quad (6)$$

where \otimes indicates the convolution operation and $w_m(n)$ is the additive white Gaussian noise for the m th subcarrier. At the receiver, the cyclic prefix is removed first from the received signal, followed by parallel conversion to frequency domain by applying fast Fourier transform:

$$Y_m(k) = \frac{1}{N} \sum_{n=1}^{N-1} y_m(n) e^{-j2\pi kn/N}, \quad k = 0, \dots, N-1. \quad (7)$$

Thus in the frequency domain, the input-output relationship can be expressed as

$$Y(k) = H(k)X(k) + W(k), \quad k = 0, \dots, N-1. \quad (8)$$

Consequently, the system can be described as a set of N independent parallel Gaussian channels:

$$y_k = h_k x_k + w_k, \quad k = 0, \dots, N-1. \quad (9)$$

For convenience, we rewrite (10) using matrix notations as [3]

$$\mathbf{y} = \mathbf{X}\mathbf{F}\mathbf{g} + \mathbf{w}, \quad (10)$$

where \mathbf{X} is a $N \times N$ diagonal matrix containing $\mathbf{x} \triangleq [x_m(0) \ x_m(1) \ \dots \ x_m(N-1)]^T$ as the main diagonal, $\mathbf{g} \triangleq [g_m(0) \ g_m(1) \ \dots \ g_m(N-1)]^T$, $\mathbf{w} \triangleq [w_0 \ w_1 \ \dots \ w_{N-1}]^T$ and

$$\mathbf{F} \triangleq \begin{bmatrix} F_{0,0} & \cdots & F_{0,N-1} \\ \vdots & \ddots & \vdots \\ F_{N-1,0} & \cdots & F_{N-1,N-1} \end{bmatrix}, \quad (11)$$

is the DFT matrix with $F_{n,k} = \frac{1}{\sqrt{N}} e^{-j2\pi nk/N}$.

The objective of this study is to estimate the CIR $H_m(k)$ from the observation of $Y_m(k)$ with known pilot signal $X_m(k)$. In the following, we will first discuss the conventional channel estimation approaches, and then develop a deep learning based channel estimation algorithm for the V2I system (7).

III. CONVENTIONAL CHANNEL ESTIMATION TECHNIQUE

In conventional communication systems, both blind and non-blind channel estimation techniques have been considered for estimating the CIR. Popular channel estimation algorithms include maximum likelihood (ML), least mean square (LMS), minimum mean square error (MMSE) and least square (LS) methods. These methods have been studied thoroughly to estimate CSI within a certain time and frequency range. Although the LS algorithm has a worse performance in time-varying environments compared to the other approaches, its implementation is very simple. On the other hand, the MMSE algorithm can be well-behaved in all general fading channels with both frequency and time selectivity.

A. The LS Algorithm

The LS algorithm tries to minimize the squared error (i.e., Euclidian distance) between the transmitted and the received signals which is expressed as

$$\|\mathbf{y} - \mathbf{X}\mathbf{F}\mathbf{g}\|^2 = (\mathbf{y} - \mathbf{X}\mathbf{F}\mathbf{g})^H (\mathbf{y} - \mathbf{X}\mathbf{F}\mathbf{g}). \quad (12)$$

Taking the derivative of (12) and equating the derivative to 0, the LS estimate of the channel frequency response is given by

$$\hat{\mathbf{h}}_{\text{LS}} = \mathbf{F} (\mathbf{F}^H \mathbf{X}^H \mathbf{X} \mathbf{F})^{-1} \mathbf{F}^H \mathbf{X}^H \mathbf{y}. \quad (13)$$

Since the optimal training sequence is orthonormal [17], [18], (13) eventually reduces to

$$\hat{\mathbf{h}}_{\text{LS}} = \mathbf{X}^{-1} \mathbf{y}. \quad (14)$$

B. The MMSE Algorithm

The MMSE channel estimator takes the effect of channel noise into consideration. Meanwhile, the MMSE estimate of \mathbf{g} is given by

$$\hat{\mathbf{g}}_{\text{MMSE}} = \mathbf{R}_{gy} \mathbf{R}_{yy}^{-1} \mathbf{y} \quad (15)$$

where

$$\mathbf{R}_{gy} = \mathbb{E} \{ \mathbf{g} \mathbf{y}^H \} = \mathbf{R}_{gg} \mathbf{F}^H \mathbf{X}^H, \quad (16)$$

$$\mathbf{R}_{yy} = \mathbb{E} \{ \mathbf{y} \mathbf{y}^H \} = \mathbf{X} \mathbf{F} \mathbf{R}_{gg} \mathbf{F}^H \mathbf{X}^H + \sigma_w^2 \mathbf{I}_N \quad (17)$$

$$(18)$$

are the corresponding covariance matrices and σ_w^2 is the noise variance $\mathbb{E} \{ |w_k|^2 \}$. Since the columns of \mathbf{F} in (11) are orthonormal, the frequency-domain MMSE estimate of \mathbf{h} is given by [3]

$$\hat{\mathbf{h}}_{\text{MMSE}} = \mathbf{F} \hat{\mathbf{g}}_{\text{MMSE}} = \mathbf{F} \mathbf{R}_{gg} \left[(\mathbf{F}^H \mathbf{X}^H \mathbf{X} \mathbf{F})^{-1} \sigma_w^2 + \mathbf{R}_{gg} \right]^{-1} \times (\mathbf{F}^H \mathbf{X}^H \mathbf{X} \mathbf{F})^{-1} \mathbf{F}^H \mathbf{X}^H \mathbf{y}. \quad (19)$$

IV. PROPOSED MACHINE LEARNING APPROACH

It has been shown in [11] that a multi-layer neural network (MLNN) can provide a very good approximation of channels. Hence in the following, we propose a deep learning based V2I channel estimation algorithm using back propagation technique.

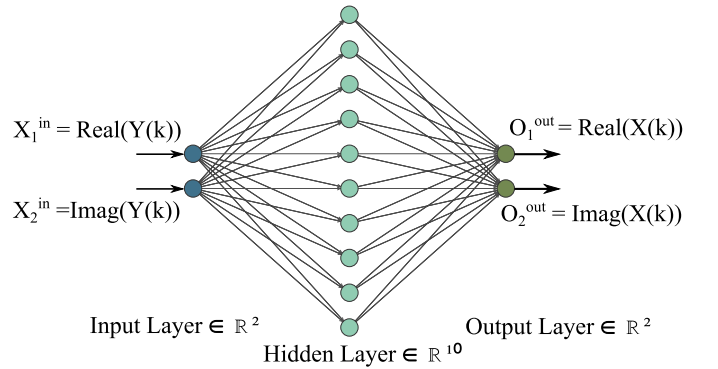


Figure 2. The proposed deep neural network for estimating the V2I channel.

A. Defining the Neural Network

As shown in Fig. 2, the neural network contains two inputs and two outputs and n_h hidden nodes. The inputs to the network are the received signals, and the estimated channel response parameters are the outputs. The two inputs and outputs of the MLNN are connected to the real and imaginary components of the corresponding complex numbers, since neural networks work only with real numbers.

For the machine learning based channel estimation scheme, we consider a supervised learning approach which estimates the CIRs using a fully connected neural network. Note that we do not consider any bias inputs to the neurons at any layer. We apply sigmoid function for the activation of the hidden

neurons, while the activation function at the output layer is linear. The output of each neuron is generated by computing the weighted sum of the inputs coming into that node and then applying the activation function.

Let us now assume that $w_{i,j}^{(1)}$ is the i th input to j th hidden neuron weight and $w_{j,k}^{(2)}$ is the j th hidden node to the k th output node weight. The hidden layer activation function is the sigmoid function

$$\Phi(z) = \frac{1}{1 + e^{-z}} = \frac{e^z}{e^z + 1}. \quad (20)$$

Thus, the hidden layer output is defined as

$$O_j^h = \Phi(\text{in}_j) = \frac{1}{1 + e^{-\text{in}_j}}, \quad j = 1, \dots, n_h, \quad (21)$$

where $\text{in}_j = \sum_{i=1}^{n_{in}} X_i^{\text{in}} w_{i,j}^{(1)}$, n_{in} is the number of input layer nodes and X_i^{in} is the i th input to the DNN. Similarly, each output node computes its net output as

$$O_k^{\text{out}} = \Phi(\text{in}_k) = \frac{1}{1 + e^{-\text{in}_k}}, \quad k = 1, 2, \quad (22)$$

where $\text{in}_k = \sum_{j=1}^{n_h} O_j^h w_{j,k}^{(2)}$. Then, the sum of weighted inputs to the nodes is applied to the output layer activation function which is assumed to be a linear function. Thus the network outputs can be calculated as

$$O_k^{\text{out}} = \Phi \left(\sum_{j=1}^{n_h} w_{j,k}^{(2)} \Phi \left(\sum_{i=1}^{n_{in}} w_{i,j}^{(1)} X_i \right) \right), \quad k = 1, 2. \quad (23)$$

B. Training the Neural Network

In training process, the weights of all the layers are adjusted according to the mismatches between the outputs and the targets. In each epoch, the direction of changes are that tending to minimize the mean-squared error (MSE). Accordingly, we define the cost function as the MSE between DNN output and the target output as

$$\varepsilon = \frac{1}{2} \sum_{k=1}^{n_o} (t_k - O_k^{\text{out}})^2, \quad (24)$$

where t_k is the k th desired output and n_o is the number of outputs (2 in this case). The gradient descent based back propagation learning rule is exploited to optimize the weights for improving the DNN performance. The gradient descent method minimizes the cost function ε by updating the weights in the opposite direction of the gradient of the objective function w.r.t. to the weights [19]. Thus the weight update in each iteration is given by

$$\nabla w = -\eta \frac{\partial \varepsilon}{\partial w}, \quad (25)$$

where $0 < \eta < 1$ is the learning rate that determines the size of the steps in each iteration. Thus the update of hidden-to-output-layer weights can be expressed as

$$\nabla w_{jk} = \eta (t_k - O_k^{\text{out}}) \frac{\partial O_k^{\text{out}}}{\partial w_{jk}} O_j^h. \quad (26)$$

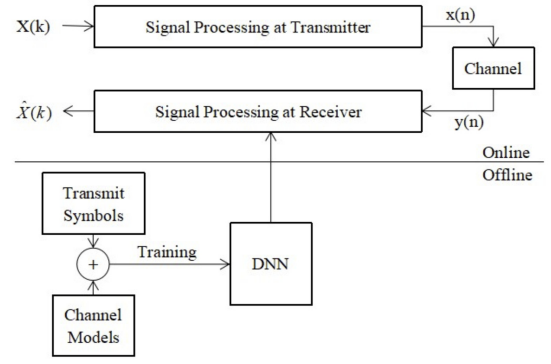


Figure 3. Proposed DNN-based channel estimation method.

Note that the gradient of output layer activation function is always unity. Similarly, the back propagated updates of the input-to-hidden-layer weights can be expressed as

$$\nabla w_{jk} = \eta \left[\sum_{k=1}^{n_o} w_{jk} (t_k - O_k^{\text{out}}) \frac{\partial O_k^{\text{out}}}{\partial w_{jk}} \right] \frac{\partial O_j^h}{\partial w_{ij}} X_i. \quad (27)$$

For training the DNN, we first generate a large set of pilot symbols following a particular modulation scheme. We then use a series of historical channel responses between moving vehicle and BS for optimizing the weights of the neural network. The training process continues until the target accuracy is reached. The overall training procedure is summarized in Fig. 3.

V. NUMERICAL SIMULATIONS

In this section, we perform numerical simulations to demonstrate the effectiveness of the proposed machine learning based channel estimation method for wireless V2I communications. Throughout this section, we compare the performance of the proposed approach against the conventional LS and MMSE based estimation schemes in [3]. We first demonstrate the estimation accuracy of the DNN method considering a low-mobility environment. We then test the trained DNN performance for V2I channel estimation.

We assume that the multipath V2I channel has time delays and data is transmitted using 4-QAM modulation. The neural network is designed to have a single hidden layer with 10 neurons and single output layer with two nodes. Sigmoid and linear activation functions has been utilized in the hidden layer and the output layer, respectively. The DNN algorithm is implemented in two ways. The first one is built from scratch using the structure illustrated in Section IV (identified as ‘build’ in the figures), and the second one is implemented based on the ANN toolbox in MATLAB, whose performance has also been compared in simulations where applicable.

Fig. 4 shows the performance of the proposed DNN approach in the training process according to normalized MSE at different SNR environment. We run for 20 epochs for each training data set and It is clear that the error between network outputs and the targets is being minimized as the training progresses. It can also be found that with higher SNR, which

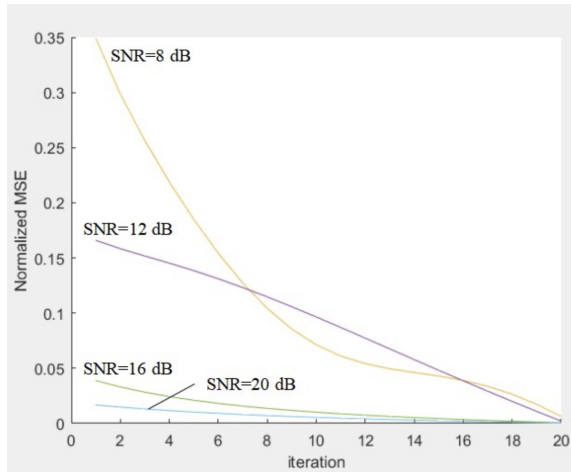


Figure 4. MSE performance during the training process.

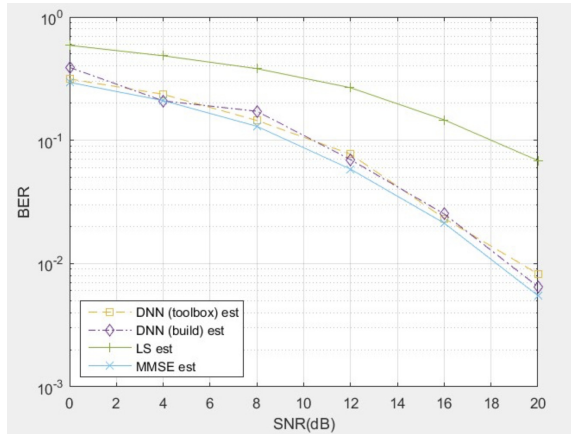


Figure 5. BER performance during the training process.

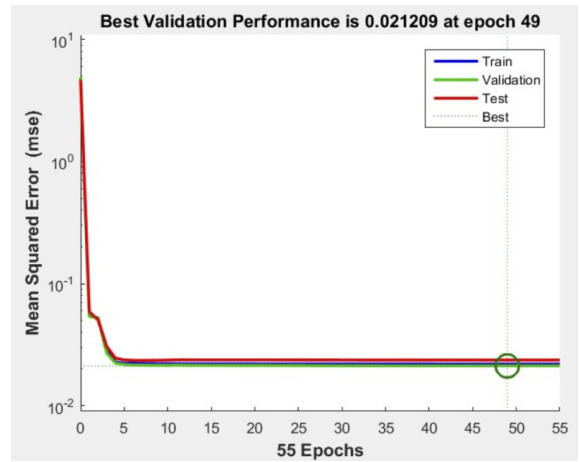


Figure 6. MSE at different phases of the DNN approach.

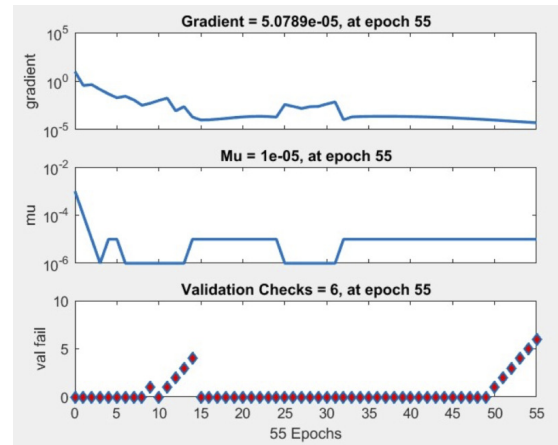


Figure 7. Training progression over different epochs.

means the noise power is low compared with signal power, the error of the estimated channel response becomes lower.

Then, Fig. 5 compares the bit error rate (BER) performances of all the methods. It can be seen that the LS algorithm results in the worst behaviour compared to the others and it cannot correctly estimate the channel especially when SNR is very low. From this figure it is clear that the performance of proposed back propagation DNN approaches are much better than the LS algorithm at low SNR values as well as high SNR values. The results also show that both DNN approaches perform close to each other in a low-mobility environment, and are very close to the MMSE algorithm. Obviously, the MMSE algorithm performs the best according to the simulation results.

Next, simulations of a V2I wireless communication system in a fast time-varying fading channel are carried out to demonstrate the performances of different channel estimation methods. Figs. 6 to 9 show the training process of DNN for V2I channel prediction. In these simulations, the entire data set is divided into three subsets namely: training (70%), validation (15%) and testing (15%), respectively. During the training process, the MSE reduces sharply and the best estimate is

achieved at the 49th epoch with the minimum error distribution. Fig. 8 shows the histogram of each subset. From Fig. 9, it can also be seen that the outputs fit the targets very well.

Finally, we plot the BER performance of the algorithms in Fig. 10. It can be seen from the results that the BER performance of the traditional DNN methods is significantly worse than the other methods. The reason for this is that the channel estimation in V2I communication is influenced by many factors, such as fast variation of channels, large Doppler-shift, high interference and noise. These results further justify the effectiveness of the proposed DL-based approach for V2I channel estimation.

VI. CONCLUSION

We have introduced a deep learning based channel estimation approach for V2I wireless communication systems. We have demonstrated that the channels in high-mobility environment can be estimated using DNN based prediction methods with a group of historical CIR to solve the outdated CSI problem. Extensive simulation results illustrate that the proposed channel prediction method is able to dramatically

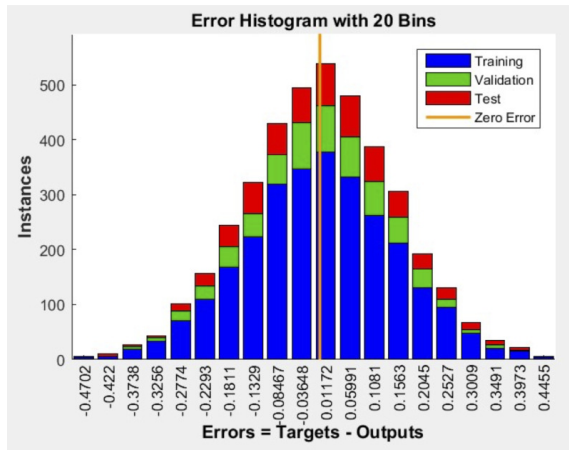


Figure 8. Histogram of errors at different processes.

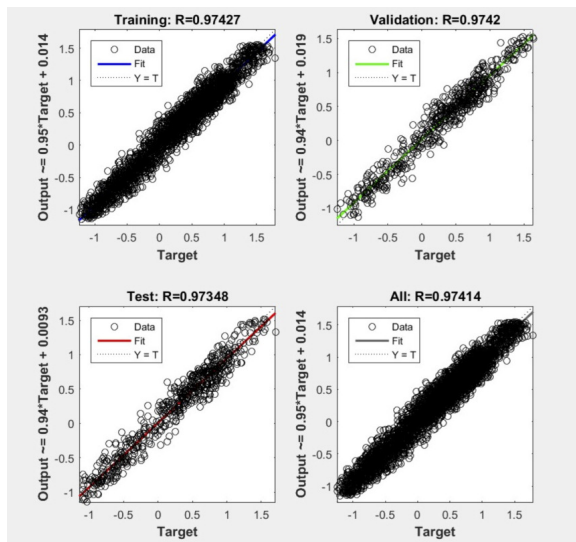


Figure 9. Training regression.

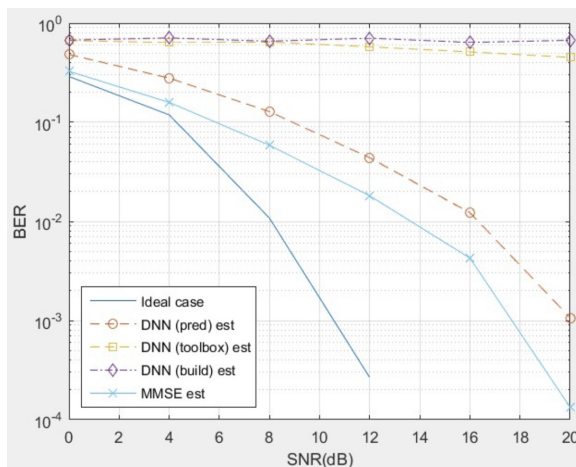


Figure 10. BER versus SNR for the testing data.

improve the performance of channel estimation in particular in high-mobility environment. Considering the nonuniform movement of vehicles, including variant position and changing velocity in the training process of DNN can be an interesting future work.

REFERENCES

- [1] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. Imran, M. Bennis, and M. rouane Debbah, "A speculative study on 6G," *IEEE Commun. Magazine*, June 2019 (submitted). Available: <https://arxiv.org/pdf/1902.06700.pdf>.
- [2] J. G. A. et al., "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, pp. 1065–1082, June 2014.
- [3] 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.*, vol. 2, Chicago, IL, USA, 1995, pp. 815–819.
- [4] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, , and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, pp. 98–105, 2016.
- [5] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys & Tutorials*, 2019.
- [6] J. Gao, M. R. A. Khandaker, F. Tariq, K.-K. Wong, and R. T. Khan, "Deep neural network based resource allocation for V2X communications," in *IEEE 90th Vehicular Technology Conference: VTC2019-Fall*, 22-25 Sep. 2019, Honolulu, Hawaii, USA (submitted). Preprint arXiv:1906.10194.
- [7] H. Ye and G. Y. Li, "Deep reinforcement learning for resource allocation in V2V communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kansas City, MO, 2018.
- [8] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," Available online: <https://arxiv.org/abs/1808.02342>, 2018.
- [9] M. Mehrabi, M. Mohammadkarimi, M. Ardakani, and Y. Jing, "Decision directed channel estimation based on deep neural network k-step predictor for MIMO communications in 5G," Available: <https://arxiv.org/abs/1901.03435>, Jan. 2019.
- [10] N. Taspinar and M. N. Seyman, "Back propagation neural network approach for channel estimation in OFDM system," in *Proc. Wireless Commun., Netw. Inf. Security (WCNIS)*, June 2010, pp. 265–268.
- [11] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, pp. 114–117, Feb. 2018.
- [12] H. He, C. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmWave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, pp. 852–855, Oct. 2018.
- [13] T. Wang, C. Wen, S. Jin, and G. Y. Li, "Deep learning-based CSI feedback approach for time-varying massive MIMO channels," *IEEE Wireless Commun. Lett.*, vol. 8, pp. 416–419, Apr. 2019.
- [14] L. Liang, H. Ye, and G. Y. Li, "Toward intelligent vehicular networks: A machine learning framework," *IEEE IoT Journal*, vol. 6, pp. 124–135, Feb. 2019.
- [15] L. Liang, G. Y. Li, and W. Xu, "Resource allocation for D2D-enabled vehicular communications," *IEEE Trans. Commun.*, vol. 65, pp. 3186–3197, July 2017.
- [16] G. L. Stüber, *Principles of Mobile Communication*, 4th ed. Springer, 2017.
- [17] Y. Rong, M. R. A. Khandaker, and Y. Xiang, "Channel estimation of dual-hop MIMO relay systems using parallel factor analysis," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 2224–2233, Jun. 2012.
- [18] Y. Rong and M. R. A. Khandaker, "Channel estimation of dual-hop MIMO relay system via parallel factor analysis," in *Proc. 17th Asia-Pacific Conf. Commun. (APCC'2011)*, Sabah, Malaysia, Oct. 2-5, 2011.
- [19] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint*. Available: [arXiv:1609.04747](https://arxiv.org/abs/1609.04747), June 2017.