IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 11, 2019, accepted June 29, 2019, date of publication July 5, 2019, date of current version August 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2926843

A CNN-Based End-to-End Learning Framework Toward Intelligent Communication Systems

NAN WU, (Member, IEEE), XUDONG WANG, (Member, IEEE), BIN LIN, (Member, IEEE), AND KAIYAO ZHANG

School of Information Science and Technology, Dalian Maritime University, Dalian 116023, China

Corresponding author: Bin Lin (binlin@dlmu.edu.cn)

ABSTRACT Deep learning has been applied in physical-layer communications systems in recent years and has demonstrated fascinating results that were comparable or even better than human expert systems. In this paper, a novel convolutional neural networks (CNNs)-based autoencoder communication system is proposed, which can work intelligently with arbitrary block length, can support different throughput and can operate under AWGN and Rayleigh fading channels as well as deviations from AWGN environments. The proposed generalized communication system is comprised of carefully designed convolutional neural layers and, hence, inherits CNN's breakthrough characteristics, such as generalization, feature learning, classification, and fast training convergence. On the other hand, the end-to-end architecture jointly performs the tasks of encoding/decoding and modulation/demodulation. Finally, we provide the numerous simulation results of the learned system in order to illustrate its generalization capability under various system conditions.

INDEX TERMS Convolutional neural network, end-to-end learning, autoencoder, communication systems.

I. INTRODUCTION

A. MOTIVATION

The fundamental problem of communication is that of finding a representation of a message, which is resilient to the channel impairments, so that it can be recovered perfectly at the other end. In order to meet this challenge, transmitter and receiver are divided into subtasks, such as, source coding, channel coding, modulation and equalization. This design method [1] has the advantage that each component can be optimized separately, leading to the reliable modular communication system as of today.

Deep Learning (DL) is one of the latest trends in the field of machine learning and artificial intelligence. DL methods have already brought revolutionary advances in computer vision and natural language processing [2]. In the area of communications, DL has been used in modulation [3], channel estimation [4]–[6], signal detection [7], [8], modulation recognition [9], [10], and channel decoding [11]. DL can also be applied to non-orthogonal multiple access schemes [12], which are essential for 5G wireless communication systems. Especially, DL based communication systems [13] are capable optimizing all the components of transceivers jointly using the concepts of autoencoders. In the pioneer work of [13], the communication system design was treated as an end-to-end reconstruction optimization task that sought to jointly optimize transmitter and receiver components, so that encoding and decoding were facilitated by learned weights of neural networks, rather than specially designed expert codes. In [14], convolutional neural layers [15] were used as the building blocks of the autoencoder-based communication systems, which has no restriction on the length of the input bit sequence. However, the performance of [14] suffered from an error floor in the high SNR regime. Furthermore, the Convolutional Neural Networks (CNNs) in [14] were designed from a machine learning perspective, which lacked of communication engineering insights, hence the network's generalization ability was limited.

Motivated by the desire of deep integration of the expertise of communications and the breakthrough abilities of CNNs, in this paper, we propose a novel CNN based autoencoder communication system, which can generalize to arbitrary block length, various throughput, different channel usage and environments. The proposed architecture can be trained easily and converge quickly at a specific SNR point, while working properly on the whole SNR range. The vision is that communication algorithms will be represented by learned weights of neural networks optimized using end-to-end

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

loss functions. The resultant intelligent communication systems shall work optimally under any communication requirements, i.e., cellular systems and underwater communications.

B. PREVIOUS WORK

Broadly speaking, there are two approaches to apply DL in the physical layer of communication systems. Namely, to use DL to replace a certain block of communication systems, or to treat the communication system design as an end-to-end autoencoder. In the former path, most notably, in [16]–[18], Recurrent Neural Network (RNNs) were employed at the receiver to decode channel-coded information bits, such as, convolutional codes, Turbo codes, and Polar codes. RNNs' ability to extract information from long sequence [17], [18] proved that neural networks can master the intricate structure of human designed codes with limited training data, while tolerating channel turbulence.

On the other hand, the concept of interpreting a communication system as an autoencoder was originated in [19]. The authors demonstrated that it was possible to jointly optimize transmitter and receiver components under a single neural network framework. The simple yet powerful autoencoder system [19] evolved further in [13] with the help of radio transformer networks [20] as a means to combine expert domain knowledge in a DL model. Reference [24] demonstrated that the end-to-end neural network schemes can be implemented on a hardware system. Moreover, [21] extended the original autoencoder system to Multiple Input Multiple Output (MIMO) communication systems and transmission throughput can be improved as shown in [22]. Furthermore, authors in [23] proposed a deep learning-enabled millimeter wave massive MIMO framework for effective hybrid precoding. Besides autoencoder system's appealing concept. However, training such a system over actual wireless channels becomes problematic, when the actual channel's gradients are missing. In order to tackle this issue, authors in [25] proposed an algorithm to iterate between supervised training of the receiver and Reinforcement Learning (RL) based training of the transmitter. Also, a number of adversarial approaches for channel response approximation and information encoding were presented in [26]–[28]. These adversarial schemes considered the actual channel as a black-box and used the structure of Generative Adversarial Networks (GANs) [29] to capture the analytic representation.

C. CONTRIBUTIONS

Inspired by autoencoder system's promising representation and classification capabilities, in this paper, we propose a novel CNN-based Autoencoder (CNN-AE) communication system, which builds on top of CNN's architecture, while integrating communication engineering expertise. More specifically, the contributions are summarized as follows:

1) The carefully designed CNN layers allow the learned network to have generalization capability while achieving optimal Block Error Rate (BLER) performance, i.e., $\Pr(\hat{s} \neq s)$, namely support any input bit length, work with flexible data rate, suitable for AWGN and Rayleigh as well as non-AWGN channels.
2) The proposed end-to-end learning system can be trained at a specific E_b/N_0, while working across the whole range. Also, the proposed CNN-AE system can converge quickly with a small number of epochs.
3) We also propose a differential coding version of the CNN-based autoencoder system, where no Channel State Information (CSI) at the receiver is required.
4) The CNN architecture is designed from a communication engineering perspective, hence theoretical explanation of the network structure and the resultant simulation results become possible.

The rests of the paper are organized as follows: Section II introduces the system architecture and settings, while Section III shows the proposed system's performance under AWGN and flat Rayleigh fading channels. Next, the proposed network's adaptivity is further examined in Section IV. Finally, conclusions are drawn in Section V.

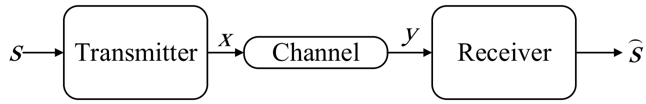

FIGURE 1. Illustration of a simple communications system.

II. END-TO-END LEARNING OF COMMUNICATION SYSTEMS

A. SYSTEM ARCHITECTURE

A simple communications system consisting of a transmitter, a channel and a receiver, as shown in Fig.1 [19]. The transmitter selects a symbol s, containing k information bits, to communicate to the receiver over a channel. More explicitly, the transmitter's job is to apply a transformation to the symbol s so that the generated transmitted signal $x \in \mathbb{C}^n$ of Fig. 1 occupies n channel time slots. The transmit vector $x \in \mathbb{C}^n$ is subjected to power constraints, e.g.,$\|x\|^2 \leq n$. At the receiver side, a noisy and possibly distorted version $y \in \mathbb{C}^n$ of x can be observed. Hence, the task of the receiver is to produce the estimate $\hat{s}$ of the original symbol s as closely as possible. Therefore, the code rate of the system is $R = k/n$*(bits/channel use)*.

Based on the architecture of Fig 1, we propose a CNN-based autoencoder learning system, as shown in Fig. 2, where the transmitter and the receiver are comprised of CNN layers, which are jointly optimized.

More specifically, the transmitter of Fig.2 consists of three 1-Dimensional Convolutional (Conv1D) layers followed by a power normalization layer that ensures that power constraints on the transmitted signal $\mathbf{X}$ are met. Although more Conv1D layers could potentially increase the representation and classification power of the neural network, deep networks also lead to gradient exploding and more parameters

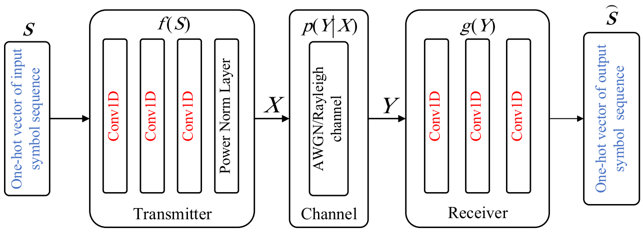

FIGURE 2. Represent a self-learning communication system as an autoencoder. The trainable layers are written in red, non-trainable layers are written in black.

TABLE 1. Autoencoder structure of fig. 2 (note: each Conv1D layer is followed by a batch normalization layer before activation).

Layer	Activation	Output dimensions
One-hot input	None	$L \times 2^k$
Conv1D	elu	$L \times 256$
Conv1D	elu	$L \times 256$
Conv1D	linear	$L \times 2n$
Power Norm Layer	None	$L \times 2n$
AWGN/ Rayleigh channel	None	$L \times 2n$
Conv1D	elu	$L \times 256$
Conv1D	elu	$L \times 256$
Conv1D	softmax	$L \times 2^k$

to train. In our work, three layers were found to be adequate to achieve the best possible BLER performance without losing any learning ability.

Furthermore, the Conv1D layers allow the transmitter to process a sequence of symbols S, rather than one symbol at a time as that of Fig. 1. Hence, a total number of $k \times L$ bits can be handled simultaneously, where k is the number of bits per symbol and L is the number of symbols (block length). Furthermore, each symbol of input sequence S is converted to a one-hot vector, which means the proposed system is aiming for minimizing BLER. The detailed parameters of the proposed CNN-based learning system are summarized in Table 1.

From the channel coding perspective, the Conv1D layers at the transmitter of Fig.2 facilitate linear/non-linear block encoding of the input symbol sequence, where convolutional operations enable linear coding and Exponential Linear Unit (ELU) activation functions of Table 1 enable potential non-linear coding. From modulation perspective, the Conv1D layers at the transmitter of Fig.2 transform the one-hot input symbol sequence to a new signal representation $\mathbf{X} = f(S)$, which occupies n channel use. In other words, signal constellation points are designed in $2n$-*dimensional* space. The idea is similar to Sphere Packing modulation [30], but the Conv1D layers can jointly optimize for higher $2n$-*dimensional* space. In the eyes of deep learning, each Conv1D layer has $d = 256$ filters as shown in Table 1, which allows mapping each one-hot vector from 2^k-*dimensional* space to 256-*dimension* space and search for the most suitable representation of the input symbol under a given communication channel so that the receiver can reconstruct the information. Note that each transmission makes use of n channel slots. Consider each time slot has I/Q channels, the Power Normalization layer of Fig.2 are required to compress 256-*dimension* representation of the symbols to $2n$-*dimensional* space signal of $\mathbf{X}$.

The channel layer of Fig.2 can be described as the conditional probability density function $p(\mathbf{Y}|\mathbf{X})$. Furthermore, an additive Gaussian white noise with a fixed variance $\sigma_1 = (2RE_b/N_0)^{-1}$ is added to the signals. For flat Rayleigh fading channels, $\mathbf{X}$ need to convolve with the channel impulse responses before arriving at the receiver.

The task of the receiver of Fig.2 is to classify each of the received signal $\mathbf{Y}$ out of 2^k possibilities based on the learned symbol characteristics. At the receiver side, the Conv1D layers of Fig.2 firstly decompress the received signal $\mathbf{Y}$ back to 256-*dimension* space in order to extract adequate information for classification. Hence, the following Conv1D layers are capable of capture the intricate signal representations. Finally, the signal is mapped to a one-hot vector length of 2^k for soft decision using soft-max activation of Table 1. The whole learning and transformation process are denoted using $g(\mathbf{Y})$ in Fig.2. Note that perfect CSI is assumed at the receiver, and is fed into the receiver network together with the received signal $\mathbf{Y}$.

B. SYSTEM PARAMETERS

The training and validation datasets are generated randomly using i.i.d. binary bit sequences, where 0 and 1 are submitted to uniform distribution. The proposed autoencoder system of Fig.2 was trained using 12800 data messages, where each message contains a block length of L symbols and each symbol has k information bits. The network was tested using 64000 data messages, while batch size was set to 64.

Because Conv1D layers have the merit of weight sharing, the proposed system of Fig.2 can be trained at any input symbol sequence length (L). In other words, the very same convolution operations are performed on every input symbol, regardless how many symbols are fed to the network at a time. For the purpose of illustration, L was set to 100 for the rest of the paper.

CNN's generalization ability allows us to train the CNN-based system at a fixed E_b/N_0, while testing at a wide range of SNRs. More explicitly, the CNN-based receiver of Fig.2 needs to witness sufficient statistics samples around the decision boundary to learn the signal representation under the influence of channel impairments. If E_b/N_0 is set at a relatively small value, the receiver might learn nothing but the noise. On the other hand, if the training E_b/N_0 is relatively high, the receiver's CNNs can only learn the perfect signal representation, and any channel induced corruption would lead to false classification. Again, a proper training E_b/N_0 value must facilitate the CNNs to have enough training samples in the vicinity of the decision boundary.

The CNN's generalization ability also blessed the proposed system to achieve optimal encoding and decoding in terms of BLER, when communicating under AWGN, Rayleigh and non-standard channels. This kind of adaptivity can be achieved because the CNNs at the transmitter are capable of learning suitable representation of the symbols for each channel setting, and the CNNs at the receiver can learn to discrimination the received signals accordingly.

Moreover, the loss function is defined as the Binary Cross Entropy (BCE) between the input symbol sequence and the output symbol sequence of Fig. 2, both of which are converted to one-hot vectors. The Adam optimized was used to train the end-to-end system, where the learning rate was set to be 0.001 and decayed by a factor of 10 when saturated. 50 epochs were used for training, since the proposed scheme can converge quickly, with the help of batch normalization layers after each Conv1D layer in Fig. 2.

For all the Conv1D layers in Table 1, the kernel size is set to 1 and stride is 1. This implies that each input symbol is handled individually, since the symbols are memoryless. Note that although each input sequence has $2^{k\times L}$ possibilities, the proposed CNN-based model was trained using merely 12800 messages. That means, the network can generalize to decode unseen codewords.

III. SYSTEM PERFORMANCE

In this section, numerous simulation results were provided in order to demonstrate the proposed CNN-based Autoencoder (CNN-AE) system's generalization capacity for block length, training E_b/N_0, code rate, channel use, when communicating over AWGN and flat Rayleigh fading channels. The system model of Fig. 2 was employed and all the neural network parameters were given in Table 1 and Section II.B, unless otherwise specified. Besides, our source codes were implemented in Keras and is available on GitHub upon publication.[1]

A. GENERALIZATION WITH BLOCK LENGTH AND TRAINING E_b/N_0

Firstly, we trained the CNN-AE system of Fig.2 using a short block length ($L = 10$) and tested the BLER over different block length ($L = 10, 100$) using the same trained network parameters.

Fig. 3 explicitly demonstrate that when testing over $L = 10, 100$ block length, the CNN-AE scheme achieves identical BLER performance, thanks to the weight sharing characteristics of the Conv1D layers. Therefore, the proposed system can be generalized to work with any block length L.

Furthermore, the CNN-AE system was trained at a single E_b/N_0, while testing over the whole E_b/N_0 range. As shown in Fig.3 that when trained at $E_b/N_0 = 9$dB, the proposed CNN-AE scheme achieves the best BLER performance, compared with trained at 3 or 20dB. Again, that is because the CNN layers of the receiver must observe enough training samples near the decision boundary in order to learn

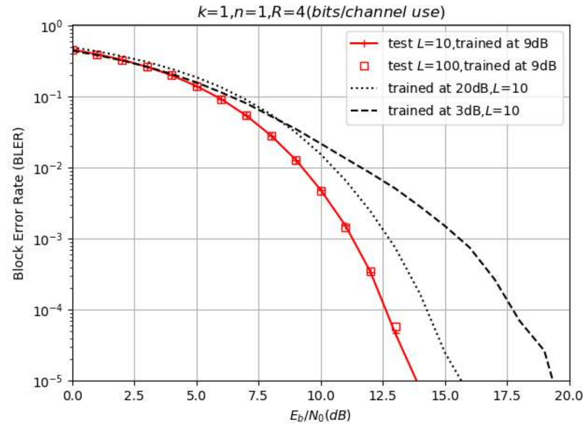

FIGURE 3. BLER performance of the CNN-AE system of Fig.2 under AWGN channels, when trained at E_b/N_0 =3,9,20dB with L =10, while testing at L =10,100.

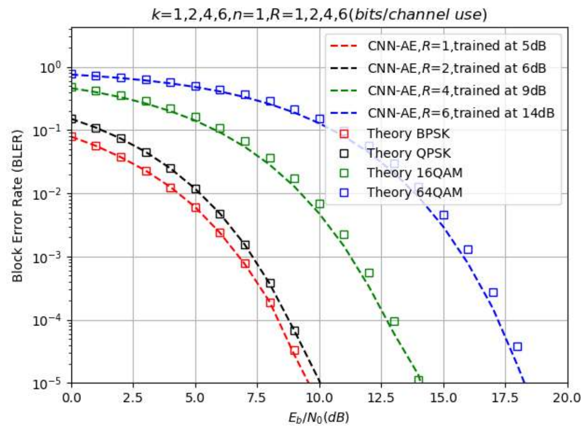

FIGURE 4. BLER performance of the CNN-AE system of Fig.2 having different rates R =1,2,4,6 (bits/channel use), when compared with corresponding BPSK, QPSK, 16QAM, 64QAM modulations under AWGN channels.

intricate signal structure under the influence of channel impairments.

B. GENERALIZATION WITH DIFFERENT CODE RATES UNDER AWGN AND RAYLEIGH CHANNELS

In this section, the proposed CNN-AE system is generalized to work with various rates under both AWGN and Rayleigh fading channels. In order to compare with classic modulation schemes i.e. BPSK, QAM, etc., which are used as benchmarks, CNN-AE's channel use was set to $n = 1$ and different code rates were achieved by changing the number of bits in a symbol (k).

Fig. 4 plots the CNN-AE's BLER performance having different rates R = 1,2,4,6 (*bits/ channel use*), when transmitting over AWGN channels. On one hand, the representation power of CNNs, which is facilitated by the 256 filters in Table 1 and deep network architecture, can learn suitable symbol transformations even for high data rates. On the other hand, the achieved BLER performance can match that of the corresponding conventional BPSK,

[1]Source codes available in: https://github.com/ZhangKaiyao/Deepcom/tree/master

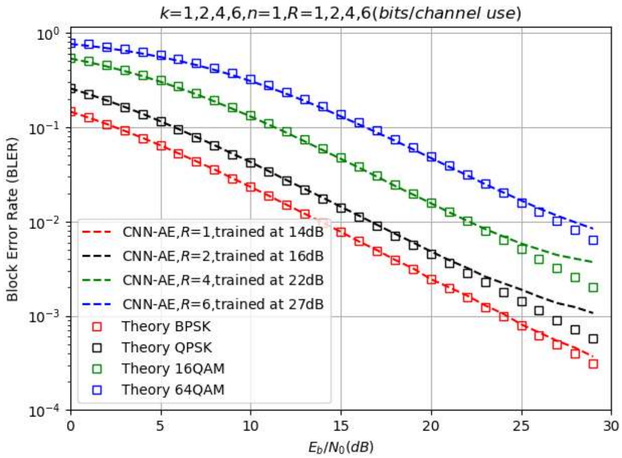

FIGURE 5. BLER performance of the CNN-AE system of Fig.2 having different rates R =1,2,4,6 (bits/channel use), when compared with corresponding BPSK, QPSK, 16QAM, 64QAM modulations under flat Rayleigh fading channels.

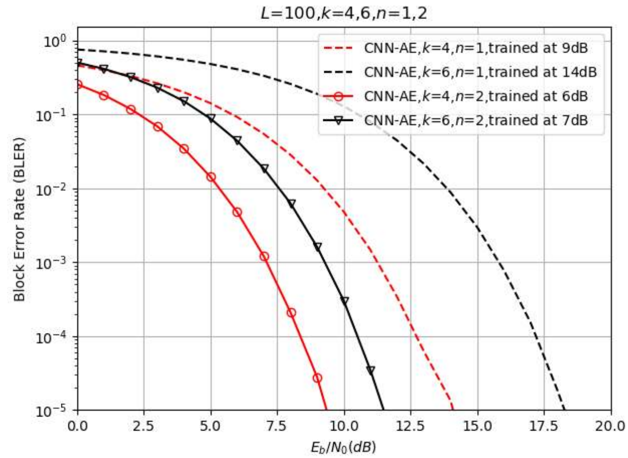

FIGURE 6. BLER performance of the CNN-AE system of Fig.2 under AWGN channels, when having different channel use n.

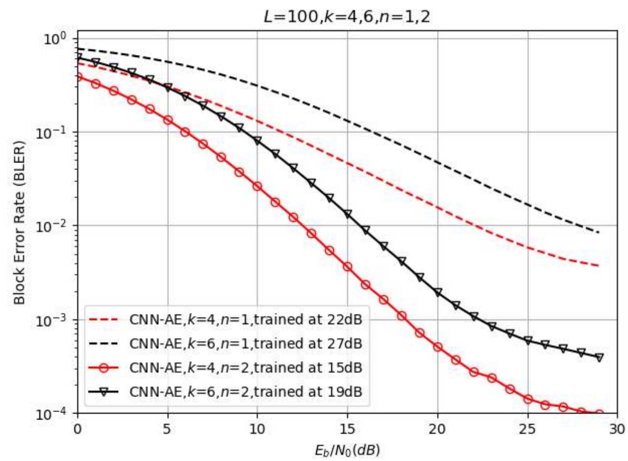

FIGURE 7. BLER performance of the CNN-AE system of Fig.2 under flat rayleigh fading channels, when having different channel use n.

QPSK, 16QAM and 64QAM schemes, which are the optimal solutions under AWGN channels.

Fig. 5 demonstrates the BLER performance of the CNN-AE system of Fig. 2 having different rates $R = 1,2,4,6$ (*bits/ channel use*), when transmitting over flat Rayleigh channels. Note that perfect CSI is assumed at the receiver, which is fed into the Conv1D layers together with the received signal $\mathbf{Y}$ as a concatenated vector. Observe in Fig.5 that the receiver learns to equalize the fading effects before decoding, and the resultant BLERs are identical to the BPSK, QPSK, 16QAM and 64QAM counterparts, as expected.

C. GENERALIZATION WITH CHANNEL USE

If multiple time slots (n>1) are available at the channel for each symbol transmission, the proposed CNN-AE system can exploit the additional time domain resources, owing to the representation power of the Conv1D layers at the transmitter.

From modulation perspective, the design of constellation points is carried out in 2n-*dimensional* space, which potentially could maximize the minimum Euclidean distance between constellation points, compared with the 2-*dimensional* I/Q space having $n = 1$. From channel coding point of view, the Conv1D layers learn appropriate block encoding across 2n-*dimensional* space that could increase the minimum hamming distance between codewords. Therefore, it is reasonable to expect that better BLER can be achieved with the increase of channel use n.

Fig. 6 and Fig. 7 compare the BLER performance of the CNN-AE system under both AWGN and flat Rayleigh fading channels, when having different channel use n. As expected, for a given k value, the increase of $n = 1$ to $n = 2$ results in substantial performance gain. Again, this coding gain was achievable, because the CNN architecture allowed us to search constellation points in 2n-*dimensional* space.

D. DIFFERENTIAL CNN-BASED AUTOENCODER SCHEME

The CNN-AE system of Fig. 2 requires the perfect CSI to assist the receiver to combat the effect of fading. The burden of channel estimation could be potentially quite significant, especially for massive MIMO systems. In order to eliminate the need of CSI at the receiver, we further propose a Differential CNN-based Autoencoder (DCNN-AE) system, as shown in Fig. 8.

Similar to the idea of conventional DPSK schemes, the signals are differentially encoded before subjected to power normalization, as seen in Fig. 8. At the receiver, the received signal $\mathbf{Y}$ is firstly differentially decoded before fed into the neural networks. The differential encoding/decoding processes are implemented as non-trainable layers in Fig. 8, which support complex number operations.

Fig. 9 illustrates the BLER performance of the DCNN-AE system of Fig. 8, when transmitting over Rayleigh block fading channels. The theoretical D-BPSK scheme's BLER is used as the bench-marker. When $k = 1$, the proposed

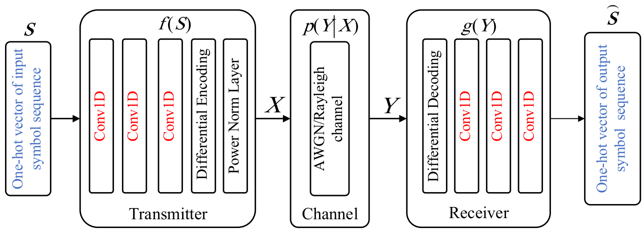

FIGURE 8. System diagram of Differential CNN-based Autoencoder (DCNN-AE) system, which eliminates the need of channel estimation.

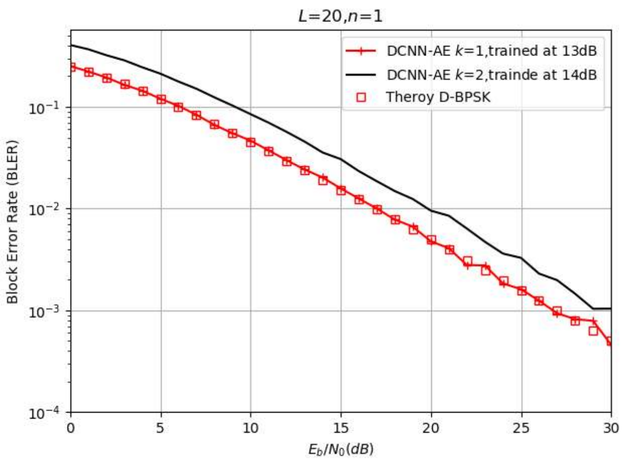

FIGURE 9. BLER performance of the differential CNN-AE (DCNN-AE) system of Fig.8 under rayleigh block fading channels.

autoencoder system's BLER performance can match to that of the D-BPSK system. When $k = 2$, a 3dB performance gap is recorded in Fig. 9, as expected.

IV. ADAPTIVITY AND TRAINING CONVERGENCE

A. ADAPTIVITY ON NON-GAUSSIAN CHANNEL

For AWGN and Rayleigh fading channels, where the optimal solutions are known, Fig. 4 and Fig. 5 have already demonstrated that the learned scheme can match to that of the conventional schemes. In this section, we demonstrate that even if the channel does not obey a clean mathematical analysis, the proposed scheme can still find the optimal solution via end-to-end learning.

More explicitly, consider a scenario, where the transmitted signal is always corrupted by AWGN, in addition, with a small probability, a further high variance noise is added. The channel model is described as follows: $y = x + z + w$, where $z \sim N(0, \sigma_1^2)$ is the background AWGN noise and $w \sim N(0, \sigma_2^2)$ denotes the bursty noise with probability ρ, and $\sigma_2^2 \gg \sigma_1^2$. This channel model accurately captures how bursty radar signals interfere with LTE signals [31].

Under the aforesaid channel model with σ_2 =1.0 and ρ =0.05, Fig. 10 plots the BLER performance of the $k = 4$ CNN-AE system of Fig. 2, when having different channel use n. Observe in Fig. 10 that the conventional 16QAM modulation record the worst BLER performance and an error

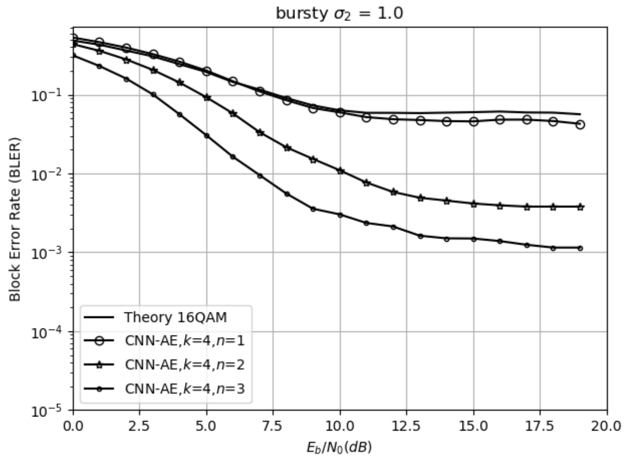

FIGURE 10. BLER performance of the *k=4* CNN-AE system of Fig.2 under bursty noise channel (with standard deviation σ_2 =1 and probability ρ =0.05), when having different channel use *n*.

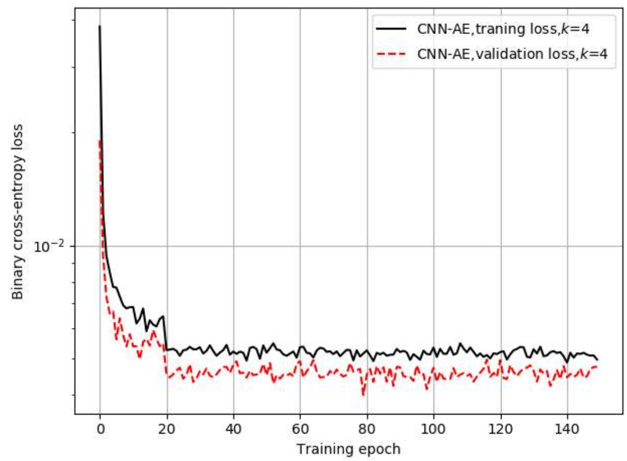

FIGURE 11. Training loss and validation loss of the CNN-AE system of Fig. 2 under AWGN channels with*k=4*, which was trained using 150 epochs.

floor existed because information symbols are permanently lost whenever bursty noise occurred. On the other hand, the CNN-AE system with n =1 achieves slightly lower error floor, owing to the end-to-end optimization of the system. Furthermore, significant performance gain can be achieved in Fig. 10, if the channel use is increased to n =2, 3. That is because the transmitted signal is designed in *2n-dimensional* space, whereas the bursty noise is limited to a single time slot when occurred. Hence, information symbol recovery remains possible under bursty channels.

B. TRAINING CONVERGENCE

The proposed end-to-end learning system also has the benefit of fast training convergence, which is exemplified in Fig. 11 and Fig. 12.

Fig. 11 and Fig. 12 show the training loss and validation loss of the CNN-AE system as a function of training epochs

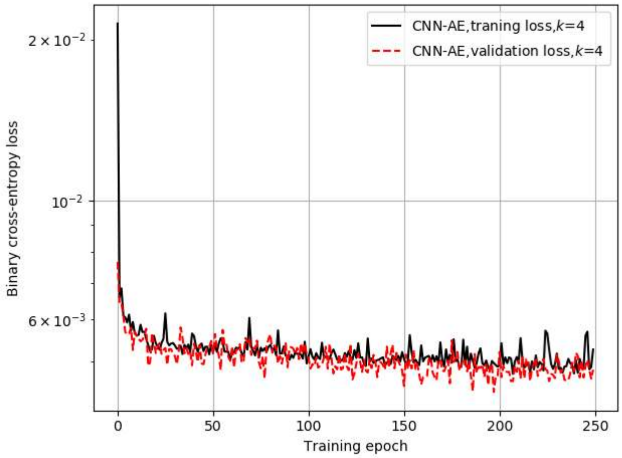

FIGURE 12. Training loss and validation loss of the CNN-AE system of Fig. 2 under Rayleigh fading channel with $k = 4$, which was trained using 250 epochs.

for AWGN and flat Rayleigh fading channels, respectively. Both loss functions converge rapidly within several epochs. Note that the proposed system was insensitive to the initialization parameters of the neural networks.

V. CONCLUSION

In this treatise, a CNN-based autoencoder communication system is proposed, which infuses communication domain expertise into neural networks. CNN layers are employed in order to learn the representations needed for transmission and detection through raw data. More explicitly, merely power constraints, time constraints (number of channel use) and the channel model are imposed on the neural layers. Then, supervised learning with sufficient data drives the CNN-based model to converge on the exact network parameters that are suitable for pre-defined communication requirements.

This generalized learning framework allow us to communicate intelligently under various conditions. Furthermore, under AWGN and flat Rayleigh fading channels, the proposed CNN-AE system can match the performance of existing optimal human engineered solutions, regardless block length, E_b/N_0, code rates, channel use. In other words, the CNN-AE scheme serves as a unified system for communications. When transmitting over non-standard bursty channels, the proposed network is able to out-perform existing schemes, since the learned representations of the signal adapts to the environment through constellation optimization in high dimensional space. Moreover, a differential CNN-AE scheme is proposed in order to eliminate the need of channel estimation. Finally, we demonstrate that the CNN-AE architecture can be trained with limited epochs, since it has fast training convergence.

REFERENCES

[1] E. Zehavi, "8-PSK trellis codes on Rayleigh channel," in *Proc. IEEE Mil. Commun. Conf., 'Bridging Gap. Interoperability, Survivability, Secur.'*, vol. 2, Oct. 1989, pp. 536–540.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[3] A. Felix, S. Cammerer, S. Dörner, J. Hoydis, and S. T. Brink, "OFDM-autoencoder for end-to-end learning of communications systems," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.

[4] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmWave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 852–855, Oct. 2018.

[5] D. Neumann, T. Wiese, and W. Utschick, "Learning the MMSE channel estimator," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2905–2917, Jun. 2018.

[6] H. Ye, L. Liang, G. Y. Li, and B.-H. F. Juang, "Deep learning based end-to-end wireless communication systems with conditional GAN as unknown channel," 2019, *arXiv:1903.02551*. [Online]. Available: https://arxiv.org/abs/1903.02551

[7] N. Samuel, T. Diskin, and A. Wiesel, "Deep MIMO detection," in *Proc. IEEE 18th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2017, pp. 1–5.

[8] 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, no. 1, pp. 114–117, Feb. 2018.

[9] H. Huang, S. Guo, G. Gui, Z. Yang, J. Zhang, H. Sari, and F. Adachi, "Deep learning for physical-layer 5G wireless techniques: Opportunities, challenges and solutions," 2019, *arXiv:1904.09673*. [Online]. Available: https://arxiv.org/abs/1904.09673

[10] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.

[11] E. Nachmani, E. Marciano, D. Burshtein, and Y. Be'ery, "RNN decoding of linear block codes," 2017, *arXiv:1702.07560*. [Online]. Available: https://arxiv.org/abs/1702.07560

[12] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440–8450, Sep. 2018.

[13] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.

[14] B. Zhu, J. Wang, L. He, and J. Song, "Joint transceiver optimization for wireless communication PHY using neural network," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1364–1373, Jun. 2019.

[15] Y. LeCun, "Generalization and network design strategies," in *Connectionism in Perspective*, vol. 19. Amsterdam, The Netherlands: Elsevier, 1989.

[16] T. Gruber, S. Cammerer, J. Hoydis, and S. T. Brink, "On deep learning-based channel decoding," in *Proc. 51st Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2017, pp. 1–6.

[17] H. Kim, Y. Jiang, R. Rana, S. Kannan, S. Oh, and P. Viswanath, "Communication algorithms via deep learning," 2018, *arXiv:1805.09317*. [Online]. Available: https://arxiv.org/abs/1805.09317

[18] Y. Jiang, H. Kim, H. Asnani, S. Kannan, S. Oh, and P. Viswanath, "DeepTurbo: Deep turbo decoder," 2019, *arXiv:1903.02295*. Accessed: Mar. 1, 2019. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2019arXiv190302295J

[19] T. J. O'Shea, K. Karra, and T. C. Clancy, "Learning to communicate: Channel auto-encoders, domain specific regularizers, and attention," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2016, pp. 223–228.

[20] T. J. O'Shea, L. Pemula, D. Batra, and T. C. Clancy, "Radio transformer networks: Attention models for learning to synchronize in wireless systems," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 662–666.

[21] T. J. O'Shea, T. Erpek, and T. C. Clancy, "Deep learning based MIMO communications," 2017, *arXiv:1707.07980*. [Online]. Available: https://arxiv.org/abs/1707.07980

[22] X. Chen, J. Cheng, Z. Zhang, L. Wu, and J. Dang, "Data-rate driven transmission strategy for deep learning based communication systems," 2018, *arXiv:1812.08869*. [Online]. Available: https://arxiv.org/abs/1812.08869

[23] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019.

[24] S. Dörner, S. Cammerer, J. Hoydis, and S. ten Brink, "Deep learning based communication over the air," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 132–143, Feb. 2018.

[25] F. A. Aoudia and J. Hoydis, "End-to-end learning of communications systems without a channel model," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 298–303.
[26] T. J. O'Shea, T. Roy, N. West, and B. C. Hilburn, "Physical layer communications system design over-the-air using adversarial networks," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 529–532.
[27] H. Ye, G. Y. Li, B.-H. F. Juang, and K. Sivanesan, "Channel agnostic end-to-end learning based communication systems with conditional GAN," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–5.
[28] T. J. O'Shea, T. Roy, and N. West, "Approximating the void: Learning stochastic channel models from observation with variational generative adversarial networks," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2019, pp. 681–686.
[29] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
[30] J. H. Conway and N. Sloane, *Sphere Packings, Lattices and Groups*. Springer, 1988.
[31] F. H. Sanders, J. E. Carroll, G. A. Sanders, and R. L. Sole, "Effects of radar interference on LTE base station receiver performance," US Dept. Commerce, NTIA, Washington, DC, USA, 2013.

NAN WU (M'05) received the B.S. degree in electrical and electronics engineering from the Dalian University of Technology, in 2003, and the M.Sc. and Ph.D. degrees from the University of Southampton, U.K., in 2004 and 2008, respectively.

From 2008 to 2009, he was a Guest Researcher with the National Institute of Standards and Technology, USA. Since 2010, he has been an Associate Professor with the School of Electrical and Information Engineering, Dalian Maritime University, China. He has authored one book and more than 50 articles. His research interests include deep learning-based communications, visible light communication, indoor positioning, and MIMO. He also holds two patents.

XUDONG WANG (M'08) received the B.S. degree in electronics engineering, the M.Sc. degree in circuits and systems, and the Ph.D. degree in electronics and science technology from Xidian University, in 1989, 1992, and 2008, respectively.

He is currently a Professor with the School of Information and Science Technology, Dalian Maritime University. He has authored one book, more than 70 articles, and more than 19 inventions. His current work is mainly focused on wireless communications, with a special emphasis on fading channel modeling and MIMO system performance analysis.

BIN LIN (M'11) received the B.S. and M.S. degrees from Dalian Maritime University, Dalian, China, in 1999 and 2003, respectively, and the Ph.D. degree from the Broadband Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, in 2009.

She was a Visiting Scholar with George Washington University, Washington, DC, USA, from 2015 to 2016. She is currently a Full Professor with the Department of Information Science and Technology, Dalian Maritime University. Her current research interests include wireless communications, network dimensioning and optimization, resource allocation, artificial intelligence, maritime communication networks, edge/cloud computing, wireless sensor networks, and the Internet of Things. She is an Associate Editor of the *IET Communications*.

KAIYAO ZHANG received the B.S. degree from the Department of Information, Shenyang Jianzhu University, Shenyang, China, in 2017. He is currently pursuing the master's degree with the Department of Information Science and Technology, Dalian Maritime University. He has published a paper on the 7th International Conference on Communications, Signal Processing, and Systems, Dalian, China. His research interest includes deep-learning-based communications.