A Novel OFDM Autoencoder Featuring CNN-Based Channel Estimation for Internet of Vessels

Bin Lin, *Member, IEEE*, Xudong Wang, *Member, IEEE*, Weihao Yuan, and Nan Wu, *Member, IEEE*

***Abstract*—This article proposes a novel orthogonal frequency-division multiplexing (OFDM) autoencoder featuring convolutional neural networks (CNNs)-based channel estimation for marine communications with complex and fast-changing environments. We demonstrate that the proposed OFDM autoencoder system can be generalized to work under various channel environments, different throughputs, while outperforming the traditional OFDM counterparts, especially when working at high throughputs. In addition, since OFDM systems require accurate channel estimations to function properly, this treatise also proposes a new channel estimation algorithm for OFDM systems that combine the power of deep learning (DL) with the philosophy of super-resolution reconstruction, which uses dense convolutional neural networks (Dense-Nets) to reconstruct low-resolution pilot information images into high-resolution full-channel impulse responses (CIRs). The Dense-Net structure has the characteristics of dense connections and feature multiplexing. The simulation results show that under slow fading, the proposed channel estimator (CE) can estimate the CIRs perfectly. Under fast fading, the proposed CE outperforms the existing learning-based algorithms with fewer neural network parameters. Therefore, the proposed novel autoencoder scheme and the powerful CE are potentially attractive approaches for the Internet of Vessels (IoV).**

***Index Terms*—Autoencoder, channel estimation, convolutional neural networks (CNNs), Internet of Vessels (IoV), orthogonal frequency-division multiplexing (OFDM).**

I. Introduction

EMERGING Internet of Vessels (IoV) [1] is expected to play an important role in the realm of intelligent marine transportation systems. Specifically, the maritime autonomous surface ship (MASS) is widely recognized as the development trend of the future shipping industry. IoV invokes a large amount of information exchange among vessels and land-shore facilities, such as position, speed, and route. As a result, IoV has the ability to realize the refinement of shipping management, the comprehensiveness of industry services, and the humanization of user experiences by providing an intelligent and safer navigating environment [1], [2]. Thus, high throughputs and reliability are essential for IoV communications and networking, which can employ orthogonal frequency-division multiplexing (OFDM) techniques [3], [4] as the information-bearing signaling. Furthermore, marine wireless communication systems are required to accommodate complex and fast-changing channel environments, as compared with their terrestrial counterparts.

Meanwhile, the effort of introducing deep learning (DL) into the field of wireless communications [5], [6] grows significantly with the breakthrough of DL in image processing, speech recognition, natural language processing, etc. The resultant DL-based wireless transceivers have already demonstrated promising system performance and provided inspirations in many aspects of communication system designs. In particular, the pioneering work of O'Shea and Hoydis [7] and O'Shea *et al.* [8] proposed a communication system based on the concept of the autoencoder optimized through end-to-end learning. In contrast to conventional communications systems, this single-carrier autoencoder architecture employs neural networks to jointly optimize the transmitter and receiver, conditioned on the surrounding channel environments. Since then, researchers around the globe have relentlessly pushed the boundary of DL into higher OSI layers, such as hatching the idea of DL in resource allocation [9], [10], routing [11], and multiple access protocols [12].

As for the multicarrier modulations, OFDM technologies are often used to combat the effect of frequency selective fading, since it can transform a frequency-selective fading channel into parallel flat-fading subchannels [13]. Naturally, the challenge of fusing DL with OFDM motivates researchers to design innovative frameworks for IoV, where an increasing number of vessels are required to be connected. Among them, Balevi and Andrews [14], Felix *et al.* [15], and Kim *et al.* [16] used the fully connected (FC) neural network layers to implement an OFDM system based on autoencoder structures. However, FC layers require a large amount of network parameters, and the vanish of gradients during training becomes problematic when building deep neural networks.

On the other hand, OFDM systems demand high-accuracy channel estimations in order to recover information correctly, where pilot-based channel estimation algorithms, such as least squares (LSs) and linear minimum mean square error (LMMSE) algorithms are often used [17]. More explicitly, Zhao *et al.* [18] implemented a channel estimation module based on the LS algorithm using deep complex convolutional networks, whereas Ye *et al.* [19] proposed an OFDM channel estimation module using FC neural networks. Gao *et al.* [20] proposed a model approach that combined DL with the expert knowledge to replace

Manuscript received January 30, 2020; revised March 14, 2020; accepted April 04, 2020. Date of publication April 8, 2020; date of current version August 12, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61971083 and Grant 51939001, and in part by the Dalian Science and Technology Innovation Fund under Grant 2019J11CY015. *(Corresponding author: Nan Wu.)*

The authors are with the Department of Information Science and Technology, Dalian Maritime University, Dalian 116026, China (e-mail: binlin@dlmu.edu.cn; wu.nan@dlmu.edu.cn).

Digital Object Identifier 10.1109/JIOT.2020.2986442

the existing channel estimation module. In addition, the channel estimation problem can also be formulated as the issue of super-resolution reconstruction in the field of image processing [21]. The pilot data across the time and frequency domains are treated as a low-resolution image and the full-channel responses that need to be estimated are considered to be the corresponding high-resolution image [22], where vanilla convolutional neural networks (CNNs) were used. However, this shallow CNN's structure prevented the channel estimation algorithm to achieve higher accuracy.

In this article, a novel autoencoder-based OFDM system using CNNs layers to facilitate end-to-end learning designed for IoV is proposed. Compared with FC layers, our CNN-based systems can learn intricate signal representations with fast convergence during training, while using fewer network parameters. Moreover, long short-term memory (LSTM) network layers [23] are also used to exploit the dependency of time sequence, so that the channel coding gain becomes achievable. Furthermore, in order to handle complex marine communication environments, we propose a novel channel estimation algorithm for OFDM systems combining the power of DL with super-resolution reconstruction, where the low-resolution pilot information image is reconstructed into a full-channel response high-resolution image using dense convolutional neural networks (Dense-Nets) [24]. The Dense-Net structure allows us to effectively solve the vanishing gradient problem caused by the excessive network layers of typical CNN networks. In addition, Dense-Net has fewer network parameters than the residual network (Res-Net) [25] architecture. We will demonstrate that the proposed algorithm's channel estimation performance is better than the state-of-the-art DL-based algorithm [18]. Apart from the performance gain, the proposed algorithm has a deeper network structure to learn high-level features with fewer total parameters than that of [18]. More explicitly, the contributions of this article are as follows.

1) The carefully designed CNN layers allow the learned OFDM autoencoder to have generalization capability while achieving optimal block error rate (BLER) performance, namely, support flexible data rates, suitable for AWGN and Rayleigh, as well as non-AWGN channels.
2) Furthermore, LSTM layers are introduced on top of CNN layers to achieve the channel coding gain, which is essential to combat hostile marine communication environments. Namely, the entire transmitted sequence is encoded using LSTM layers before transmission, while the LSTM layers at the receiver are responsible for extracting adequate correlated information for detection.
3) To the best of our knowledge, this is the first study in the literature to propose a novel Dense-Net channel estimation scheme based on the idea of image super resolution, which not only achieves the state-of-the-art estimation accuracy but also solves conventional CNN's problem of gradient vanishing and excessive parameters through dense connections and feature multiplexing.

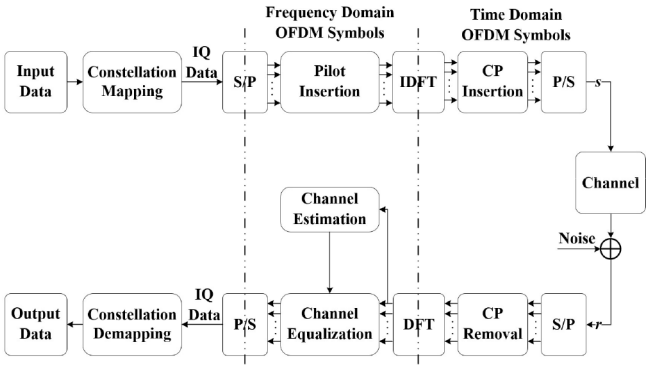


Fig. 1. Block diagram of a conventional OFDM communication system.

4) The proposed OFDM autoencoder and the novel Dense-Net-based channel estimator (CE) can be trained jointly with fast convergence, hence becomes an appealing solution to dynamic IoV communications.

The remainder of this article is organized as follows. The introduction of a conventional OFDM system is provided in Section II. Section III introduces the proposed OFDM autoencoder and the CNN-based CE. Section IV provides numerous simulation results as well as numerical analysis, where the summary is made in Section V.

II. Typical OFDM Communication Systems

In this section, a conventional OFDM communication system is briefly described, which is used as the benchmarker for the proposed DL-based OFDM autoencoder.

The block diagram of a conventional OFDM communication system is shown in Fig. 1. First, the input bits are mapped to QAM constellation points, which consist of in-phase and quadrature components (IQ) in the constellation plane. Then, the IQ signals are mapped to N (equals the number of subcarriers used to transmit data) sets of parallel data streams via serial to parallel (S/P) conversion. Later, the pilot data are inserted in each frequency-domain OFDM symbol. The insertion pattern of the pilot signals is chosen to be comb mode in order to combat the effect of fast fading [26], which is shown in Fig. 2. Hence, the resultant frequency-domain OFDM symbols are converted to their time-domain representations using inverse discrete Fourier transform (IDFT) operations, followed by adding cyclic prefix (CP) to deal with intersymbol interference (ISI). The CP's length depends on the size of the multipath delay.

The multipath fading channel is modeled as

$$r = s \otimes h + n \tag{1}$$

where the vectors s and r represent the transmitted and received signals in the time domain, respectively, and n represents the white noise. h is the channel impulse responses (CIRs) of the multipath channel and $\otimes$ denotes the convolution operation. Assuming that the channel response remains static within one OFDM symbol while changing between difference OFDM symbols, the tapped delay line

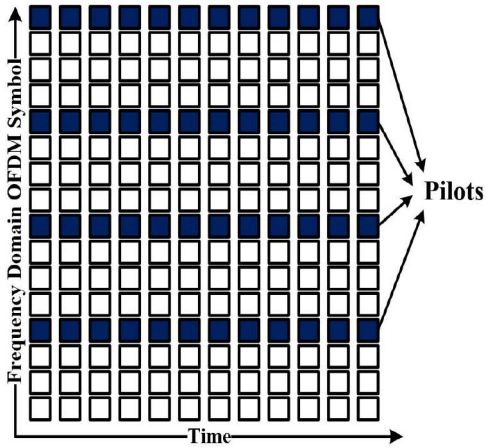


Fig. 2. Insertion of pilot data in a frequency-domain OFDM symbol.

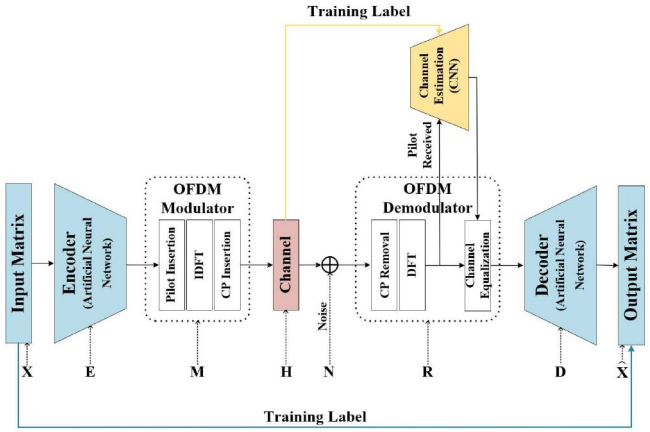


Fig. 3. Block diagram of the proposed OFDM autoencoder with a novel CNN-based CE.

model [27] of h is given by

$$h(\tau, t) = \sum_{k=1}^{K} a_k(t)\delta(\tau - \tau_k) \tag{2}$$

where $a_k(t)$ is the gain of the signal on the kth path, τ_k is the delay of the kth path, and K is the total number of paths.

At the receiver, after removing the CPs, the received time-domain OFDM symbols r of (1) will be converted to its frequency domain representations via discrete Fourier transform (DFT) operations. Then, the channel estimation module of Fig. 1 estimates the full CIRs using the received pilot data, which assist the channel equalization module to equalize the received frequency-domain symbols. Finally, the output signal is demodulated into binary data streams after the P/S operation.

III. Proposed OFDM Autoencoder

In this section, the design and training methods of the proposed OFDM autoencoder for IoV is introduced, where the system diagram is illustrated in Fig. 3. The purpose of the end-to-end structure is to find appropriate signal representations for the transmitter that can adapt to the surrounding marine channels, such that the receiver can recover the information with minimum error probability. More explicitly, two autoencoder structures are proposed. One uses CNN layers to facilitate the task of modulation, the other is designed using both the CNN and LSTM layers to perform "joint" channel coding and modulation.

Besides, the problem of channel estimation is tackled using the idea from imagine super resolution [28], [29], where pilot data are the low-resolution signals and neural networks are used to reconstruct full CIRs.

A. CNN-Based OFDM Autoencoder

The proposed OFDM autoencoder based on CNN is shown in Fig. 3. First, a group of $N \times \log_2 m$ information bits are converted into one-hot matrix format $\boldsymbol{X} \in \mathbb{C}^{N\times m}$, which corresponds to N number of conventional m-array QAM symbols in conventional OFDM systems. The one-hot data format can assist the transmitter to converge to the optimal solution quickly. Then, the CNN-based autoencoder $\boldsymbol{E} \in \mathbb{C}^{m\times 2}$ transforms the input matrix $\boldsymbol{X}$ into $\boldsymbol{XE} \in \mathbb{C}^{N\times 2}$, where the real and imaginary parts of the signal are represented as two real numbers. Furthermore, the information signal $\boldsymbol{XE}$ is protected by adding P number of pilot signals and C number of CPs, before modulated to multiply carriers using IDFT operations, as seen in Fig. 3. This process is characterized using a matrix operation denoted as $\boldsymbol{M} \in \mathbb{C}^{S\times N}$, where $S = N+P+C$. Hence, the learned signal $\boldsymbol{MXE} \in \mathbb{C}^{S\times 2}$ propagates through the channel $\boldsymbol{H} \in \mathbb{C}^{S\times S}$ and $\boldsymbol{MXE}$ is corrupted by AWGN $\boldsymbol{N} \in \mathbb{C}^{S\times 2}$ before arriving at the receiver.

Next, the OFDM demodulator at the receiver of Fig. 3 performs an inverse matrix operation denoted as $\boldsymbol{R} \in \mathbb{C}^{N\times S}$ to the received signal, where CP removal, DFT calculations, and channel equalizations are processed. More importantly, a novel CNN-based CE of Fig. 3 is used in order to accurately estimate the CIRs based on the pilot signals. Then, the CNN-based decoder transforms the received baseband signal into suitable representations using $\boldsymbol{D} \in \mathbb{C}^{2\times m}$ of Fig. 3, so that the recovery of data $\widehat{\boldsymbol{X}} \in \mathbb{C}^{N\times m}$ using supervised learning becomes feasible.

Finally, the whole OFDM autoencoder of Fig. 3 can be expressed using matrix operations as

$$\widehat{\boldsymbol{X}} = (\boldsymbol{R}(\boldsymbol{H}(\boldsymbol{MXE}) + \boldsymbol{N}))\boldsymbol{D}. \tag{3}$$

Unlike the traditional OFDM communication systems of Fig. 1 that are designed block-by-block, the proposed end-to-end learning architecture can globally find the optimal solution for the OFDM signaling representation and detection in the overparametrized neural network space, while subject to challenging channel conditions.

Compared to FC neural networks, the employed convolutional layers of Fig. 3 can significantly reduce the number of trained parameters through weights sharing. From the communication perspective, this means the information is processed locally or in a block manner. In addition, the convolutional layers are capable of mapping bitstreams to higher dimensions by increasing the number of kernels. The detailed parameters of CNN layers used in Fig. 3 are shown in Table I. More explicitly, 1-D convolutional (Conv-1-D) layers are used for both encoding and decoding, where B in the first dimension

TABLE I
PROPOSED CNN-BASED AUTOENCODER

Layer Name	Type	Output Shape
Input	One-hot vector	(B, N, m)
Encoder1	Conv1D, kernel size: 1	(B, N, 256)
Encoder2	Conv1D, kernel size: 1	(B, N, 256)
Encoder3	Conv1D, kernel size: 1	(B, N, 2)
Modulator	Non-trainable	(B, N+P+C, 2)
Channel	Non-trainable	(B, N+P+C, 2)
Demodulator	Non-trainable	(B, N+P, 2)
Equalization	Non-trainable	(B, N, 2)
Decoder1	Conv1D, kernel size: 1	(B, N, 256)
Decoder2	Conv1D, kernel size: 1	(B, N, 256)
Output	Conv1D, activation: softmax	(B, N, m)

TABLE II
PROPOSED CODED CNN-BASED AUTOENCODER USING LSTM

Layer Name	Type	Output Shape
Input	One-hot vector	(B, N*e, m)
Encoder1	Conv-1D, kernel size: 1	(B, N*e, 32)
Encoder2	Conv-1D, kernel size: 3	(B, N*e, 32)
Encoder3	Conv-1D, kernel size: 1	(B, N*e, 32)
Encoder4	LSTM, units: 128	(B, N*e, 128)
Encoder5	Time Distributed	(B, N*e, 2/e)
Reshape I	Reshape	(B, N, 2)
Modulator	Non-trainable	(B, N+P+C, 2)
Channel	Non-trainable	(B, N+P+C, 2)
Demodulator	Non-trainable	(B, N+P, 2)
Equalization	Non-trainable	(B, N, 2)
Reshape II	Reshape	(B, N*e, 2/e)
Decoder1	LSTM, units: 64	(B, N*e, 64)
Decoder2	LSTM, units: 32	(B, N*e, 32)
Decoder3	Conv-1D, kernel size: 1	(B, N*e, 32)
Decoder4	Conv-1D, kernel size: 3	(B, N*e, 32)
Output	Time Distributed	(B, N*e, m)

of the output shape is the mini-batch size, the second output dimension N in Table I is the number of OFDM symbols, and the last dimension is the number of kernels.

Note that since the kernel size of all the Conv-1-D layers of Table I is set to 1, hence, each information vector is processed individually rather than jointly, which means that there is no correlation between the information vectors. In the language of traditional communications, the proposed autoencoder of Table I replaces the traditional QAM modulation/demodulation using Conv-1-D layers, which enable the autoencoder system to search for the optimal modulation solution in an overparameterized space for different channel conditions.

B. Coded CNN-Based OFDM Autoencoder Using LSTM

Furthermore, it is well known that channel coding can improve the system's performance by correlating the information bits to the adjacent bits using convolutional codes (CCs), turbo codes, etc. In this section, we propose an OFDM autoencoder empowered by LSTM layers, so that the architecture has the capacity of achieving channel coding gain, where the specific neural network structure is given in Table II and illustrated in Fig. 4.

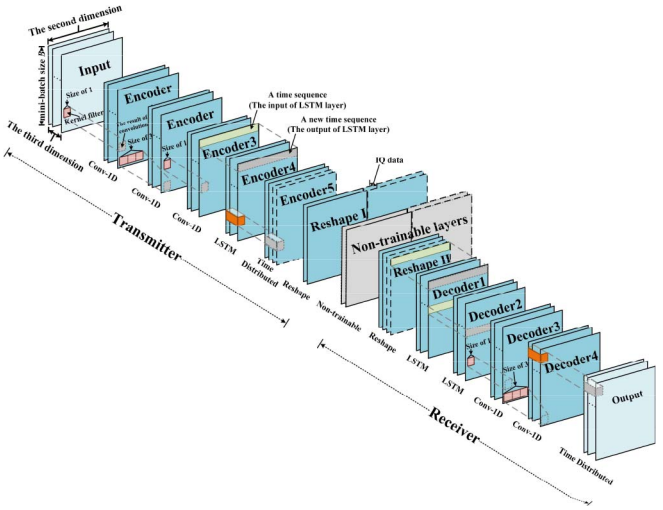


Fig. 4. Illustration of the proposed coded CNN-based autoencoder using LSTM, as listed in Table II.

At the transmitter side, three Conv-1-D layers are employed to facilitate signal transformation, however, the second Conv-1-D layer's kernel size is set to 3 to correlate the information symbols, as seen in Table II. Later, an LSTM layer having 128 units is added after the Conv-1-D layers in order to further spread the information vector to the adjacent transmit sequence, so that having CC-like channel coding becomes possible. More specifically, the output data sequence of the last Conv-1-D layer is fed into the LSTM cell to recurrently encode the current information to all the past time steps through LSTM's input, output, and forget gates. In addition, 128 units are used to ensure that adequate features are encoded into the whole sequence. In other words, it is the CNN layers together with the additional LSTM layer jointly carry out the task of modulation and channel coding for the autoencoder system. In addition, the time-distributed and reshape layers of Table II are used to conform the transmit signals to the appropriate OFDM frame structure.

At the receiver, compared with the pure CNN-based autoencoder of Table I, two LSTM layers are added to extract the correlated information from the received sequence, followed by two Conv-1-D layers, so that the channel decoding and demodulation are carried out jointly. Finally, the time-distributed layer is used to reconstruct the extracted features at each time step into the one-hot vector to complete the task of data recovery.

Unlike the traditional communication system, where channel coding and modulation are optimized separately, the proposed coded CNN-based OFDM autoencoder using LSTM of Table II uses a stack of Conv-1-D and LSTM layers to encode and correlate the data jointly. Furthermore, Conv-1-D layers have limited correlation capability of time sequence, since the kernel size is small. On the other hand, increasing the kernel size makes the neural network difficult to converge. Therefore, the additional LSTM layers are necessary to complement the Conv-1-D layers in terms of correlating long-time sequence signals, given that the LSTM layer has a broader data receptive and processing field than the Conv-1-D layers.

TABLE III
PROPOSED DENSE-NET CHANNEL ESTIMATION IN FIG. 3

Layer name	Type	Output Shape
Input	Received Pilot signals	(B, P, 2)
Reshape1	Reshape	(B, 2P, 1)
Up-Sampling	FC, activation: linear	(B, 2(N+P), 1)
DN1	Conv-2D, 1×1, activation: elu	(B, 2(N+P), 8)
Concat1	Concat [DN1, Up-Sampling]	(B, 2(N+P), 9)
DN2	Conv-2D, 3×3, activation: elu	(B, 2(N+P), 8)
Concat2	Concat [DN2, Concat1]	(B, 2(N+P), 17)
DN3	Conv-2D, 1×1, activation: elu	(B, 2(N+P), 8)
⋮	⋮	⋮
DN11	Conv-2D, 1×1, activation: elu	(B, 2(N+P), 8)
Concat11	Concat [DN11, Concat10]	(B, 2(N+P), 89)
DN12	Conv-2D, 3×3, activation: linear	(B, 2(N+P), 1)
Reshape2	Reshape	(B, N+P, 2)
Denoise1	Lambda (IDFT)	(B, N+P, 2)
Denoise2	Lambda (Zero-Setting)	(B, N+P, 2)
Output	Lambda (DFT)	(B, N+P, 2)

Besides, the autoencoder system's code rate e can be controlled by changing the number of input bits, that is, each group of $(N * e) \times \log_2 m$ information bits is converted into a one-hot matrix as the input.

C. CNN-Based Channel Estimation

We also propose a novel CNN-based CE for dynamic marine channel environments as shown in Fig. 3, where the specific network structure is given in Table III and illustrated in Fig. 5. The received pilot signals are served as the input of the CE. Then, the data are reshaped through the reshape layer of Table III, based on which the following 2-D convolutional (Conv-2-D) layers can infer full CIRs. The pilot data are further upsampled using an FC layer, therefore the pilot signals' dimension is expanded to $2(N+P)$. Furthermore, 12 consecutive Conv-2-D layers together with the shortcuts of the learned features of the previous layers are used, namely, the concatenation layers in Table III. This Dense-Net-type structure allows the proposed CE to accurately reconstruct the full-channel frequency responses from the upsampled pilot data.

Furthermore, the effect of the pilot signal's AWGN is minimized by three specially designed denoising layers in Table III. More specifically, Denoise1 layer of Table III converts the estimated frequency channel response to the time-domain CIRs through IDFT. Then, the last $(N+P)-K$ samples are set to zero since the full CIR's length is restricted to K. Finally, the time-domain CIRs are converted back to the frequency domain through DFT, which will be used in the channel equalizer.

The design philosophy of the proposed CNN-based CE originates from image super resolution in image processing. That is, treating the pilot data having a size of $B \times 2P$ as a low-resolution 2-D image, and the full-channel frequency response as the corresponding high-resolution image having a size of $B \times 2(N+P)$. Then, we innovatively design a Dense-Net network structure to facilitate the super-resolution process, where the input of each Conv-2-D layer is the concatenation of all the previous layers, which means the features learned by each layer are directly passed to all the latter layers. In contrast, the shallow CNN networks are unable to deliver the precision of the estimation required by the CE module, even though they are adequate for the autoencoder itself. In other words, the channel perturbation is more complex to learn than the encoding/decoding of the information sequence, hence a deeper and powerful Dense-Net architecture is necessary. Additionally, the proposed Dense-Net CE has the ability of fast convergence, owing to inherited shortcut connections. This structure allows the CE to piece together all the learned features to solve the "puzzle." Also, through feature multiplexing, the parameters required by the Dense-Net are greatly reduced. For example, only eight filters using the kernel of size 1×1 or 3×3 are used in each Conv-2-D layer in the DN layers of Table III. In addition, the dense connections alleviate the vanishing gradient phenomenon usually seen during the training of deep CNN networks.

D. Model Training

For the proposed CNN-based OFDM autoencoder system and the coded CNN-based OFDM autoencoder system using LSTM of Fig. 3, random one-hot data are generated as training data. The cross entropy of (4) is used as the loss function, which is given by

$$L_{\text{Cross-Entropy}} = -\sum_{i=1}^{U}\sum_{j=1}^{g} y_{ij}\log x_{ij} \tag{4}$$

where x and y are the training and label data sets, g is the number of categories, and U is the number of samples.

The training data of the CNN-based channel estimation of Fig. 3 is obtained after the pilot signals propagate through multipath fading channels, and the perfect CIRs are used as the label data. Besides, the mean square error (MSE) is used as the loss function

$$L_{\text{MSE}} = \frac{1}{U}\sum_{i=1}^{U}(y_i - x_i)^2. \tag{5}$$

The *elu* and linear activation functions are used in Tables I–III to perform nonlinear and linear transformations, respectively, and the SoftMax activation function is used at the decoder. We found that *elu* activation is more suitable to solve the problem of vanishing gradient and cell "dying" for the proposed autoencoder and CE networks, compared with *tanh*, *relu*, and other commonly used activation functions. We also find that using the early stopping and dropout mechanism with a rate of 0.05 are useful to prevent overfitting. ADAM optimizer [30] is used to optimize the backpropagation process and to control the step and direction of the gradient during training. In order to speed up the training process and facilitate the neural network convergence, the initial learning rate is set to be 0.01 and the learning rate will be reduced by a factor of ten, whenever the loss function of (4) or (5) is saturated for five consecutive epochs. In addition, the mini-batch gradient descent algorithm is used to speed up the training process, which is set to $B = 64$.

Besides, the training E_b/N_0 is also an important factor to converge to optimal solutions. When training at low E_b/N_0, neural networks have difficulty to learn intricate the data structure because of high noise power. On the other hand, if the

TABLE IV
OFDM PARAMETERS OF THE PROPOSED OFDM AUTOENCODER

Carrier Frequency	157.2MHz (International VHF Marine Radio Channel 24)
DFT Size	64
Pilots per OFDM Symbol	8
Number of used Subcarriers	56
CP Duration	0.1ms
Data Symbol Duration	0.4ms
Total Symbol Duration	0.5ms
Pilot Pattern	Comb of Fig. 2
Pilot Value	$1+1i$
Channel Model	AWGN, Multipath Fading
Channel Coherence	taps updated per OFDM symbol
Doppler Frequency	Flat or 8.73Hz

training E_b/N_0 is too high, neural networks lose the generalization ability to cope with noise-contaminated data. In this article, the CNN-based OFDM system of Table I was trained in the range of 4–10 dB, the coded CNN-based OFDM autoencoder system using LSTM of Table II was trained in the range of 10–20 dB which is higher than the uncoded system, while the CNN-based CE of Table III was trained at 20 dB.

IV. SIMULATION RESULTS

After the detailed OFDM parameters are given in Table IV, the BLER performance of the proposed CNN-based OFDM autoencoder under AWGN and multipath fading channels are provided, when the receiver has perfect CIRs. The CNN-based OFDM autoencoder of Fig. 3 is benchmarked by the conventional OFDM systems of Fig. 1 having corresponding m-array QAM modulations. The performance of the coded autoencoder using LSTM is compared with the traditional half-rate convolutional coded OFDM system. Furthermore, the OFDM autoencoder's BLER performance using the proposed CE of Table III is also provided under slow or fast fading channels. Our source codes are implemented in Keras and will be available on GitHub upon publication.

For the marine communication channel, since the relatively low-onboard antenna's height on a vessel and local scatters around the user introduce multiple paths in wireless channels [31], multipath Rayleigh fading models are used with flat fading or having a Doppler frequency of 8.73 Hz for fast fading [32], which corresponds to the maximum relative speed between ships of 60 km/h. We will show that, first, the OFDM autoencoders of Tables I and II can adapt to marine channel environments through learning, second, the proposed CE scheme is capable of outperforming the existing DL-based channel estimations [18] for marine channels.

A. AWGN and Fading Channels

Fig. 6 demonstrates the BLER performance of the CNN-based OFDM autoencoder of Fig. 3 under AWGN channels. In the case of $m = 2$ and 4, the proposed CNN-based OFDM system achieves nearly identical BLER performance

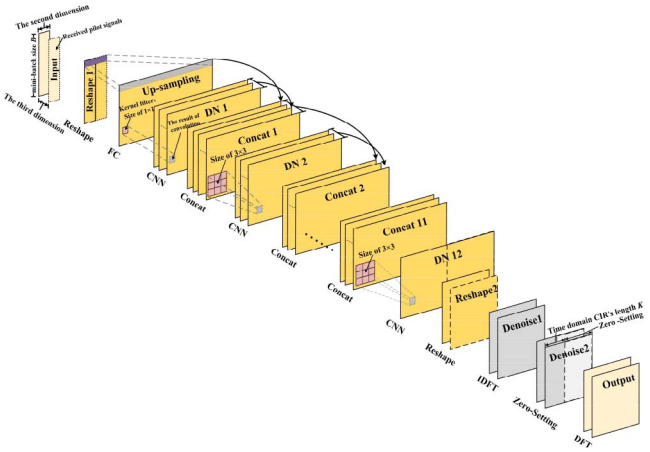


Fig. 5. Illustration of the proposed Dense-Net channel estimation network, as listed in Table III.

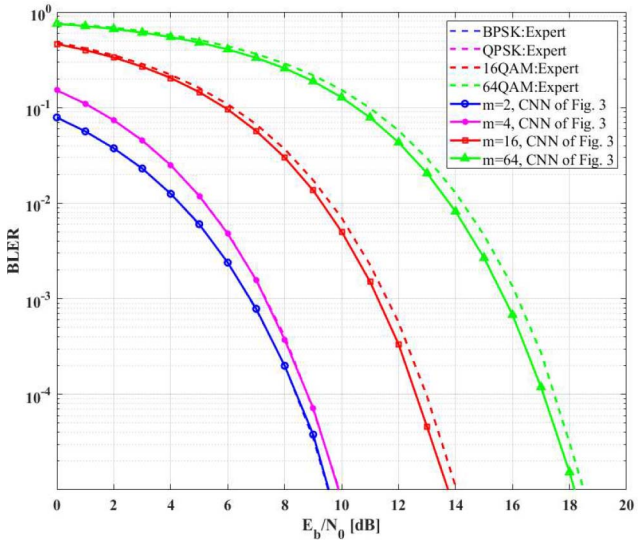


Fig. 6. BLER performance of the proposed CNN-based OFDM autoencoder under AWGN channels, when compared with corresponding expert OFDM systems.

compared with conventional expert OFDM systems of Fig. 1 having the same throughput, when using BPSK and QPSK. However, when the system throughput increases as $m = 16$ and 64, the CNN-based OFDM achieves slightly better BLER performance than that of the expert OFDM systems using 16-QAM and 64-QAM. That is because the CNN structure can optimize the data symbols jointly in higher dimensions. In addition, the CNN operations inheritably enable hamming-type block coding to the information bits [7], which becomes more transparent when the increase of m.

Fig. 7 shows the BLER performance of the CNN-based OFDM autoencoder of Fig. 3 under multipath Rayleigh fading channels of Doppler frequency 8.73 Hz, when having perfect CIRs. Again, we observe that for $m = 2$ and 4, the BLER performance of the CNN-based OFDM is in agreement with the baseline systems. At $m = 16$ and 64, the CNN system's BLER performance is slightly better than their conventional counterparts. This phenomenon indicates that the CNN-based autoencoder architecture can converge quickly

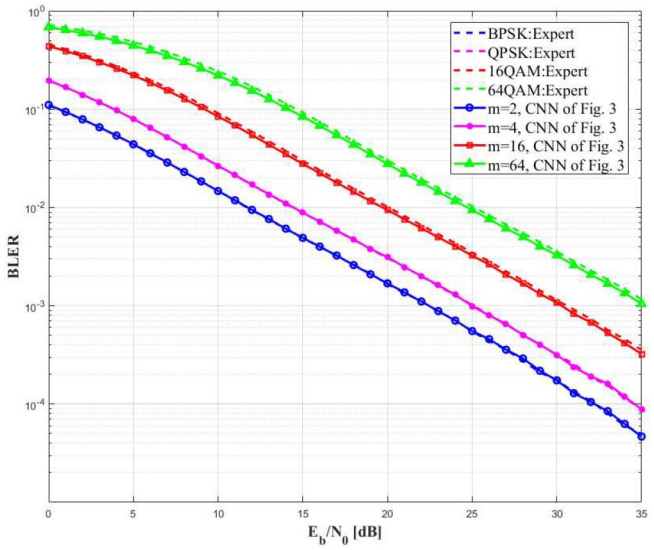


Fig. 7. BLER performance of the CNN-based OFDM autoencoder of Fig. 3 under multipath Rayleigh fading channels of Doppler frequency 8.73 Hz when having perfect CIRs and compared with corresponding expert OFDM systems.

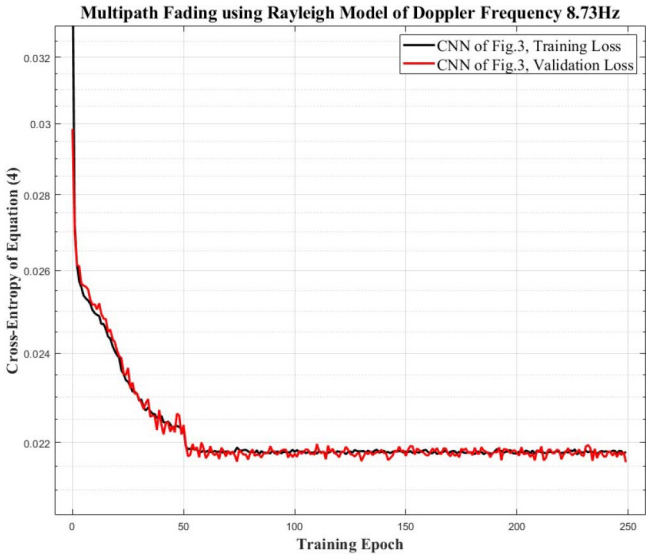


Fig. 9. Training loss and validation loss of the CNN-based OFDM autoencoder of Fig. 3 with $m = 64$ under multipath Rayleigh fading channels with Doppler frequency of 8.73 Hz.

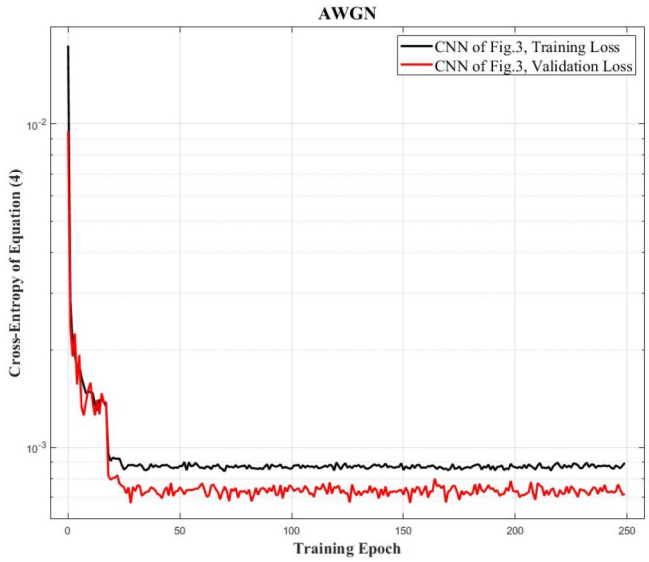


Fig. 8. Training loss and validation loss of the CNN-based OFDM autoencoder of Fig. 3 under AWGN channels with $m = 64$.

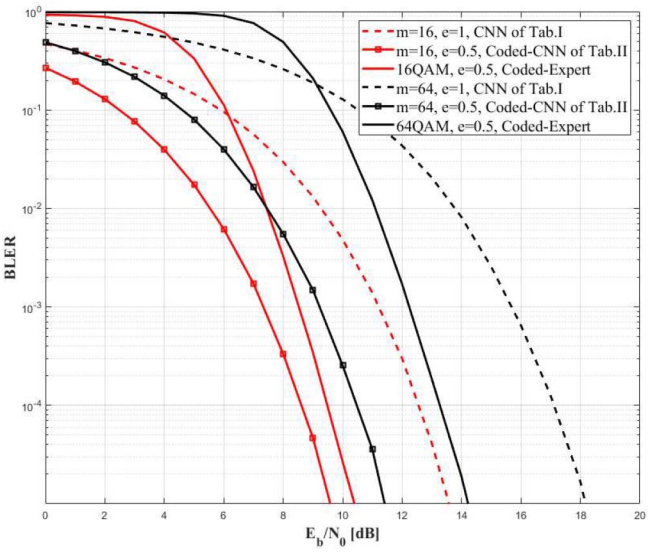


Fig. 10. BLER performance of the proposed CNN-based OFDM autoencoder of rate $e = 1$ and coded CNN-based OFDM autoencoder using LSTM of $e = 1/2$ under AWGN channels, when compared with corresponding expert OFDM systems using CC of $e = 1/2$.

subject to not only simple AWGN conditions but also to more dynamic multipath fading environments. In other words, CNNs are flexible enough to learn useful signal representations under different channel variations at the transmitter, while distinguishing different symbols at the receiver after channel impairment.

In order to further illustrate the CNN structure's quick convergence, Figs. 8 and 9 plot the training loss and validation loss of the proposed CNN-based OFDM autoencoder under AWGN and multipath fading channels with a Doppler frequency of 8.73 Hz. CNN's learning process is further complicated by setting $m = 64$. However, despite high data rate and dynamic channel environments, we observe in Figs. 8 and 9 that the convergence happens around training over 25 and 50 epochs. Observe that the autoencoder working under fading channels of Fig. 9 converges slower than that under AWGN channels of Fig. 8. That is because the fading channels are more complicated than the AWGN channels, the proposed autoencoder requires more data to gain comprehensive knowledge of the surrounding environments. Besides, the training loss and validation loss match quite well, which means the issue of overfitting can be alleviated by a proper design of the CNN model.

Fig. 10 shows the BLER performance of the coded CNN-based OFDM autoencoder using LSTM of Table II under AWGN channels. In the case of $m = 16$ and $m = 64$, it can be clearly seen that the BLER performance of the coded OFDM autoencoder having a rate of $e = 1/2$ is greatly improved

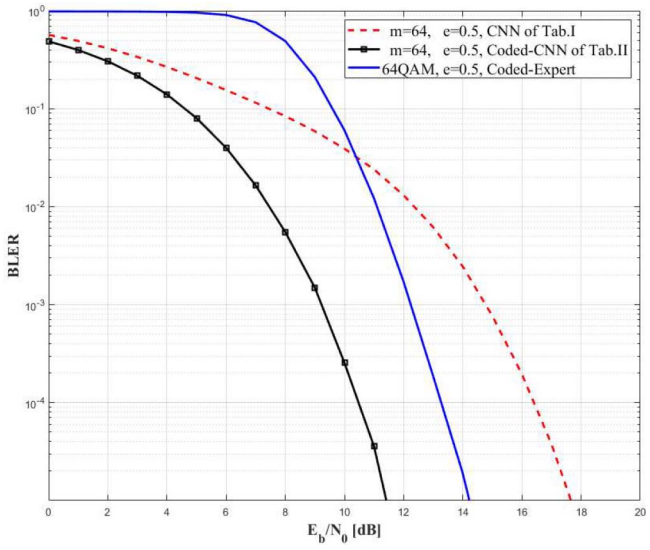


Fig. 11. BLER performance of the proposed CNN-based OFDM autoencoder of rate $e = 1/2$ and coded CNN-based OFDM autoencoder using LSTM of $e = 1/2$ under AWGN channels, when compared with corresponding expert OFDM systems using CC of $e = 1/2$.

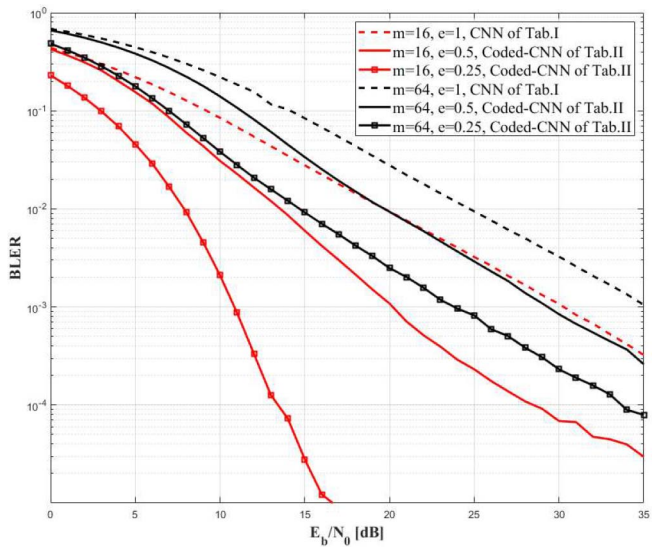


Fig. 12. BLER performance of coded CNN-based OFDM autoencoder using the LSTM of Table II having a rate of $e = 1, 1/2,$ and $1/4$, when transmitting over multipath Rayleigh fading channels of Doppler frequency 8.73 Hz, while having perfect CIRs.

over the uncoded OFDM autoencoder of Table I having a rate of $e = 1$. This means the proposed LSTM-aided model can achieve the coding gain. Even better, when compared with the corresponding expert OFDM system using CC of $e = 1/2$, the proposed half-rate-coded OFDM autoencoder archives a better BLER performance in the whole range of E_b/N_0. This is because the proposed combination of CNN and LSTM layers can better exploit and extract the features of the entire transmission sequence than the individually designed convolutional coded expert systems. That is, the autoencoder carries out the search for an optimal transmission sequence in a much larger overparameterized space, compared with the traditional handcrafted constellation points. Also, the constraint length of a CC is limited to 7, whereas the LSTM layer can correlate current input information to the whole vector having a length of $N*e$, which also assists the LSTM layer at the receiver to observe and extract information from the entire received signals. So, the LSTM layers at the transmitter are able to correlate each information vector much further in a transmission sequence than a simple CC, and the LSTM layers at the receiver can exact more information than a Viterbi decoder.

Then, Fig. 11 shows the BLER performance of three half-rate ($e = 1/2$) systems under AWGN channels, which are the proposed CNN-based OFDM autoencoder of Table I, the coded CNN-based OFDM autoencoder of Table II, and the corresponding CC-coded expert OFDM system. It can be clearly seen that the coded OFDM autoencoder outperforms both its expert OFDM counterpart and the uncoded OFDM autoencoder significantly, which exemplifies LSTM's ability of properly encoding and decoding information across a total number of 1000 time steps through learning [23], whereas the CC has a limited constraint length of 7. On the other hand, we also note that the uncoded CNN-based OFDM autoencoder beats the BLER performance of the CC-coded expert OFDM system in the low-SNR range, before the power of CC kicks in for high SNRs. Again, this proves that the key to approach the Shannon capacity is to correlate the transmitting sequence through coding.

Furthermore, the BLER performance of the coded CNN-based OFDM autoencoder using the LSTM of Table II having a rate of $e = 1, 1/2,$ and $1/4$ is shown in Fig. 12, when transmitting over multipath Rayleigh fading channels of Doppler frequency 8.73 Hz. It is obvious that our proposed autoencoder can indeed function under multipath fading channels and working with various rates, owing to the learning ability of neural networks. What is more, the BLER performance keeps improving significantly, with the help of increasing redundancy from $e = 1$ to 1/4. This means that the proposed coded CNN-based OFDM autoencoder using LSTM can exploit the redundancy efficiently through sufficient training.

B. Channel Estimation

There are two factors that contribute to the accuracy of the CE, namely, the noise level in the pilot data and the fast-changing nature of the marine wireless channel, characterized by the Doppler frequency.

Fig. 13 shows the BLER performance of the proposed CNN-based OFDM autoencoder of Fig. 3 using the CNN-based CE of Table III, when communicating over flat multipath Rayleigh fading channels. In order to demonstrate the proposed CE's learning ability, the OFDM system's BLER performance using perfect CIRs and using the state-of-the-art DL-based channel estimation [18] are also given as the benchmarkers. When $m = 2$ and 4, we observed that both [18] and the proposed CE can achieve the same performance as that of using the perfect CIRs.

Also observe in Fig. 13 that, in the case of $m = 16$ and 64, there exists an error floor for CE of [18] at high E_b/N_0 region. However, our proposed CNN-based CE of Table III remains capable of accurately reconstructing CIRs, which results in

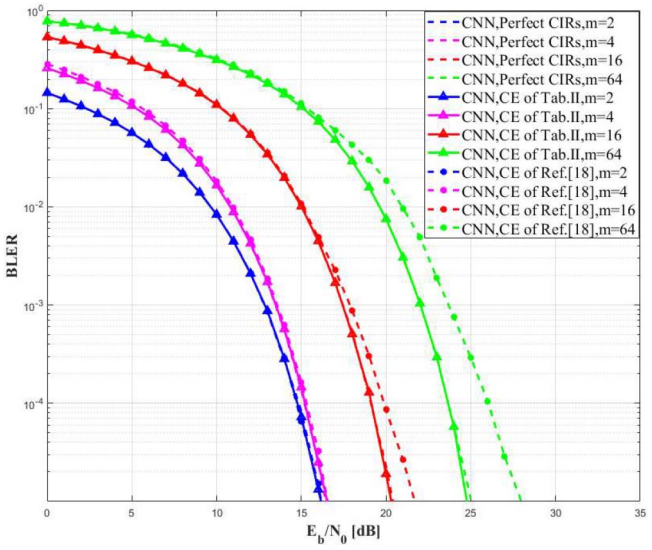


Fig. 13. BLER performance of the CNN-based OFDM autoencoder of Fig. 3 with CNN-based channel estimation (CE) of Table III, when communicating over flat multipath Rayleigh fading channels, while compared with OFDM systems having perfect CIRs and using CE of [18].

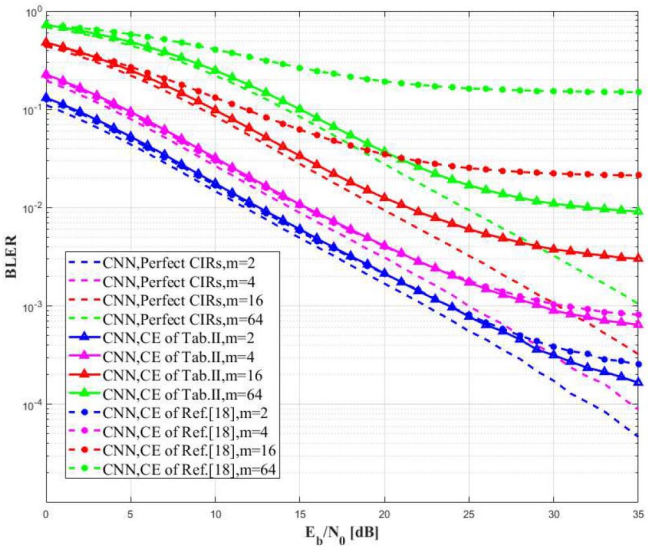


Fig. 14. BLER performance of the CNN-based OFDM autoencoder of Fig. 3 with CNN-based channel estimation (CE) of Table III, when communicating over fast multipath Rayleigh fading channels of Doppler frequency 8.73 Hz, while compared with OFDM systems having perfect CIRs and using CE of [18].

a close BLER performance with the case of using perfect CIRs. This improvement over [18] comes from the Dense-Net structure's deepened layers and multiple shortcut connections, which allow the proposed CE to infer better based on the pilot data.

For fast multipath Rayleigh fading channels having a Doppler frequency of 8.73 Hz, Fig. 14 plots the BLER performance of the CNN-based OFDM system of Fig. 3 using the CNN-based CE of Table III, when compared with using perfect CIRs and [18]. First, observe in Fig. 14 that all the BLER performance under fast fading environments are worse than that of flat fading channels in Fig. 13, owing to the increasing unpredictability of the channel. In this case, the comb pattern pilot signals of Fig. 2 are insufficient to capture the change of the environment. Second, all the CE algorithms in Fig. 14 exhibit error floors at high-SNR regions, and the error floors become more apparent for large m. However, the proposed CE algorithm of Table III still manages to beat the performance of [18]. Again, this means that the proposed Dense-Net structure can estimate the CIRs better than that of [18]. In addition, the proposed CNN-based CE uses fewer neural network parameters (total number of parameters = 19 192) than [18] (total number of parameters = 64 292). In other words, the proposed CNN-based CE invokes less computational cost.

V. Conclusion

In this article, in order to design an intelligent IoV communication system for challenging maritime channel environments, a novel OFDM autoencoder is proposed, which employs the CNN and LSTM layers to facilitate end-to-end learning. Furthermore, the proposed learning-based autoencoder is best suitable for communication systems that having a precise channel model is either impractical or impossible, which is not limited to maritime communications. Our OFDM autoencoder consists of multiple CNN layers in order to learn intricate signal representations for optimal transmissions while using fewer parameters than using FC layers. In addition, the proposed CNN architecture is able to converge to optimal solutions easily during training as of Figs. 8 and 9. When operating at low throughput, the proposed autoencoder exhibits similar BLER performance comparable to conventional QAM-assisted OFDM systems. The proposed system is able to outperform the traditional OFDM systems in high throughput settings as seen in Figs. 6 and 7. Furthermore, we also propose an LSTM-aided OFDM autoencoder of Table II, which has the capacity of achieving a better channel coding gain than the convolutional coded counterparts, which is shown in Figs. 10 and 12.

In the meantime, we also propose a new channel estimation algorithm for OFDM systems that marries the power of DL with the idea of super-resolution reconstruction, where a specially designed Dense-Net architecture is used to reconstruct low-resolution pilot information image into a high-resolution image. The Dense-Net architecture solves the traditional CNN's problem of vanishing gradient and excessive parameters through dense connections and feature multiplexing. Under slow fading channels, the proposed CE can estimate the CIRs near perfection. In the case of fast fading channels, the proposed CE achieves better performance than that of the existing DL-based algorithms with lower computational complexity. Finally, the proposed coded OFDM autoencoder and the proposed CNN-based CE require a significant amount of computation power to train offline, however, once the training is complete, the proposed autoencoders enjoy simple deployment by loading trained parameters and readily online training update.

In summary, we believe that the proposed OFDM autoencoder's generalization ability married with superior BLER

performance and adaptability to complex marine environments, which are blessed by LSTM-aided CNNs, hold to key to the evolution of IoV.

REFERENCES

[1] W. Zhen and B. Lin, "Maritime Internet of Vessels," in *Encyclopedia of Wireless Networks*, X. Shen, X. Lin, and K. Zhang, Eds. Berlin, Germany: Springer, 2019, pp. 1–9. [Online]. Available: https://link.springer.com/referenceworkentry/10.1007/978-3-319-32903-1_344-1

[2] L. Jiang, G. Huang, C. Huang, and W. Wang, "Data mining and optimization of a port vessel behavior behavioral model under the Internet of Things," *IEEE Access*, vol. 7, pp. 139970–139983, 2019, doi: 10.1109/ACCESS.2019.2943654.

[3] D. Chen, Y. Tian, D. Qu, and T. Jiang, "OQAM-OFDM for wireless communications in future Internet of Things: A survey on key technologies and challenges," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3788–3809, Oct. 2018, doi: 10.1109/JIOT.2018.2869677.

[4] S. Gao, M. Zhang, and X. Cheng, "Precoded index modulation for multi-input multi-output OFDM," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 17–28, Jan. 2017, doi: 10.1109/TWC.2017.2760823.

[5] 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, no. 2, pp. 98–105, Apr. 2017, doi: 10.1109/MWC.2016.1500356WC.

[6] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019, doi: 10.1109/COMST.2019.2904897.

[7] 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, doi: 10.1109/TCCN.2017.2758370.

[8] T. 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)*, Limassol, Cyprus, 2016, pp. 223–228.

[9] Q.-Z. Li, L.-W. Zhao, J. Gao, H.-B. Liang, L. Zhao, and X.-H. Tang, "SMDP-based coordinated virtual machine allocations in cloud-fog computing systems," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1977–1988, Jun. 2018, doi: 10.1109/JIOT.2018.2818680.

[10] H.-B. Liang, X. Zhang, J. Zhang, Q.-Z. Li, S. Zhou, and L. Zhao, "A novel adaptive resource allocation model based on SMDP and reinforcement learning algorithm in vehicular cloud system," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10018–10029, Oct. 2019, doi: 10.1109/TVT.2019.2937842.

[11] B. Mao *et al.*, "Routing or computing? The paradigm shift towards intelligent computer network packet transmission based on deep learning," *IEEE Trans. Comput.*, vol. 66, no. 11, pp. 1946–1960, Nov. 2017, doi: 10.1109/TC.2017.2709742.

[12] N. Ye, X.-M. Li, H. Yu, L. Zhao, W. Liu, and X. Hou, "DeepNOMA: A unified framework for NOMA using deep multi-task learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2208–2225, Apr. 2020, doi: 10.1109/TWC.2019.2963185.

[13] L. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. C-33, no. 7, pp. 665–675, Jul. 1985, doi: 10.1109/TCOM.1985.1096357.

[14] E. Balevi and J.-G. Andrews, "One-bit OFDM receivers via deep learning," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4326–4336, Jun. 2019, doi: 10.1109/TCOMM.2019.2903811.

[15] 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)*, Kalamata, Greece, 2018, pp. 1–5.

[16] M. Kim, W. Lee, and D.-H. Cho, "A novel PAPR reduction scheme for OFDM system based on deep learning," *IEEE Commun. Lett.*, vol. 22, no. 3, pp. 510–513, Mar. 2018, doi: 10.1109/LCOMM.2017.2787646.

[17] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001, doi: 10.1109/78.969514.

[18] Z. Zhao, M.-C. Vuran, F. Guo, and S. Scott. (2018). *Deep-Waveform: A Learned OFDM Receiver Based on Deep Complex Convolutional Networks*. [Online]. Available: https://arxiv.org/abs/1810.07181

[19] 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, doi: 10.1109/LWC.2017.2757490.

[20] X. Gao, S. Jin, C.-K. Wen, and G. Y. Li, "ComNet: Combination of deep learning and expert knowledge in OFDM receivers," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2627–2630, Dec. 2018, doi: 10.1109/LCOMM.2018.2877965.

[21] A. Brifman, Y. Romano, and M. Elad, "Unified single-image and video super-resolution via denoising algorithms," *IEEE Trans. Image Process.*, vol. 28, no. 12, pp. 6063–6076, Dec. 2019, doi: 10.1109/TIP.2019.2924173.

[22] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 652–655, Apr. 2019, doi: 10.1109/LCOMM.2019.2898944.

[23] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[24] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, 2017, pp. 4700–4708.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778.

[26] M.-H. Hsieh and C.-H. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels," *IEEE Trans. Consum. Electron.*, vol. 44, no. 1, pp. 217–225, Feb. 1998, doi: 10.1109/30.663750.

[27] M. Patzold, A. Szczepanski, and N. Youssef, "Methods for modeling of specified and measured multipath power-delay profiles," *IEEE Trans. Veh. Technol.*, vol. 51, no. 5, pp. 978–988, Sep. 2002, doi: 10.1109/TVT.2002.801747.

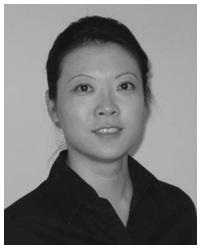
[28] J. Yu, X. Gao, D. Tao, X. Li, and K. Zhang, "A unified learning framework for single image super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 4, pp. 780–792, Apr. 2014, doi: 10.1109/TNNLS.2013.2281313.

[29] H. S. Mousavi and V. Monga, "Sparsity-based color image super resolution via exploiting cross channel constraints," *IEEE Trans. Image Process.*, vol. 26, no. 11, pp. 5094–5106, Nov. 2017, doi: 10.1109/TIP.2017.2704443.

[30] Z. M. Fadlullah *et al.*, "State-of-the-art deep learning: Evolving machine intelligence toward tomorrow's intelligent network traffic control systems," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2432–2455, 4th Quart., 2017, doi: 10.1109/COMST.2017.2707140.

[31] J. Wang *et al.*, "Wireless channel models for maritime communications," *IEEE Access*, vol. 6, pp. 68070–68088, 2018, doi: 10.1109/ACCESS.2018.2879902.

[32] J. I. Smith, "A computer generated multipath fading simulation for mobile radio," *IEEE Trans. Veh. Technol.*, vol. VT-24, no. 3, pp. 39–40, Aug. 1975, doi: 10.1109/T-VT.1975.23600.



Bin Lin (Member, IEEE) 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 is currently a Full Professor with the Department of Information Science and Technology, Dalian Maritime University. She was a Visiting Scholar with George Washington University, Washington, DC, USA, from 2015 to 2016. 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 Internet of Things.

Prof. Lin is an Associate Editor of *IET Communications*, and a Guest Editor of Special Issue on Internet of Things for Smart Ocean of the IEEE INTERNET OF THINGS JOURNAL.

Xudong Wang (Member, IEEE) received the B.S. degree in electronics engineering, the M.Sc. degree in circuit and system, and the Ph.D. degree in electronics and science technology from Xidian University, Xi'an, China, in 1989, 1992, and 2008, respectively.

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

Weihao Yuan received the B.S. degree from the Department of Electronic Information Science and Technology, Dalian Maritime University, Dalian, China, in 2018, where he is currently pursuing the master's degree with the Department of Information and Communication Engineering.

His research interest includes deep-learning-based communications.

Nan Wu (Member, IEEE) received the B.S. degree in electrical and electronics engineering from Dalian University of Technology, Dalian, China, in 2003, and the M.Sc. and Ph.D. degrees from the University of Southampton, 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, Gaithersburg, MD, USA. Since 2010, he has been an Associate Professor with the School of Electrical and Information Engineering, Dalian Maritime University, Dalian. He has authored one book, more than 50 articles. He also holds two patents. His research interests include deep-learning-based communications, visible light communication, indoor positioning, and MIMO.