

ComNet: Combination of Deep Learning and Expert Knowledge in OFDM Receivers

Xuanxuan Gao, *Student Member, IEEE*, Shi Jin, *Senior Member, IEEE*,
Chao-Kai Wen, *Member, IEEE*, and Geoffrey Ye Li, *Fellow, IEEE*

Abstract

In this article, we propose a model-driven deep learning (DL) approach that combines DL with the expert knowledge to replace the existing orthogonal frequency-division multiplexing (OFDM) receiver in wireless communications. Different from the data-driven fully connected deep neural network (FC-DNN) method, we adopt the block-by-block signal processing method that divides the receiver into channel estimation subnet and signal detection subnet. Each subnet is constructed by a DNN and uses the existing simple and traditional solution as initialization. The proposed model-driven DL receiver offers more accurate channel estimation comparing with the linear minimum mean-squared error (LMMSE) method and exhibits higher data recovery accuracy comparing with the existing methods and FC-DNN. Simulation results further demonstrate the robustness of the proposed approach in terms of signal-to-noise ratio and its superiority to the FC-DNN approach in the computational complexities or the memory usage.



Index Terms

Deep learning, wireless communications, OFDM.

I. INTRODUCTION

Deep learning (DL) has gained great successes in the fields of computer vision and natural language processing, among others, and has been considered for application in wireless communications since then. The potential applications of DL in physical layers are discussed in [1], [2] and

X. Gao and S. Jin are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: gaoxuanxuan@seu.edu.cn; jinshi@seu.edu.cn).

C.-K. Wen is with the Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 804, Taiwan (e-mail: chaokai.wen@mail.nsysu.edu.tw).

G. Y. Li is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (email: liye@ece.gatech.edu).

[3]. The data-driven DL approach in [4] adopts a fully connected deep neural network (FC-DNN) and replaces all modules at the orthogonal frequency-division multiplexing (OFDM) receiver. The abovementioned literature challenges the conventional OFDM receiver and treats the receiver as a black box. But, it does not exploit the expert knowledge¹ in wireless communications, which in turn renders the FC-DNN-based receiver unexplainable and unpredictable. In addition, the data-driven method relies on a huge amount of data to train a large number of parameters, thus converges slowly and has high computational complexity.

To address the above issues, the model-driven DL approach can be used instead. A general model-driven DL framework in [5] can overcome the heavy demand of a huge amount of training data. Moreover, the model-driven DL network can clearly explain the specially designed model family by using domain knowledge to facilitate further performance improvement. In the field of wireless communications, all modules in transceivers have been rigorously developed, which subsequently enables to set the existing algorithms as the fundamentals of the model family in model-driven DL approaches. The superiority of introducing expert knowledge into wireless communications to form a model-driven DL solution has been demonstrated in the examples of the radio transformer network (RTN) [1], the channel state information (CSI)-aided MIMO communication [6] and the PAPR reducing network (PRNet) [7]. A comprehensive overview of model-driven DL for physical layer communications can be found in [8].

In this article, we propose a model-driven DL architecture, called ComNet, to replace the conventional or FC-DNN OFDM receiver [4], which combines DL with expert knowledge in wireless communications. The proposed ComNet receiver uses DL to facilitate existing receiver models, such as channel estimation (CE) module and signal detection (SD) module, rather than replacing the receiver with an entire DL architecture and then integrating communication information, such as the RTN [1], CSI-aided MIMO communications [6] and PRNet [7]. This model-driven DL approach shows better performance comparing with the traditional LMMSE-MMSE method and FC-DNN [4] and exhibits relatively faster convergence speed with fewer parameters comparing with the FC-DNN OFDM receiver [4].

¹The term “expert knowledge” used in this article means block-based architectures and algorithms in conventional wireless communications and differs from the knowledge-based expert system in the artificial intelligence field.

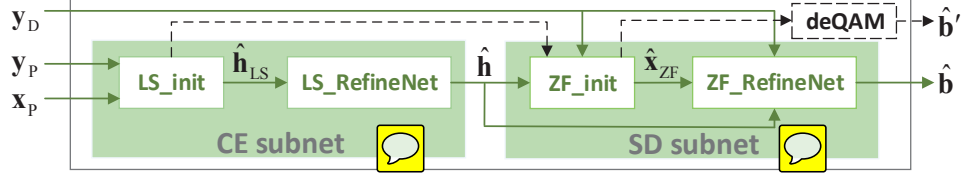


Fig. 1: ComNet receiver architecture. The two subnets use traditional communication solutions as initializations, and apply DL networks to refine the coarse inputs. The dotted short-path provides a relatively robust candidate of the binary symbols recovery.

II. COMNET

This section presents the proposed ComNet receiver for the OFDM system. The architecture and the details of the DL-based subnets, including the CE and SD subnets, are elaborated in Section II A. In Section II B, initializations of network weights, the choice of cost function and optimizer, and configurations of hyper-parameters are explained.

A. ComNet Architecture

In the OFDM system, the transmitted signals consist of the transmitted data vector, \mathbf{x}_D , and the pilot symbol vector, \mathbf{x}_P , which is known to the receiver. Correspondingly, the received signals include the received data vector, \mathbf{y}_D , and the received pilot symbol vector, \mathbf{y}_P . The conventional OFDM receiver recovers the estimation of the transmitted binary data $\hat{\mathbf{b}}$ given frequency domain signals, \mathbf{y}_D , \mathbf{y}_P , and \mathbf{x}_P through CE, SD, and quadrature amplitude modulation (QAM) demodulation sequentially.

Fig. 1 illustrates the architecture of the ComNet receiver. The inputs and outputs of the ComNet receiver are similar to that of the conventional OFDM receiver while the ComNet receiver adopts two cascaded DL-based subnets to replace the conventional OFDM receiver. Instead of using straightforward FC-DNN as in [4] to estimate the transmitted data in a brute force manner, in the proposed ComNet receiver, the CE and SD subnets use traditional communication solutions as initializations and apply DL networks to refine the coarse inputs. The ComNet receiver also takes full advantage of the conventional methods and connects them to form a relatively robust recovery to adapt to various scenarios.

Fig. 2 shows the CE subnet. Its input is the least-square (LS) CE, obtained by

$$\hat{h}_{LS}(k) = \frac{y_P(k)}{x_P(k)}, \quad (1)$$

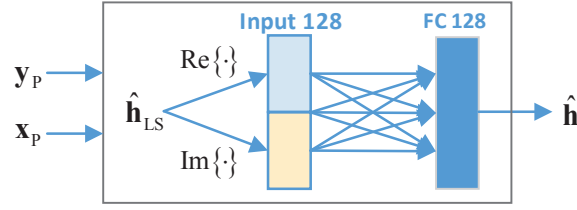


Fig. 2: CE subnet. A subnet type initialized by LS CE. Then the real-valued initialization is refined by LS_RefineNet.

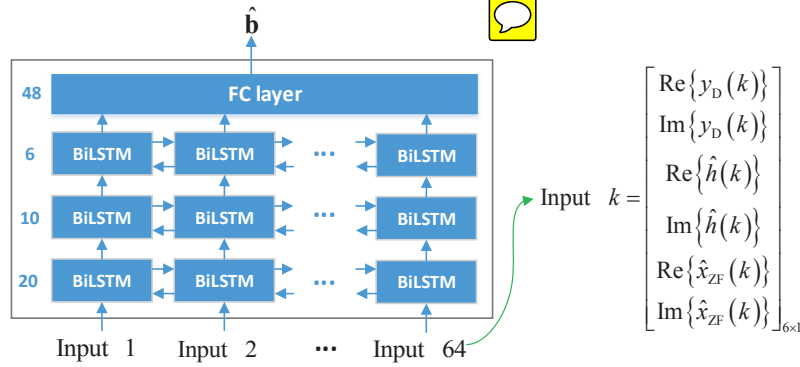


Fig. 3: Detection subnet with BiLSTM-Detection as ZF_RefineNet. A subnet type initialized by ZF solution, in which ZF_RefineNet adopts BiLSTM-Detection with a FC-DNN layer.

where $x_p(k)$ and $y_p(k)$ are the pilot symbol and the corresponding received symbol at the k -th subcarrier. Then $\hat{\mathbf{h}}_{LS}$ is used by the LS_RefineNet to generate accurate CE $\hat{\mathbf{h}}$, where the LS_RefineNet is a one-layer DNN. Its input is a 128-dimensional real-valued signal vector that consists of the real and imaginary parts of $\hat{\mathbf{h}}_{LS}$. The number of neurons of the following layer is 128 and these neurons have no activation function, that is, it is a linear channel estimator.

In the SD subnet, the input is just zero-forcing (ZF) SD of the transmit symbol, obtained by

$$\hat{x}_{ZF}(k) = \frac{y_D(k)}{\hat{h}(k)}. \quad (2)$$

The \hat{x}_{ZF} is used by the ZF_RefineNet to predict the binary data from 8 symbols on 8 consecutive subcarriers. For an OFDM system with 64 subcarriers, 8 independent SD subnets are needed. In summary, the ZF_RefineNet uses \hat{x}_{ZF} , $\hat{\mathbf{h}}$, and \mathbf{y}_D to get more accurate estimation of the transmit data. Depending on different requirements on receiver complexity and data recovery accuracy, we propose two different forms of ZF_RefineNet.

- **FC-SD** involves a two-layer FC-DNN with 120 and 48 neurons on each layer. The input is the concatenation of the real and imaginary parts of \hat{x}_{ZF} . The activation function of the

hidden layer uses a ReLU function, $f_{\text{Re}}(a) = \max(0, a)$, whereas that of the output layer is the logistic sigmoid function, $f_{\text{Si}}(a) = \frac{1}{1+e^{-a}}$.

- **Bi-directional long short-term memory (BiLSTM)-SD** involves a three-layer 64-time steps BiLSTM network [9] with 20, 10 and 6 hidden units for each layer, followed by an one-layer FC-DNN with 48 neurons, as in Fig. 3. Considering the performance degradation of the ZF SD, the inputs of BiLSTM-SD integrate \mathbf{y}_D and $\hat{\mathbf{h}}$ as well. The activation function of the output layer is the logistic sigmoid function as before.

The abovementioned 48 outputs correspond to 48 bits to be estimated from eight consecutive subcarriers with 6 bits for each symbol for 64-QAM. Because the logistic sigmoid function maps the input to the interval, $[0, 1]$, the received binary symbol will be “1” if the output is larger than 0.5 and “0” otherwise. However, the number of layers and that of neurons in each layer, except for the last layer, depend on the empirical trials.

B. Training Specification

To accelerate the training process, the initialization of the network weights is considered. The CE subnet is initialized by the real-valued linear minimum mean-squared error (LMMSE) CE weight matrix $\tilde{\mathbf{W}}_{\text{LMMSE}}$ from

$$\tilde{\mathbf{h}}_{\text{LMMSE}} = \tilde{\mathbf{W}}_{\text{LMMSE}} \tilde{\mathbf{h}}_{\text{LS}}, \quad (3)$$

where

$$\tilde{\mathbf{W}}_{\text{LMMSE}} = \begin{bmatrix} \text{Re}\{\mathbf{W}_{\text{LMMSE}}\} & -\text{Im}\{\mathbf{W}_{\text{LMMSE}}\} \\ \text{Im}\{\mathbf{W}_{\text{LMMSE}}\} & \text{Re}\{\mathbf{W}_{\text{LMMSE}}\} \end{bmatrix}. \quad (4)$$

$\tilde{\mathbf{h}}_{\text{LMMSE}}$ and $\tilde{\mathbf{h}}_{\text{LS}}$ are the concatenation of real and imaginary parts of LMMSE CE $\hat{\mathbf{h}}_{\text{LMMSE}}$ and LS CE $\hat{\mathbf{h}}_{\text{LS}}$, respectively. In particular, the LMMSE CE weight matrix, $\tilde{\mathbf{W}}_{\text{LMMSE}}$, adopts the method in [10]. The multiplicative weights in the FC layers of the SD subnet are initialized by the method in [11].

After initialization, the DL networks are trained by minimizing the cost between $\hat{\mathbf{b}}$ and the raw transmit binary symbol \mathbf{b} to adjust the network parameters. The training data are obtained from simulation under system configurations in the next section. We adopt the mean-squared error cost function as in [4] and the adaptive moment estimation (Adam) optimizer [12] for both subnets. The two subnets are trained sequentially in TensorFlow, in which the CE subnet is trained for 2,000 epochs and then fixed, followed by 5,000 epochs training for the SD subnet. The end-to-end comparison shows that sequential training can guarantee the optimality of each

block and speed up the training process with fewer network parameter requirements. Each epoch utilizes 50 mini-batches for a total batch size of 1,000. The learning rate is set as a staircase function to realize training epochs with initial values of 0.001 and decreased 10-fold every 1,000 epochs for the CE subnet and decreased 5-fold every 2,000 epochs for the SD subnet.

III. NUMERICAL RESULTS

Simulation is conducted under three cases as in [4], linear, cyclic prefix (CP) removal, and clipping, which are corresponding to no postfix, “_CP” postfix and “_CR” postfix in figures. Since the main contribution of this article is to propose a novel model-driven OFDM receiver architecture instead of resolving the nonlinearity, traditional nonlinearity compensatory methods are not delved. The “SameSNR” markers represent the results when the ComNet receiver is trained and deployed both under $\text{SNR} = 5 \text{ dB}$. Simulation results are compared in terms of accuracy and complexity among the proposed ComNet receiver, FC-DNN receiver [4] and the traditional communication methods [10].

System configurations of the simulation are similar to [4] as follows. The OFDM system contains 64 subcarriers with 16 samples of CP and each frame contains one pilot OFDM symbol and one data OFDM symbol. The mapping of 64-QAM adopts the long-term evolution (LTE) standard. The channel is a WINNER II channel under C1 scenario NLOS case in 2.6 GHz. The ComNet receiver utilizes FC-SD in the linear case, whereas BiLSTM-SD is adopted in the nonlinear cases.

We employ the following concise conventions in the subsequent discussion:

- **ComNet-BiLSTM**: Proposed ComNet architecture with BiLSTM-Detection ZF_RefineNet
- **ComNet-FC**: Proposed ComNet architecture with FC-Detection ZF_RefineNet
- **FC-DNN**: FC-DNN in [4], but modified such that the number of neurons in the output layer is changed from 16 to 48 to render it suitable for 64-QAM
- **LMMSE-MMSE**: Traditional LMMSE CE and minimum mean-squared error (MMSE) SD
- **Y/H_true**: The quotient of \mathbf{y}_D and the true frequency domain channel \mathbf{h} , which can realize the maximum-likelihood solution in the linear case.

A. CE Subnet

A benefit of using ComNet receiver against FC-DNN is to access accurate CSI, which is useful for channel analysis and CSI feedback in downlink transmission. Fig. 4 shows MSE performance of the ComNet CE subnet and the LMMSE method under linear and CP removal cases. From the

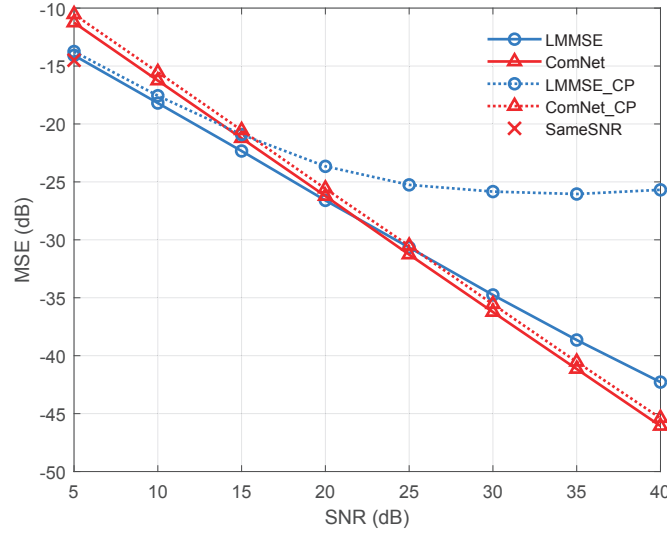


Fig. 4: MSE curves of ComNet and traditional methods under linear case and CP removal case.

figure, the CE subnet can better rectify the effect introduced by CP removal compared with the traditional LMMSE CE. This is as a result that the training process of LS_RefineNet modifies the network multiplicative weights from the initialized value, $\tilde{\mathbf{W}}_{\text{LMMSE}}$, to proper values through minimizing the channel MSE with the Adam optimizer.

B. SD Subnet

1) *Linear Case:* Fig. 5 compares the bit-error rate (BER) curves of ComNet-FC and existing methods under the linear case, where a basic OFDM system without nonlinear effects is considered. From the figure, the BER of the proposed ComNet receiver is closest to the ideal bound $\mathbf{Y}/\mathbf{H}_{\text{true}}$ compared with FC-DNN and LMMSE-MMSE. The required signal-to-noise ratio (SNR) for the ComNet receiver to reach $\text{BER} = 10^{-3}$ is 1 dB better than FC-DNN and LMMSE-MMSE. But there is also an 1 dB gap between the ComNet receiver and the ideal bound. Extra simulation suggests that the ComNet receiver obviously outperforms the FC-DNN under longer delay spread.

The model-driven approach, ComNet-FC receiver, has only one-eighth amount of parameters compared with the data-driven approach of FC-DNN [4]. Furthermore, the ComNet-FC receiver needs only 200 epochs to converge while FC-DNN needs approximately 2,000 epochs to reach the same BER level. This result demonstrates the superior convergent speed and minimal parameter requirements of the model-driven DL approach.

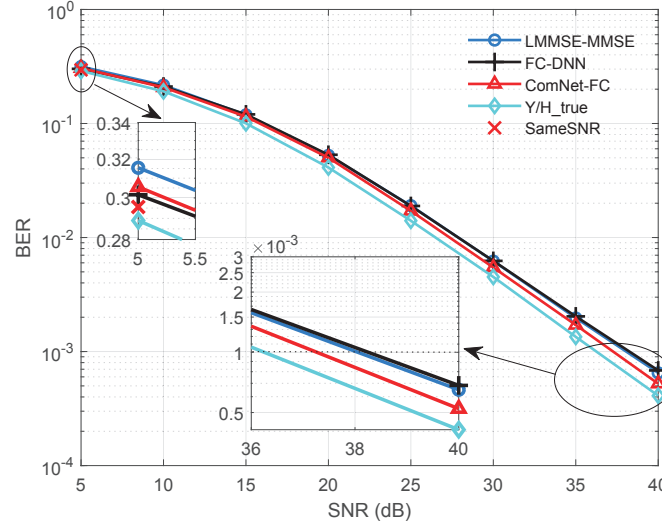


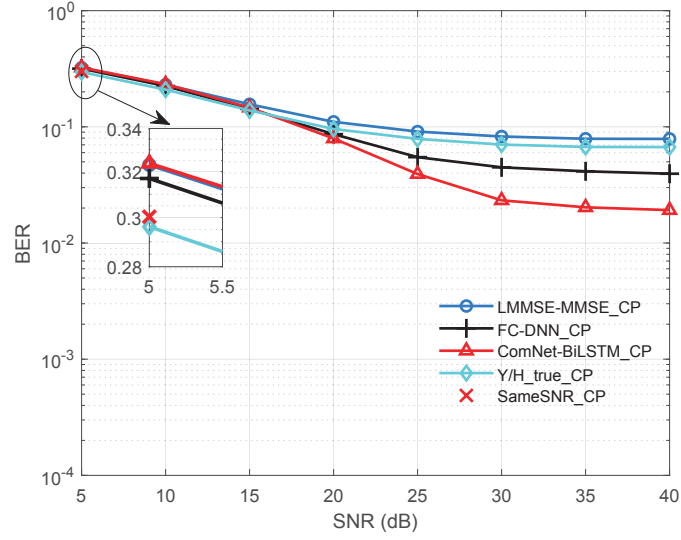
Fig. 5: BER curves of ComNet and competing methods under linear case.

2) *CP Removal*: The CP is introduced into the OFDM system to mitigate the inter-symbol interference (ISI) caused by multipath channels, which also decreases transmission efficiency and increases energy costs. Fig. 6(a) compares the BER performance of ComNet-BiLSTM with the other methods when the CP is omitted. In this case, the traditional LMMSE-MMSE method becomes saturated when SNR equals 20dB while the DL-based approaches including FC-DNN and ComNet-BiLSTM perform better in resolving ISI. In particular, ComNet-BiLSTM has about 50% BER of FC-DNN when SNR is over 25dB, which suggests that ComNet-BiLSTM has the ability to recover transmit symbols more accurately than the other approaches for OFDM systems without CP. This ability benefits from the BiLSTM recurrent neural network that is designed to utilize the inner-relationship of ISI of the sequential data.

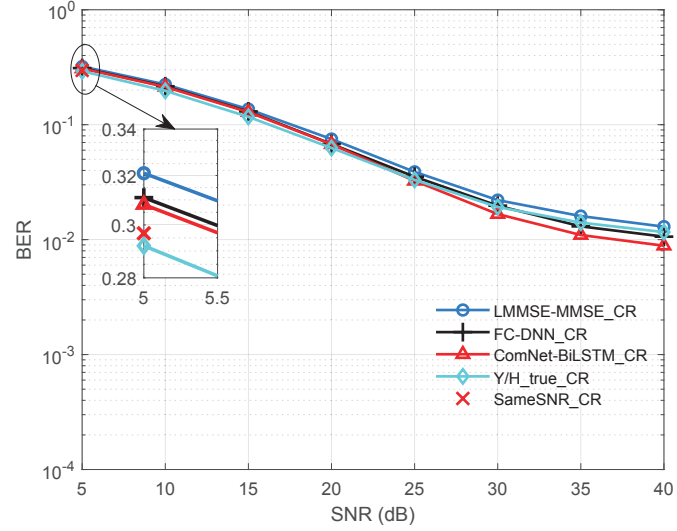
3) *Clipping*: One of the most detrimental characteristics in OFDM is the high peak-to-average power ratio (PAPR) [10]. A common method to reduce PAPR is the clipping operation, which is applied to the time-domain transmitted signal as in [4] and at the same time causes nonlinear distortion on the signal. Fig. 6(b) shows the BER curves of ComNet-BiLSTM and competing methods with nonlinear distortion of clipping with the clipping ratio of 1.6. From the figure, the ComNet-BiLSTM obtains the lowest BER among all methods.

C. Performance Analysis

1) *Robustness of SNR Mismatching*: The abovementioned results are obtained by training the ComNet receiver offline under SNR = 40 dB while deploying it online under arbitrary SNRs,



(a)



(b)

Fig. 6: BER curves of ComNet and competing methods under nonlinear cases: (a) CP removal case marked as “_CP”, and (b) clipping case denoted as “_CR”.

which are SNR mismatched results, whereas “SameSNR” markers in Figs. 4-6 represent SNR matched results. The difference of SNR mismatched and matched results indicates that the SNR mismatching causes an MSE loss of approximately 3 dB when SNR = 5 dB. Nonetheless, the SNR mismatching leads to slight BER performance loss, which suggests the robustness of the ComNet receiver against SNR mismatching when recovering binary symbols.

TABLE I: Complexity analysis for ComNet and competing methods

	# of FLOPs	Memory	Intensity	Time
ComNet-BiLSTM	10.40M	2.40MBytes	4.33	7.2e-6s
ComNet-FC	0.37M	1.22MBytes	0.30	1.2e-6s
FC-DNN	4.62M	9.30MBytes	0.49	1.2e-6s
LMMSE-MMSE	1.6M	–	–	–

2) *Application Complexity*: TABLE I compares the complexities of the receivers in terms of the amount of floating-point multiplication-adds (FLOPs), memory usage, computational intensity, and time consumption required to complete a single-forward pass of one OFDM symbol. To achieve better BER performance than FC-DNN within the same time period, ComNet-BiLSTM needs more than twice FLOPs than FC-DNN with approximately one-fourth memory whereas ComNet-FC only needs 0.37 million FLOPs and 1.22 MBytes of memory. Compared with traditional methods, the ComNet-FC consumes fewer FLOPs than LMMSE-MMSE because the parameters in ComNet receiver are settled once the training process is completed, whereas the weighted matrix of LMMSE channel estimation has to be recomputed along with the varying of channel states.

IV. CONCLUSION

In this article, we demonstrate the benefits of the proposed ComNet receiver architecture to recover the transmit data in the OFDM system with linear and nonlinear distortions. Although the coarse-to-fine idea in the ComNet is intuitive, it offers deeper insights into its implications. From a communication viewpoint, the nonlinear activation functions in DL neural network introduce nonlinearity into the SD module, which constitutes a nonlinear signal detector. From the perspective of model-driven DL assisted by communication intelligence, the useful novel features can be created manually. Moreover, these novel features can accelerate the training process, which then results in efficient deployment performance. The idea of combining DL with expert knowledge in the ComNet receiver inspires the future work of applying model-driven DLs to the physical layers in wireless communications.

REFERENCES

- [1] 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.

- [2] T. Wang, C. K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *China Commun.*, vol. 14, no. 11, pp. 92–111, Nov. 2017.
- [3] Z.-J. Qin, H. Ye, G. Y. Li, and B.-H. Juang, "Deep learning in physical layer communications," *submitted to IEEE Commun. Mag.*, July 2018/ also at <https://arxiv.org/abs/1807.11713>.
- [4] 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.
- [5] Z. Xu and J. Sun, "Model-driven deep-learning," *National Sci. Rev.*, vol. 5, no. 1, pp. 22–24, 2018. [Online]. Available: <http://dx.doi.org/10.1093/nsr/nwx099>
- [6] T. J. O'Shea, T. Erpek, and T. C. Clancy, "Deep learning based MIMO communications," *arXiv preprint arXiv:1707.07980*, 2017.
- [7] 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, 2018.
- [8] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Model-driven deep learning for physical layer communications," *submitted to IEEE Wireless Commun.*, Sep 2018/ also at <https://arxiv.org/abs/1809.06059>.
- [9] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM networks," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IJCNN)*, vol. 4, July 2005, pp. 2047–2052.
- [10] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Wiley Publishing, 2010.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.
- [12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Comput. Sci.*, 2014.