Multi-modem Implementation Method Based on Deep Autoencoder Network

Peng Wei[1], Ruimin Lu[1(✉)], Shilian Wang[2], and Shijun Xie[1]

[1] 63rd Research Institute, National University of Defense Technology, NanJing 210007, China
weipengcss@163.com, Luruiminpaper@163.com, xsjxsj520@163.com

[2] College of Electronic Science, National University of Defense Technology, Changsha 410073, China
wangsl@nudt.edu.cn

Abstract. With the fierce competition for electromagnetic spectrum, the development of intelligent satellite communication systems with intelligent waveform generation and reconstruction capabilities is an effective means to adapt satellite communication system to the harsh electromagnetic environment. In this paper, a 10-layer deep autoencoder network (DAN) is designed, and 2-ary to 64-ary modem are implemented based on this 10-layer DAN. During this process, a unified loss function and a unified optimization algorithm are utilized to train and test the 10-layer DAN. Finally, the demodulation performance, that is close to, consistent with or better than that of traditional MPSK or QAM is obtained. The above-mentioned DAN and its training method provide a new way for waveform generation and reconstruction in intelligent communication satellites. In addition, the high-order modulation constellation generated by this 10-layer DAN is quite different from the traditional modulation method and very difficult to distinguish linearly, which is beneficial to improve the anti-intercept ability of the satellite communication waveform.

Keywords: Satellite communications · Deep learning · Deep autoencoder network · Modem · Anti-interception

1 Introduction

Satellite communication has been more and more widely used due to its advantages of wide coverage, long distance transmission, and freedom from terrain. However, there are many aspects that need to be resolved, such as single satellite function, difficulty in upgrading and maintaining, and weak ability to adapt to complex electromagnetic environments. On the other hand, the third wave of artificial intelligence represented by deep learning and reinforcement learning is infiltrating or even subverting the traditional industry at an unprecedented

Q. Wu et al. (Eds.): WiSATS 2020, LNICST 358, pp. 487–501, 2021.
https://doi.org/10.1007/978-3-030-69072-4_40

speed, depth and breadth. In strategic games, image detection classification, target recognition, voice processing, automatic translation, unmanned driving and other fields, it has shown "wisdom" beyond the human brain [1]. Therefore, combining artificial intelligence with satellite communications and using the most advanced artificial intelligence research results to improve or reshape traditional satellite communications can effectively improve the performance and security of satellite communication systems and increase their ability to adapt to complex electromagnetic environments. This is critical to winning the future intelligent warfare.

Adaptive code modulation (ACM) as a technology that can achieve a good compromise between signal capacity and reliability by adjusting the coding rate and modulation order under different channel conditions has been widely used in wireless communication systems [2]. However, this method usually needs to implement multiple coding and modulation with different orders and different rates separately, and then select a certain combination of coding and modulation to complete the communication according to the perception and decision of the channel environment. This way makes the design, implementation and hardware cost of the communication system multiply. To this end, a realization method of 2-ary to 64-ary modem based on the unified DAN [3], which is different from the traditional modem based on expert design, is proposed in this paper. In this new method, a 10-layer DAN is adopted, training set is generated using random number, and a unified optimization algorithm, a unified loss function, and some certain signal-to-noise ratio (SNR) are utilized to train the 10-layer DAN. Test results show that, the 10-layer DAN can convergence in a relatively short time for all of 2-ary to 64-ary modem, and the demodulation performance is close to, consistent with or batter than that of traditional modem. So it provides a new way of realizing the waveform generation and reconstruction or ACM for satellite communication intelligent.

The research related to this paper mainly includes the end-to-end communication system based on unsupervised learning proposed by T. J. O'Shea et al. [4–7]. Among them, unsupervised representation learning of radio communication signals was studied in raw sampled time series representation, and convolutional autoencoder was used to learn the modulation basis function and visually recognize their relationship to the analytic bases used in digital communications in [4]. By optimizing the reconstruction loss during channel autoencoder through a series of channel regularizes, new modulation schemes were learned in [5], which blurs the boundary between modulation and error correction decoding, and provides similar capacity and error performance with lower complexity and without expert design. By Combining with random delay, frequency difference, phase difference, delay extension and other channel impairments, the bit error performance under different loss functions, different network connection methods (DNN, CNN), and different activation functions was evaluated in [5]. Results showed that the bit error performance can exceed the traditional QPSK demodulation under certain conditions. Furthermore, in order to solve the performance degradation in the case of large delay expansion, an attention mechanism was introduced to expand the use of the model. In [7], the advantages of deep learning in the field of communication, such as complex channel learning, overall optimization, efficient approximation of arbitrary functions, distributed parallel architecture and dedicated high-efficiency processing chips were pointed out, end-to-end modeled with autoencoder network was setup, and performance close to that of traditional communication systems was achieved under some specific parameters. In [6], the design and training method of the wireless communication physical layer based on the deep generation adversarial network (GAN) were proposed, in which the deep GAN was used to learn wireless channel characteristics and establish a passed back from receiver to sender. Hao He et al. also completed a similar work in [8], and the performance similar to the traditional communication was acquired under additive Gaussian noise or Rayleigh fading channels. Using software radio and open source deep learning software library, S. Dorner et al. [9] built, trained and run a complete communication system based entirely on deep neural networks (DNN), extended block-based transmissions to continuous data transmission. An additional frame synchronization module based on deep network was introduced to solve the synchronization problem. Without a lot of overparameter tuning, the performance deteriorated within 1 dB compared with traditional software radio demodulation. Yang Yaodong et al. [10] applied the stack sparse autocoder (SSAE) network to the demodulation of multi-position phase shift keying (MPPSK), and obtained a demodulation performance of 1–2 orders of magnitude better than the traditional demodulation. Huang Yuanyuan et al. [11] applied the deep confidence network to the feature extraction and recognition of communication signals, and the simulation showed a 0.4 dB performance improvement over the traditional modem method for MPSK modem.

In the overview of applying deep learning to wireless communication, Q. Mao et al. [1] comprehensively introduced the advantages of deep learning and deep reinforcement learning applied to wireless networks. On this basis, they listed the application of deep reinforcement learning at all layers of the system and analyzed the future research trends and challenges. Zhang Jing et al. [12] summarized the development history of wireless communication. Based on the introduction of various deep learning network structures, they focused on summarizing the channel estimation of deep learning in large-scale MIMO scenarios, signal detection in OFDM systems, CSI feedback and reconstruction, channel decoding, and end-to-end wireless communication system applications. From the perspective of system design mode, adaptability to channel changes and the current powerful computing power based on parallel GPU, Guiguan et al. [13] pointed out the potential of deep learning in the field of communication, provided the application of deep learning network in modulation recognition and beam forming, and analyzed its structure and superior performance. Finally, they pointed out the lack of common data sets, the lack of common models suitable for communication scenarios, the excessive number of parameters in the application to small terminals, as well as the physical layer security problems, and gave a preliminary solution.

In above work, the most similar to the research in this paper is [5,6]. In these two references, modems of multiple (n, k) combinations, including (2, 2), (4, 4), (7, 4), (8, 8), were realized through autoencoder network. When a symbol $s \in M$ is sending, the width of the input layer is $M = 2^k$, the width of the output layer the sending end is n, and the receiving end performs M classification output to restore the sending symbol $\hat{s} \in M$. A total $(2M + 1)(M + N) + 2M$ of training parameters are required. The minimization cross-entropy loss function is used as the optimization goal, and the minimum gradient descent method (SGD) is used for training. Although the performance of this method consistent with or better than that of traditional code modulation method. But a different network structure was utilized for each (n, k), which is not conducive to the efficient implementation and reconstruction of multiple code modulations. In addition, only I and Q signals can be transmitted on the actual channel, while in [5,6], n-channels of data are transmitted directly, which brings about spatial diversity gain, thus obtains superior performance to the traditional encoding and modulation. Therefore, this comparison is little far from reasonable.

In view of the deficiencies of above work, we focus on the realization of multi-modem based on a unified DAN architecture, in the hope of providing a new implementation method for intelligent and efficient waveform generation and reconstruction for satellite communication systems, and improving the anti-intercept ability of the waveform. Practical work is as follows:

1) A 10-layer DAN is designed, and base on this 10-layer DAN, 2-ary to 64-ary modems are realized with a unified DL network architecture, a unified loss function and a unified optimization algorithm. The parameters of each layer and the optimization training method are given.
2) For each of 2-ary to 64-ary modem, the modulation constellation, symbol mapping relationship, network convergence speed and performance comparison with traditional demodulation methods are simulated. The results show that, under the white noise channel, the performance of 2-ary and 4-ary modem based on the 10-layer DAN is completely consistent with that of BPSK and QPSK respectively, the performance of 8-ary and 64-ary modem is better than that of 8QAM and 64QAM respectively, and the performance of 16-ary and 32-ary modem is close to that of 16QAM and 32QAM respectively.

The following chapters of this paper are arranged as follows: In Sect. 2, the 10-layer DAN architecture is designed, and the parameter configuration of each layer, loss function and optimization algorithm are tried and selected. The training results and demodulation performance of multi-modem based on the 10-layer DAN are presented in Sect. 3. In Sect. 4, a brief summary of the paper is given, and further research directions is proposed.

2 Multi-modem Model Based on DAN

After many trial and error, we find that the 10-layer DAN model shown in Fig. 1 can achieve the best performance for 2-ary to 64-ary modem with fewer parameters. The 10-layer DAN is mainly composed of a compression encoder for modulation at the sending end, a decompression decoder for demodulation at the receiving end, a optimization goals which minimize the loss function, and a optimization algorithm for updating all trainable parameter.

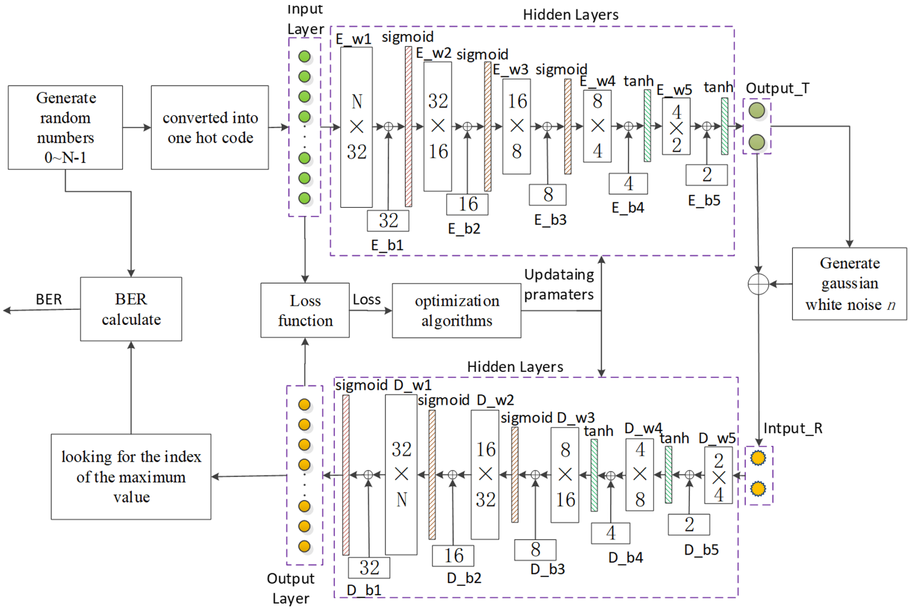

Fig. 1. Multi-modem model based on DAN.

2.1 Composition and Working Principle of the Sending End

At the sending end, random transmission symbols $s \in \{0, 1, 2, \cdots, N-1\}$ are generated according to N, which is the ary of the transmission symbol. Then, they are converted into a one hot code and input the compression encoder for modulation. The compression encoder consists of an N-node input layer (Input_Layer), 5 hidden layers (Hidden_Layers) and a 2-node output layer (Output_T). Among them, the five hidden layer weight parameters E_w1, E_w2, E_w3, E_w4 and E_w5 are $N \times 32$, 32×16, 16×8, 8×4 and 4×2 matrices respectively, and the offset parameters E_b1, E_b2, E_b3, E_b4 and E_b5 are 32, 16, 8, 4 and 2 column vectors respectively. The output of the Hidden_Layers 1, 2, and 3 uses the sigmoid activation function [14], as shown in (1), and the layers 4 and 5 use the tanh activation function [14], as shown in (2). After passing through Hidden_Layers, Output_T can be equivalent to the modulated I and Q signals of the sending end, as shown in (3), its value is limited to −1–1 by the tanh activation function.

$$\sigma_{sigmoid}(x) = \frac{x}{1+e^{-x}} \tag{1}$$

$$\sigma_{\tanh}(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{2}$$

$$\begin{aligned} Output\_T = \sigma_{\tanh}(\sigma_{\tanh}(\sigma_{sigmoid}(\sigma_{sigmoid}(\sigma_{sigmoid}(Input\_Layer \\ \times E\_w1 + E\_b1) \times E\_w2 + E\_b2)\, E\_w3 + E\_b3) \\ \times E\_w4 + E\_b4) \times E\_w5 + E\_b5) \end{aligned} \tag{3}$$

2.2 Composition and Working Principle of the Receiving End

After the Output_T is transmitted through the white noise channel, assuming that the receiving end has completed channel estimation such as carrier synchronization and bit synchronization, the signal entering decompression decoder for demodulation at the receiving end can be expressed as $Input\_R = Output\_T + n$, where $n \sim CN\left(0, \sqrt{\frac{P_s}{SNR}}\right)$ is the zero mean and $\sqrt{\frac{P_s}{SNR}}$ variance complex gaussian white noise, P_s is the signal power and the SNR is a specified signal-to-noise ratio. The structure and parameter settings of the decompression decoder for demodulation are completely symmetrical with the compression encoder for modulation. The N classification output of the N node output layer (Output_Layer) can be expressed as (4), and the sigmoid activation function limits the output value of each node between 0–1.

$$\begin{aligned} Output\_Layer = \sigma_{sigmoid}(\sigma_{sigmoid}(\sigma_{sigmoid}(\sigma_{\tanh}(\sigma_{\tanh}(Input\_R \\ \times D\_w5 + D\_b5) \times D\_w4 + D\_b4)\, D\_w3 + D\_b3) \\ \times D\_w2 + D\_b2) \times D\_w1 + D\_b1) \end{aligned} \tag{4}$$

The estimate of transmitted symbols $\hat{s}$ are obtained by looking for the index of the maximum value in the N nodes of $Output\_Layer$, and the bit error rate (BER) can be obtained by comparing $\hat{s}$ with transmit symbols s.

2.3 Optimization Objective and Optimization Algorithm

The optimization objective is to minimize the mean square error (MSE) loss function between the Input_Layer and the Output_Layer of the 10-layer DAN, as shown in (5). Where π is the parameter set that contains all the weight parameters and bias parameters to be optimized, and B is the number of symbols for each training in the mini-batch training process.

$$\min_{\pi}(loss) = \frac{1}{B}\min_{\pi}\left(\sum_{n=1}^{B}\sum_{i=1}^{N}(Output\_Layer[i] - Input\_Layer[i])^{2}\right) \tag{5}$$

AdaDelta algorithm [15] is chosen as the optimization algorithm, whose parameter update formula is shown in (6)

$$\pi_{t+1} = \pi_t - \frac{\eta}{\sqrt{E(g^2)_t + \varepsilon}} \times g_t \tag{6}$$

Where η is the initial learning rate and g_t represents the gradient at the t th iteration, ε is a minimum value added to prevent the denominator from being 0, $E(\cdot)$ represents expectation operation, and $E\big(g^2\big)_t = \rho E\big(g^2\big)_{t-1} - (1-\rho)\, g_t^2$ is the weighted average of the historical gradient squared and the current gradient squared, ρ represents the attenuation coefficient, and its value range is 0–1.

3 Training and Test Verification of the Multi-modem Based on DAN

3.1 Training of the Multi-modem Based on DAN

All parameters to be trained are initialized to small random numbers. The training set includes 4000 random integers ranging from 0 to N−1. After many experiments, it has been verified that 7 dB SNR and 10,000 rounds of training, 12 dB SNR and 20,000 rounds of training, and 13 dB SNR and 30,000 rounds of training are suitable for 16-ary and its below, 32-ary, and 64-ary modem respectively.

Using small batch training method [16], each batch is trained with 1000 symbols, and the white noise added into Output_T is regenerated during each batch training process. This method of generating white noise can be used as a Tikhonov regularization [17], which effectively prevents the 10-layer DAN from overfitting. The initial learning rate is set to $\eta = 0.2$, the attenuation coefficient is set to $\rho = 0.95$, and the minimum constant to prevent the denominator from being 0 when updating gradient is set to $\varepsilon = 1 \times 10^{-8}$. The test set contains regenerated 100,000-symbol random numbers. During the process of training, the MSE loss is outputted and the BER is calculated every each 200 rounds of training. simultaneously, the generalization ability of the 10-layer DAN is also detected.

The 10-layer DAN for Multi-modem uses TensorFlow [18] and single CPU for training. It takes about 5–8 totally for 200 rounds of training, a MSE loss calculating, a performance testing of 0–14 dB SNR, and an outputting of observed variables and graphics, which is increasing gradually from 2-ary to 128-ary. Therefore, the training for a certain modem can be completed within a few minutes.

Every set of parameters, which are trained and verified against one certain modem, can be stored, and then loaded according to the recognition and decision results of the channel environment in practice use. Transfer learning can be carried out against a specified layer or newly added layers, so that the trained Multi-modem based on DAN can quickly adapt to the new channel environment, and even be able to suppress malicious interference. These contents will be studied in depth in the follow-up works.

In addition, we also tried to use the minimized cross-entropy loss function as the optimization target to train the 10-layer DAN. When N is small, the 10-layer DAN can quickly converge, and the performance is consistent with that of the traditional modem, but when N is large, The 10-layer DAN cannot converge. For example, when $N = 16$, only 9 or 10 constellation points are formed with the cross-entropy loss function. When the Gradient Descent optimization algorithm is used, the loss value is often invalid, which result in the all parameters invalid after updating.

3.2 Training and Test Results of Multi-modem Model Based on DAN

The training and test results based on the above 10-layer DAN for 2-ary, 4-ary, 8-ary, 16-ary, 32-ary, 64-ary, and 128-ary are shown in Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6, Fig. 7, and Fig. 8 respectively. Among each figure, a constellation diagram for the modulation under training SNR is shown in subfigure (a). Relationship between the MSE loss and the number of training rounds, which is output once every 200 training rounds, is shown in subfigure (b). Training symbol modulation coding constellation without noise is shown in subfigure (c). And the performance comparison between modem based the 10-layer DAN and traditional method is displayed in subfigure (d), where, "DL (N = X)" represents the performance of X-ary modem based on the 10-layer DAN under white noise channel, and "theory of XXXX" represents the theoretical performance of XXXX modem under white noise channel.

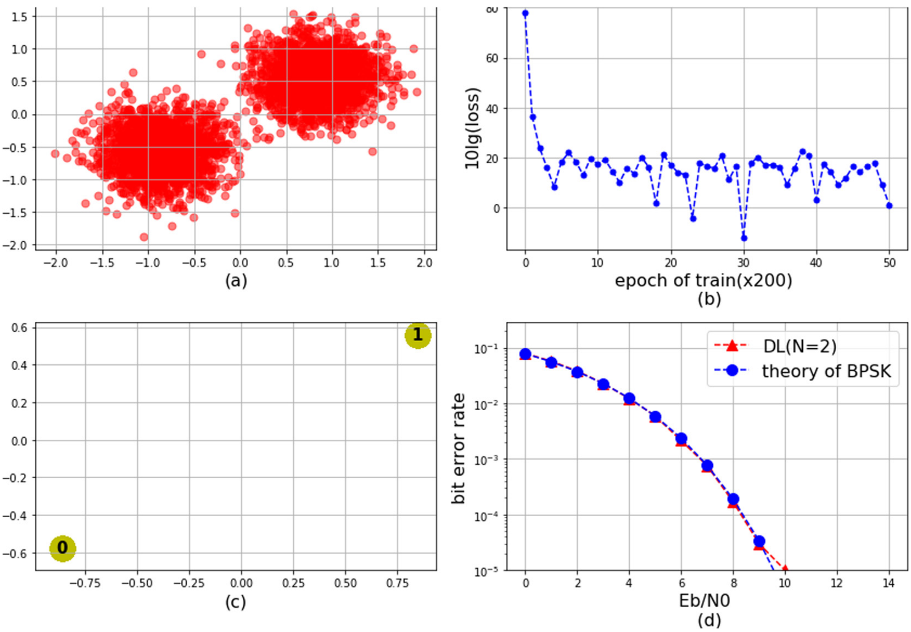

Fig. 2. 2-ary modem training results and performance comparison with BPSK.

Figure 2 and Figure 3 show that the constellation of 2-ary and 4-ary modulation based on the 10-layer DAN are consistent with traditional BPSK and QPSK

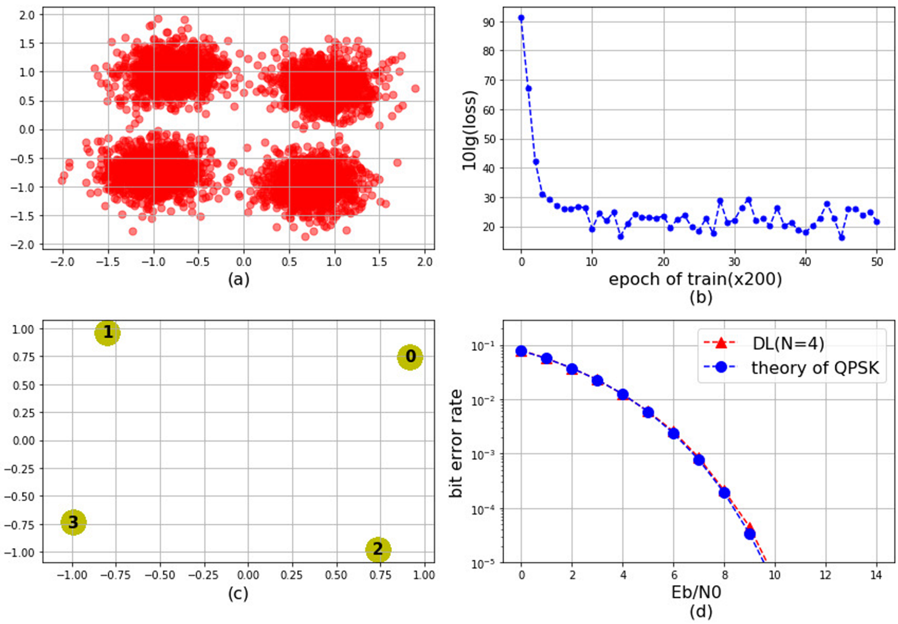

Fig. 3. 4-ary modem training results and performance comparison with QPSK.

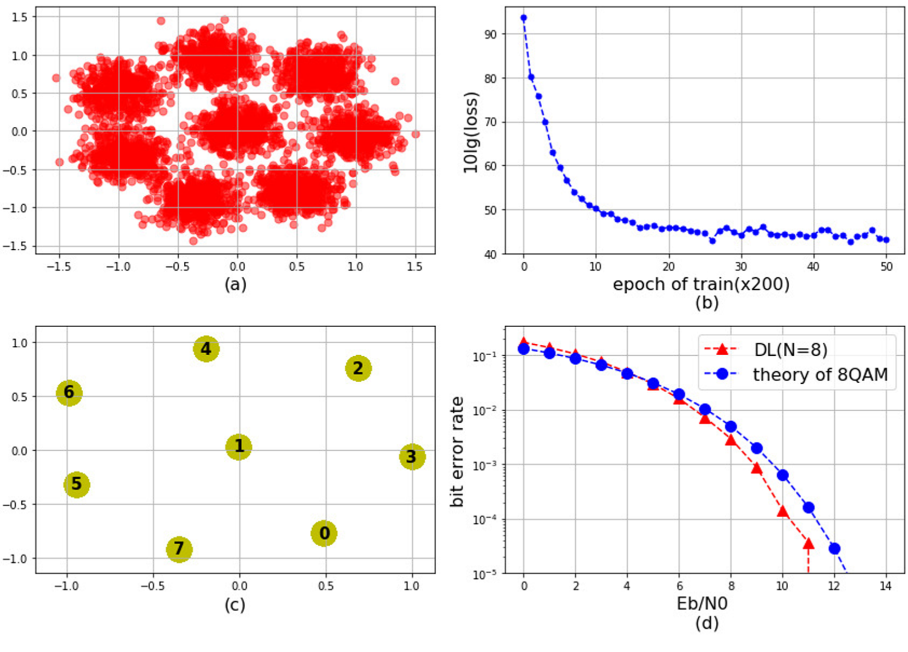

Fig. 4. 8-ary modem training results and performance comparison with 8QAM.

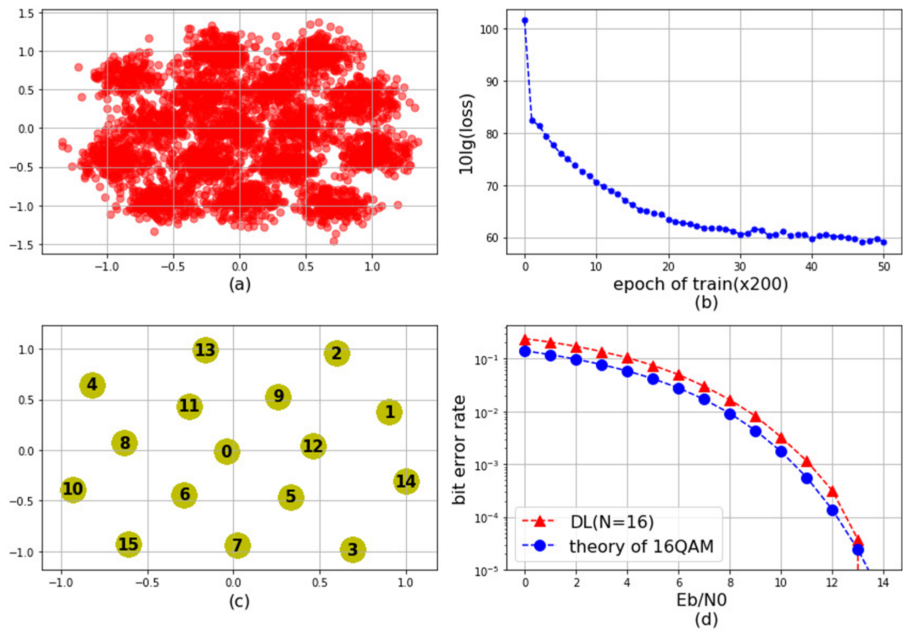

Fig. 5. 16-ary modem training results and performance comparison with 16QAM.

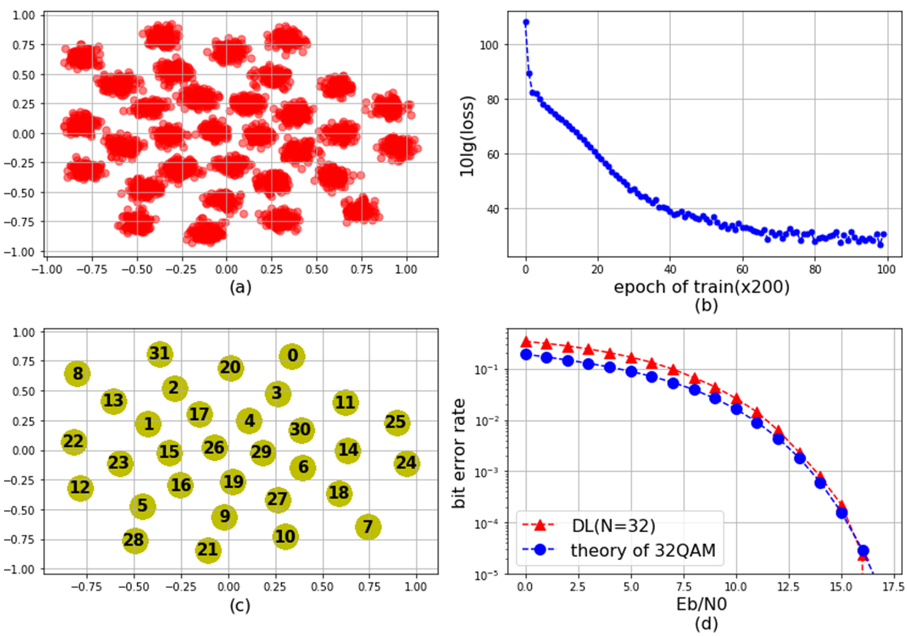

Fig. 6. 32-ary modem training results and performance comparison with 32QAM.

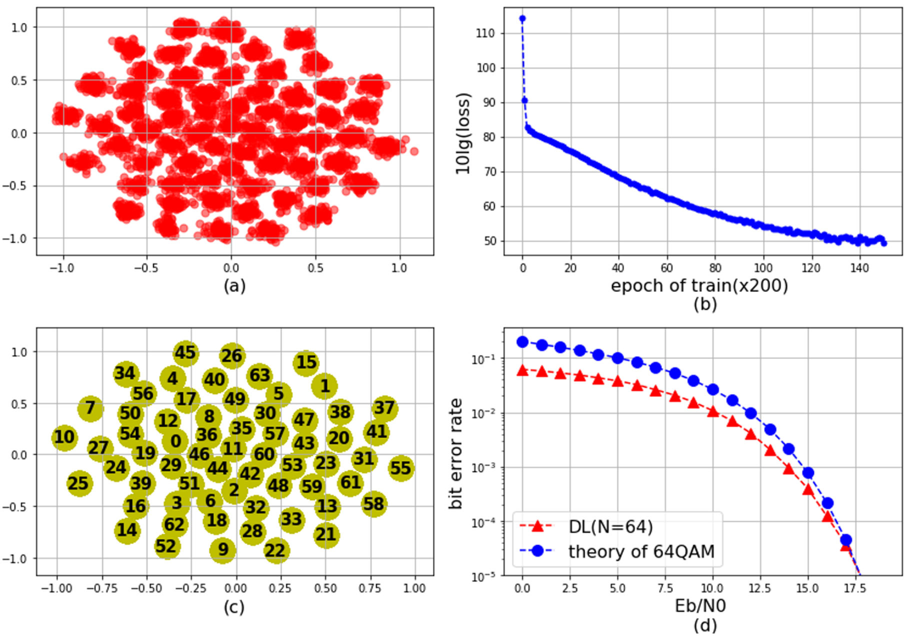

Fig. 7. 64-ary modem training results and performance comparison with 64QAM.

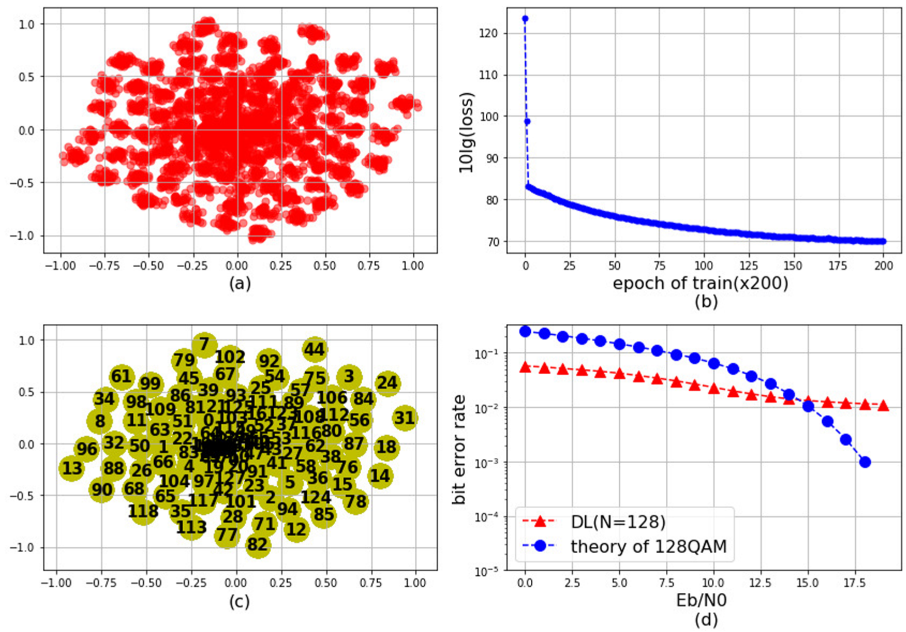

Fig. 8. 128-ary modem training results and performance comparison with 128QAM.

modulation respectively, and their demodulation performance is also completely consistent with the theoretical performance of BPSK and QPSK respectively. It should be noted that the constellation phase rotation and bit coding generated in each training are random. The coding result generated in Fig. 3(c) is consistent with the Gray code, so the demodulation performance is completely consistent with the QPSK theoretical performance. When the encoding result is not Gray code, the performance will be slightly reduced.

Figure 4 shows that the constellation of 8-ary base on the 10-layer DAN is different from the traditional 8PSK or 8QAM. A constellation point is placed at (0, 0), which results in an increase in the code interval. Therefore, the performance shows some improved compared with 8QAM.

For the training of 32-ary and 64-ary modem, the training SNR is increased to 13 dB, and the number of training rounds is increased to 20,000 and 30,000 respectively. The training results in Fig. 5, Fig. 6, and Fig. 7 show that the constellation diagrams generated by the 32-ary and 64-ary based on the 10-layer DAN are quite different from the traditional 16QAM, 32QAM, and 64QAM, and they are difficult to be distinguished linearly. Therefor they are difficult to demodulate by traditional methods and increase the signal's ability to resist interception. In terms of performance, the modem based on the 10-layer DAN for 16-ary and 32-ary is close to but slightly worse than 16QAM and 32QAM respectively (when the BER is lower than 10^{-3}, the performance degradation is less than 0.5 dB and 0.2 dB respectively). For 64-ary, the modem based on the 10-layer DAN is slightly better than that of 64QAM.

For higher-order 128-ary modem, after several attempts with the 10-layer DAN of Fig. 1, combing with different training number of symbols per batch, different training SNR, different training rounds, etc., no distinguishable symbol constellation can be formed. The typical training results are shown in Fig. 8. This result shows that the 10-layer DAN shown in Fig. 1 can only complete modem of 64-ary and below. For modem of 128-ary and above, a deeper network structure and more optimization parameters are required.

Extend the 10-layer DAN in Fig. 1 to 12 layers, that is, a new hidden layer is added after the Input_Layer of the sending end and before the Output_Layer of receiving end respectively. The weight parameter and offset parameter of the new hidden layer are a N64 matrix and a 64 column vector respectively. Sigmoid activation function is adopted.E_w1 and D_w1 of the original 10-layer self-coding network are set as a 6432 matrix respectively, E_b1 and D_b1 are set as a 64 column vectors respectively, other layer parameters remain unchanged.

The training method, loss function, training set and test set are still unchanged, and the training SNR is set as 18dB. As shown in Fig. 9, the 12-layer DAN basically converges after 40,000 rounds of training, and its performance is better than 128QAM demodulation. This results show that the Multi-modem model based on deep autoencoder network designed in this paper has flexible expansion ability and can adapt to higher order modem by appropriately increasing network depth and optimization parameters.

In order to show the training parameters and performance of the 10-layer DAN for Multi-Modem, a summary is presented in Table 1.

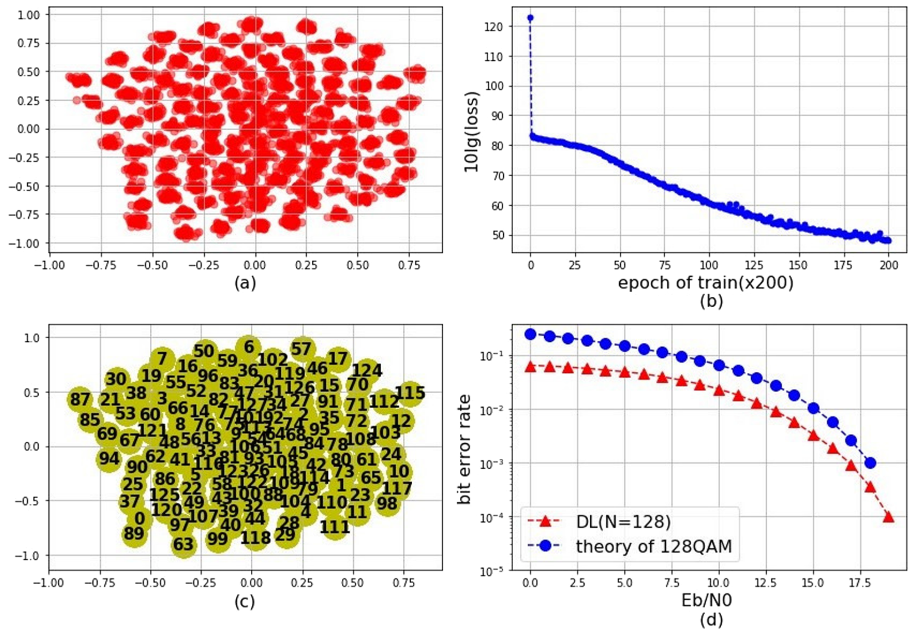

Fig. 9. 128-ary modem training results and performance comparison with 128QAM with 12-layer DAN.

Table 1. Training parameters and performance of the 10-layer DAN for Multi-modem.

10-layer DAN	Training SNR	Training epoch	Performance compared with traditional modulation	Constellation features
2-ary modem	7 dB	10,000	Consistent with BPSK	Consistent with BPSK
4-ary modem	7 dB	10,000	Consistent with QPSK	Consistent with QPSK
8-ary modem	7 dB	10,000	Better than 8QAM	Different from 8QAM, but easy to be differentiated linearly
16-ary modem	13 dB	10,000	Close to 16QAM	Different from 16QAM, and difficult to be differentiated linearly
32-ary modem	13 dB	20,000	Close to 32QAM	Different from 32QAM, and difficult to be differentiated linearly
64-ary modem	13 dB	30,000	Better than 64QAM	Different from 64QAM, and difficult to be differentiated linearly
128-ary modem	18 dB	40,000	Failure with 10-layer DAN, but better than 128QAM with 12-layer DAN	Different from 128QAM, and difficult to be differentiated linearly

4 Conclusion

In this paper, a 10-layer DAN is designed to realize modems from 2-ary to 64-ary with a unified network architecture, a unified loss function and a unified optimization algorithm, and the demodulation performance, that is close to, consistent with or better than that of traditional MPSK or QAM are obtained, which provides a new way for intelligent generation and reconstruction of satellite communication waveform and improves the anti-interception capability simultaneously. In addition, the 10-layer DAN network has the ability to be flexibly extended to accommodate higher order modems. Our further tests show that the Multi-modem model based on DAN can adapt to multipath environment and effectively suppress narrow-band interference by adding a convolutional layer at the receiving end. Further research directions include overall learning and training of coding and modulation based on deep learning network, generation of modem with specified frequency domain constraints (such as continuous phase modem), signal synchronization at the receiving end (carrier synchronization, symbol synchronization, etc.) and intelligent interference detection and suppression, etc.

References

1. Mao, Q., Fei, H., Hao, Q.: Deep learning for intelligent wireless networks: a comprehensive survey. IEEE Commun. Surv. Tutor. **20**(4), 2595–2621 (2018)
2. Zhang, J., Sheng, C., Maunder, R.G., Rong, Z., Hanzo, L.: Adaptive coding and modulation for large-scale antenna array-based aeronautical communications in the presence of co-channel interference. IEEE Trans. Wirel. Commun. **17**(2), 1343–1357 (2018)
3. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006)
4. O'Shea, T.J., Corgan, J., Clancy, T.C.: Unsupervised representation learning of structured radio communication signals. In: First International Workshop on Sensing (2016)
5. O'Shea, T.J., Karra, K., Charles Clancy, T.: Learning to communicate: channel auto-encoders, domain specific regularizers, and attention (2016)
6. O'Shea, T.J., Roy, T., West, N., Hilburn, B.C.: Physical layer communications system design over-the-air using adversarial networks (2018)
7. Oshea, T., Hoydis, J.: An introduction to deep learning for the physical layer. IEEE Trans. Cogn. Commun. Netw. **3**(4), 563–575 (2017)
8. Ye, H., Li, G.Y., Juang, B.-H.F., Sivanesan, K.: Channel agnostic end-to-end learning based communication systems with conditional GAN (2018)
9. Dorner, S., Cammerer, S., Hoydis, J., ten Brink, S.: Deep learning based communication over the air. IEEE J. Sel. Top. Sign. Process. **12**(1), 132–143 (2018)
10. Yaodong, Y., Le'nan, W.: Realization of SSAE demodulator of MPPSK communication system. J. Electro. Meas. Instrum. **32**(4), 144–150 (2018)
11. Huang, Y., Zhang, J., Zhou, X., Lu, J.: Demodulation with deep learning. Telecommun. Eng. **57**(7), 741–744 (2017)
12. Zhang, J., Jin, S., Wen, C., Gao, F., Jiang, T.: An overview of wireless transmission technology utilizing artificial intelligence. Telecommun. Sci. **34**(8), 46–55 (2018)

13. Gui, G., Wang, Y., Huang, H.: Deep learning based physical layer wireless communication techniques: opportunities and challenges. J. Commun. **4**(2), 19–23 (2019)
14. Gomar, S., Mirhassani, M., Ahmadi, M.: Precise digital implementations of hyperbolic tanh and sigmoid function. In: Asilomar Conference on Signals (2016)
15. Zeiler, M.D.: Adadelta: an adaptive learning rate method. Computer Science (2012)
16. Masters, D., Luschi, C.: Revisiting small batch training for deep neural networks (2018)
17. Bishop, C.: Training with noise is equivalent to Tikhonov regularization. Neural Comput. **7**(1), 108–116 (1995)
18. Haque, Z., Cheng, H.-T., Hong, L.: Tensorflow estimators. In: the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1763–1771 (2017)