GAN-SNR-Shrinkage-Based Network for Modulation Recognition with Small Training Sample Size

Shuai Zhang, Yan Zhang, Mingjun Ma, Zunwen He(✉), and Wancheng Zhang

Beijing Institute of Technology, BIT, Beijing, People's Republic of China
hezunwen@bit.edu.cn

Abstract. Modulation recognition plays an important role in non-cooperative communications. In practice, only a small number of samples can be collected for training purposes. The limited training data degrade the accuracy of the modulation recognition networks. In this paper, we propose a novel network to realize the modulation recognition on basis of the few-shot learning. Generative adversarial networks (GANs) and a signal-to-noise ratio (SNR) augment module are introduced to expand the training dataset. In addition, a preprocessing module and residual shrinkage networks are used to improve the capability of characterizing signal features and the anti-noise performance. The proposed network is evaluated using the RML2016.10a dataset. It is illustrated that the proposed network outperforms the baseline method and the method without data augment with a small number of training samples.

Keywords: Modulation recognition · GAN · SNR · Few-shot learning

1 Introduction

Non-cooperative communication plays an important role in radio resource management, communication spectrum monitoring, and other civilian and military fields. Modulation recognition, which is between signal detection and demodulation, has important research value in the field of non-cooperative communication. It is also of great significance to signal demodulation, information extraction, and signal detection in communication systems.

Traditional modulation recognition techniques can be divided into two categories: likelihood-based methods and signal-characteristic-based methods. Likelihood-based methods [1–3] have good performance in recognition accuracy. The main ideas of these methods are to realize the recognition of signals according to the modulation mode of signals with the help of probability theory and Bayesian theory. However, they suffer from the problem of computational complexity, i.e., the technical calculation complexity is very high in the recognition process. The signal-characteristic-based methods which are mainly based on manually extracted characteristics of signal processing are generally more widely employed. In these methods, signal features, such as instantaneous feature

H. Gao et al. (Eds.): ChinaCom 2021, LNICST 433, pp. 80–90, 2022.
https://doi.org/10.1007/978-3-030-99200-2_7

quantity [4], spectral correlation [5], wavelet transform [6], etc., are firstly designed and extracted, and then classification rules are designed for modulation recognition of signal features. However, the recognition accuracy of manually designed signal features and classification rules is usually limited in complex channel environments.

In recent years, modulation recognition techniques based on neural networks and deep learning have been proposed to automatically learn classification rules from feature data to improve recognition accuracy. In the field of modulation recognition, feature transformation based on Convolutional Neural Networks, Recurrent Neural Networks, or Convolutional, Long-Short-Term-Memory, Deep-Neural-Network is proposed in [7–9], and they have achieved good recognition rates. Although the deep-learning-based methods outperform the traditional manual extracted-based methods in recognition performance, they require a large number of labeled samples for training. If the number of labeled samples is insufficient, the recognition performance of the deep-learning-based methods will degrade sharply.

Therefore, few-shot learning is introduced to solve the modulation recognition problem with few training samples. Capsule networks [10], modulated filters, a manual-feature-based method [11], and a data preprocessing method [12] are proposed in order to improve recognition rates. Various networks based on the GAN have been applied in [13] to generate the samples for training purposes. However, these methods assume that training data contains all SNR cases which is difficult to satisfy in practice.

In this paper, we propose a GSS (GAN-SNR-Shrinkage) network structure for modulation recognition based on a small sample set. An SNR augment module and a deep convolutional GAN (DCGAN) module are introduced to this network. According to the characteristics of the modulated signal, we extend the data according to the SNR by the designed SNR augment module. This module does not need complicated operations and ensures the feasibility of extended data.

Then, we preprocess the signal by shifting and splicing in the time-domain, which increases the dimensions of data to be beneficial for the recognition. Finally, we train an effective classifier that is specially designed including residual shrinkage block to realize modulation recognition of the input data.

We verify our network performance on dataset RML2016.10a [14]. Multiple cross-experiments and comparison of results show that the proposed network has higher accuracy than the baseline method and the method without data augment under the condition of small sample size.

The rest of this paper is organized as follows: Sect. 2 and Sect. 3 introduce the system model and the structure of the network. In Sect. 4, we present the test results on the RML2016.10a and analyze the advantages of our network. Section 5 presents the conclusion.

2 System Model

In the general communication channel model with the additive white Gaussian noise (AWGN), signals of different modulation modes are expressed as

$$s(t) = h(t) * f(t) + n(t) \tag{1}$$

where is the modulated signal, which may be analog signal or digital signal. is the channel impulse response and represents the convolution operation. is the additive white Gaussian noise with a zero mean and a variance of . is the received signal.

To train an effective recognition network, received signals under different SNR conditions should be collected. The collected dataset can be expressed as

$$\{s_{original}\} = \bigcup_{i=1,2,\ldots,n} \{s_{ri}\} \tag{2}$$

where is the dataset with SNR and are the different SNR values.

3 GSS Network Structure

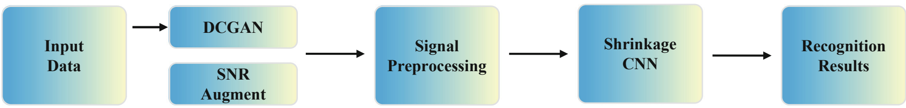

Fig. 1. Proposed GSS network structure.

The proposed network structure of the system is shown in Fig. 1 is the input of the network including different SNRs. DCGAN is set up as the first data enhancement module, which is used to generate some pseudo samples with an input. These pseudo samples are directly generated on the I/Q dimension to expand the number of training samples.

Then, the original signal is also processed through the SNR augment module to generate new samples, which is able to transform the data of high SNRs into the data of low SNRs by adding artificial Gaussian noise, and to increase the samples quantity. Finally, the entire training samples include.

$$\{s_{train}\} = \{f_{DCGAN}(z)\} \bigcup \{f_{SNR}(s_{original})\} \bigcup \{s_{original}\}. \tag{3}$$

Moreover, a signal preprocessing module [12] is applied in the third part of the GSS network structure. On the one hand, the I/Q expression is transferred into the amplitude-phase (AP) expression. On the other hand, signal samples are regularly shifted and spliced to improve the feature expression ability of the samples.

The last module is the specially designed CNN module with residual shrinkage block. Residual shrinkage block has shown a good effect in mechanical fault signal diagnosis [14]. Therefore, we introduce it into our network structure to replace the original CNN for the modulation recognition.

3.1 DCGAN

DCGAN model mainly consists of a generating network G and a discriminant network D. G is responsible for generating samples. G accepts a random noise z, and the generated samples are denoted as G(z). D is responsible for judging whether the sample is real or generated, represented by 1 and 0, respectively. The process of model training constitutes a dynamic game between G and D, and the cross-entropy loss function [15] in the training process is constructed as

$$\min_G \max_D V(D, G) = E_{x\sim pdata(x)}[\ln D(x)] + E_{x\sim P_z(z)}[\ln(1 - D(G(z)))] \tag{4}$$

where is the real data used for training, and the distribution of data is. represents the distribution of noise, and G(z) represents the samples generated by G.

To generator G, the optimization problem is

$$\min_G V(D, G) = E_{x\sim P_z(z)}[\ln(1 - D(G(z)))]. \tag{5}$$

On the other hand, D wants to have a strong discriminating ability, the mixed loss function of D can be depicted as

$$\max_D V(D, G) = E_{x\sim pdata(x)}[\ln D(x)] + E_{x\sim P_z(z)}[\ln(1 - D(G(z)))]. \tag{6}$$

Then, the discriminator D will update its parameters by descending its gradient.

$$\nabla_{\theta_d} \frac{1}{m} \sum_{l=1}^{m} [\ln D(x^{(l)}) + \ln(1 - D(G(z^{(l)})))] \tag{7}$$

where is training data batch size, is the noise vector in the minibatch.

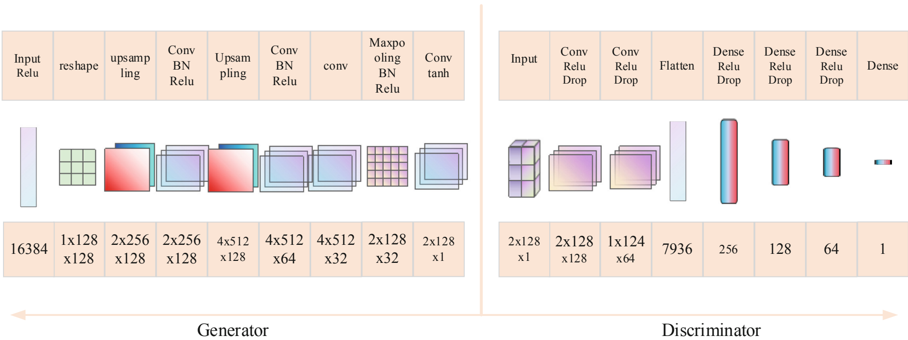

Fig. 2. Network structure parameters of DCGAN.

The network structure and parameter settings of G and D are shown in Fig. 2. In DCGAN, the input of G is random noise, and its output is a vector with dimensions 2 × 128 × 1. Meanwhile, the input of D is a vector with dimensions 2 × 128 × 1, and its output is the discriminant result (1 represents sample is real, in contrast, 0 means sample is generated). The training of DCGAN is realized by alternate optimization of D and G. DCGAN can generate any amounts of samples at different SNRs.

3.2 SNR Augment

The input of the SNR augment module is the data under the existing SNRs. These samples are enhanced to produce more samples with a lower SNR, and the amounts of generated samples present a stepwise type. The generated samples are represented as

$$\{s_{SNR}\} = \bigcup_{n=r_1,r_2,\ldots,r_n} \bigcup_{i\in\{r_1,\ldots,n\}} \{s_i\}. \tag{8}$$

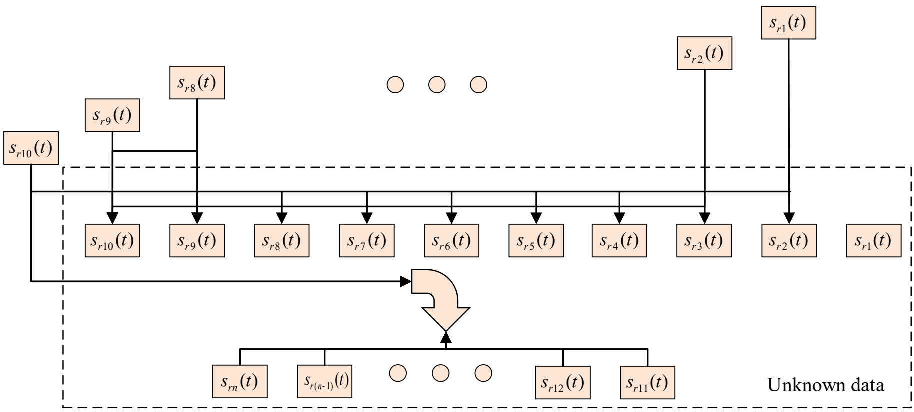

Fig. 3. Description of the process of SNR augment.

As shown in Fig. 3, each sample with a high SNR can produce different samples with low SNR. For the existing SNR signal, the energy and SNR of signal and noise are defined below

$$\frac{S_0}{\sigma^2} = \frac{The\ power\ of\ signa}{The\ power\ of\ noise} \tag{9}$$

$$\sigma^2 = S_0 * 10^{(-\frac{SNR}{10})} \tag{10}$$

where is the power of the signal, is the variance of Gaussian noise which is equal to its power. Therefore, the difference in the signal powers between low SNR and high SNR is

$$\left|\sigma_{high}^2 - \sigma_{low}^2\right| = \left|S_0 * (10^{(-\frac{SNR_{high}}{10})} - 10^{(-\frac{SNR_{low}}{10})})\right|. \tag{11}$$

Through the stepwise data enhancement method of the SNR augment, low -SNR samples can be obtained from high -SNR samples, and the reliability and validity of the samples can be guaranteed. When the SNR is below 0 dB, the signal itself is already greatly polluted by the noise. Thus, the signal samples whose SNRs are below 0 dB will not be used to carry out such corresponding data enhancement.

3.3 Signal Preprocessing

According to the characteristics of the signal, the signal preprocessing module is used before entering CNN to make contributions for recognition, and the specific processing process is shown in Fig. 4.

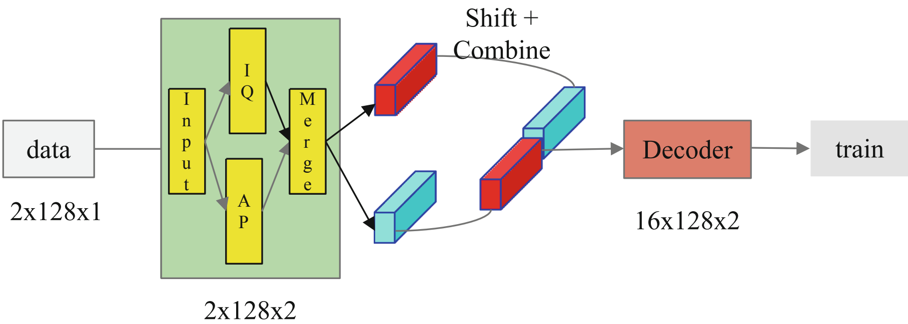

Fig. 4. Structure of signal preprocessing.

In Fig. 4, the amplitude-phase format of the data is.

$$\begin{cases} A = \sqrt{I^2 + Q^2} \\ P = \arctan(\dfrac{Q}{I}) \end{cases} \tag{12}$$

The dimension of the input data is expanded to.

$$s_l = [I_l, Q_l]^T \bigcup [A_l, P_l]^T = \left[\begin{bmatrix} I_l \\ Q_l \end{bmatrix}, \begin{bmatrix} A_l \\ P_l \end{bmatrix} \right] \tag{13}$$

In this part, the input I/Q is originally one channel, and the data are converted to amplitude-phase data and then regrouped into two channels. In addition, the sample size has been expanded through cross-shift combinations, and the combined samples are used as training data.

3.4 Shrinkage CNN

Samples expanded by the previous module will be input into the CNN for classification. The convolution kernels in CNN have good feature extraction capability and fewer parameters. BN (Batch Normalization) is set to reduce the internal covariant shift and prevent gradient disappearance during the training process. Residual shrinkage blocks are inserted into the designed CNN. The structure of Specially designed CNN is shown in Fig. 5.

The function of the first two different convolution layers is to extract horizontal and vertical features, and the purpose of BN is to reduce the internal covariant shift and

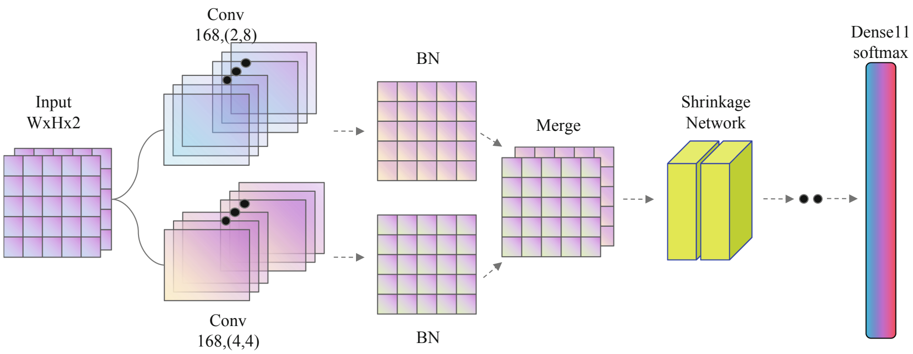

Fig. 5. Specially designed Shrinkage CNN Network.

prevent gradient disappearance during the training process. Identity shortcut is used in the residual shrinkage block, which is proved superior to the traditional CNN in [16]. To effectively eliminate the features related to noise, the shrinkage block inserts soft threshold [17] into the structure as a nonlinear transformation layer as.

$$y = \begin{cases} x - \tau, x > \tau \\ 0, -\tau \leq x \leq \tau \\ x + \tau, x < -\tau \end{cases} \tag{14}$$

where is the input vector of the soft threshold module, is the output vector, is the threshold.

In addition, specially designed sub-networks are used in the network to determine the threshold adaptively [18], so that each signal can have its own set of thresholds as.

$$\tau = \alpha * \underset{u,v,w}{average}\left|x_{u,v,w}\right| \tag{15}$$

where ,, and are indexes of width, height, and channel of the feature map, respectively.

The discriminator outputs a dimensional vector of, which is turned into by the softmax function. is the number of the signal modulation types. The expression of is

$$p_j = \frac{e^{n_j}}{\sum_{q=1}^{k+1} e^{n_q}}, j \in \{1, 2, ...k\} \tag{16}$$

4 Experimental Results

4.1 Dataset

The RML2016.10a dataset [14] is widely used for model performance verification in modulation recognition. It contains 11 different modulation modes, i.e., BPSK, QPSK, 8PSK, QAM16, QAM64, CPFSK, GFSK, 4PAM, WBFM, AM-DSB, and AM-SSB. Both analog modulations and digital modulations are included. Each signal contains 2200 baseband I/Q data with 128 sampling points.

These samples are collected from simulated signals through wireless channels and are affected by multipath fading, sampling rate offset and center frequency offset. The SNR of the data ranges from −20 dB to 18 dB, with 2 dB intervals. We randomly select 20% samples as the test set, and the evaluation of the model is carried out under the same test set.

4.2 Results

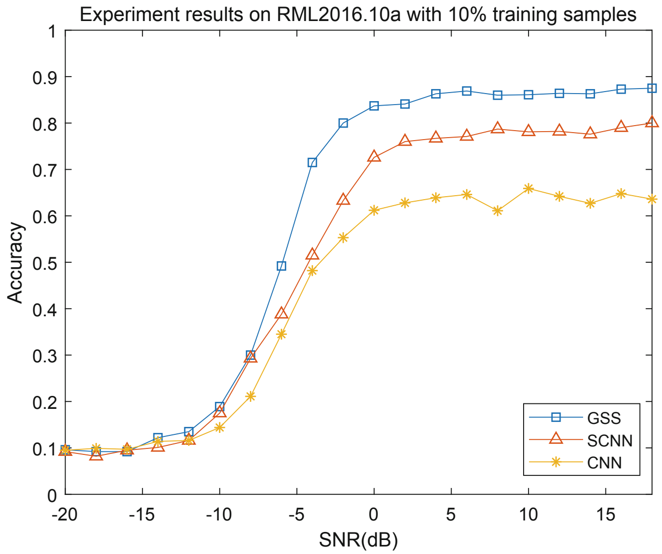

Fig. 6. Recognition accuracy of our proposed GSS network, SCNN, and CNN with 10% of the number of training samples.

Considering the few-shot learning requirement, only 10% of the number of samples are selected as the training sample. We verify the performance of the GSS network and compare it with a baseline CNN classifier and the specially designed CNN (SCNN) [12] which achieves high- accuracy in modulation recognition. Figure 6 shows the accuracy of different networks at varying SNR. Based on our experiment results, our proposed method can achieve the highest recognition accuracy of 88.1% at 6dB, which is 6.4% and 23.7% higher than the baseline CNN and the SCNN. In addition, the GSS network shows great advantages over the other two methods starting from −8 dB. The accuracy rate of our method has reached a high level when the SNR is above −4 dB.

Figure 7 shows the recognition accuracy of our proposed GSS network, SCNN, and CNN with 20% of the number of training samples. As a result of Fig. 6 and Fig. 7, interestingly, the three recognition rate curves do not rise all the time. It is noted that there are some slight fluctuations in the GSS curve when the SNR value is between 0 dB to 18 dB. The reason may come from two aspects. On the one hand, the number of training samples is small, so the fitting effect of the network on the test set at each SNR

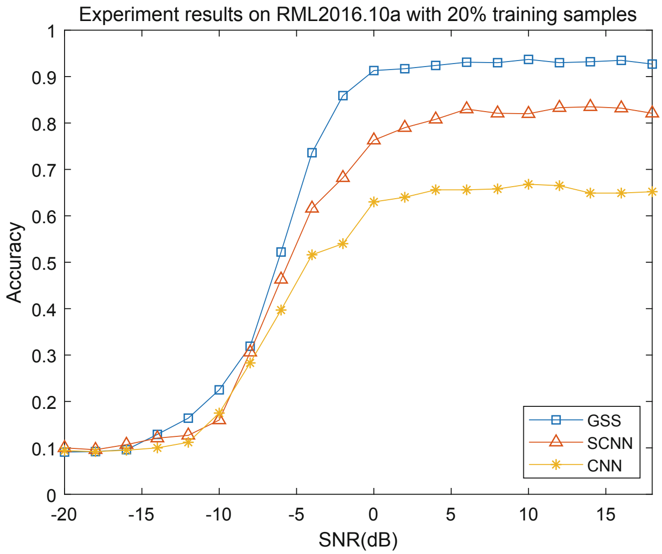

Fig. 7. Recognition accuracy of our proposed GSS network, SCNN, and CNN with 20% of the number of training samples.

may not be optimal. On the other hand, due to the stepwise characteristics of the SNR augment module, the number of training samples with high SNRs is less than that with low SNRs, which makes the training model works well with the SNR values around 6 dB.

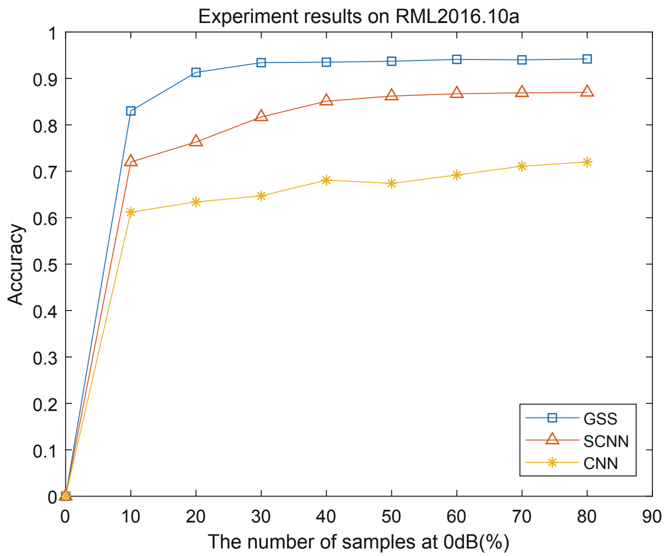

Fig. 8. Recognition accuracy with different numbers of training samples at 0 dB.

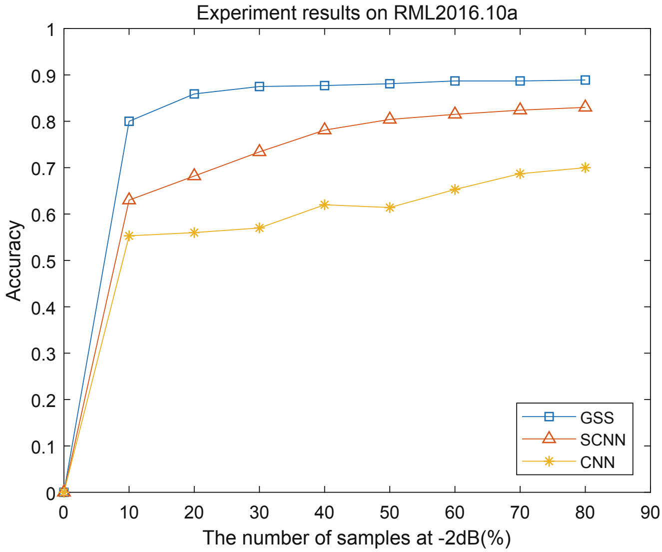

Fig. 9. Recognition accuracy with different numbers of training samples at −2 dB.

Then, we also do comparison experiments when SNR is fixed and the number of samples changes. Figure 8 and Fig. 9 illustrate the recognition accuracy of the above models at 0 dB and −2 dB. The GSS network achieves 91.3% accuracy with only 20% of the samples when the SNR is 0 dB, whereas the maximum accuracy rate of SCNN and CNN is 87.3% and 70.1%. Although the recognition rate of GSS at −2 dB is lower than that of 0 dB, it is still much higher than the other two methods. In addition, with the increase of the training sample size, the recognition accuracy of the GSS network increases faster than the other two methods.

5 Conclusion

In this paper, we have proposed the GSS network for automatic modulation recognition with a small training sample set. An SNR augment module, as well as DCGAN, have been introduced to realize the expansion of the training dataset. Specially designed CNN with the signal preprocessing module has been used to achieve recognition of signals. It has been proved that the proposed method has a high recognition rate when the training data are in a small size. Experiments on the RML2016.10a dataset have shown that the proposed GSS network can achieve a recognition rate of over 90% with only 20% training samples.

Acknowledgements. This work was supported by the National Key R&D Program of China under Grant (2019YFE0196400), the National Natural Science Foundation of China under Grant (61871035), and the National Defense Science and Technology Innovation Zone.

References

1. Dobre, O., Abdi, A., Bar-Ness, Y., Su, W.: Survey of automatic modulation classification techniques: classical approaches and new trends. IET Commun **1**(2), 137 (2007)
2. Hameed, F., Dobre, O.A., Popescu, D.C.: On the likelihood-based approach to modulation classification. IEEE Trans. Wirel. Commun. **8**(12), 5884–5892 (2009)
3. Chen, W., Xie, Z., Ma, L., Liu, J., Liang, X.: A faster maximum-likelihood modulation classification in flat fading non-Gaussian channels. IEEE Commun. Lett. **23**(3), 454–457 (2019)
4. Nandi, A.K., Azzouz, E.E.: Algorithms for automatic modulation recognition of communication signals. IEEE Trans. Commun. **46**(4), 431–436 (1998)
5. Ma, J., Qiu, T.: Automatic modulation classification using cyclic correntropy spectrum in impulsive noise. IEEE Wirel. Commun. Lett. **8**(2), 440–443 (2019)
6. Wu, Z., Zhou, S., Yin, Z., Ma, B., Yang, Z.: Robust automatic modulation classification under varying noise conditions. IEEE Access **5**, 19733–19741 (2017)
7. Li, R., Li, L., Yang, S., Li, S.: Robust automated VHF modulation recognition based on deep convolutional neural networks. IEEE Commun. Lett. **22**(5), 946–949 (2018)
8. Peng, S., et al.: Modulation classification based on signal constellation diagrams and deep learning. IEEE Trans. Neural Netw. Learn. Syst. **30**(3), 718–727 (2019)
9. Hong, D., Zhang, Z., Xu, X.: Automatic modulation classification using recurrent neural networks. In: 2017 3rd IEEE International Conference on Computer and Communications (ICCC), pp. 695–700 (2017)
10. Li, L., Huang, J., Cheng, Q., Meng, H., Han, Z.: Automatic modulation recognition: a few-shot learning method based on the capsule network. IEEE Wirel. Commun. Lett. **10**(3), 474–477 (2021)
11. Zhang, D., Ding, W., Liu, C., Wang, H., Zhang, B.: Modulated autocorrelation convolution networks for automatic modulation classification based on small sample set. IEEE Access **8**, 27097–27105 (2020)
12. Zhang, H., Huang, M., Yang, J., Sun, W.: A data preprocessing method for automatic modulation classification based on CNN. IEEE Commun. Lett. **25**(4), 1206–1210 (2021)
13. Gong, J., Xu, X., Lei, Y.: Unsupervised specific emitter identification method using radio-frequency fingerprint embedded InfoGAN. IEEE Trans. Inf. Forensics Secur. **15**, 2898–2913 (2020)
14. O'Shea, T.J., Corgan, J., Clancy, T.C.: Convolutional radio modulation recognition networks. In: Jayne, C., Iliadis, L. (eds.) EANN 2016. Communications in Computer and Information Science, vol. 629, pp. 213–226. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44188-7_16
15. De Boer, P.T., Kroese, D.P., Mannor, S., et al.: A tutorial on the cross-entropy method. Ann. Oper. Res. **134**(1), 19–67 (2005)
16. Zhang, Z., Li, H., Chen, L.: Deep residual shrinkage networks with self-adaptive slope thresholding for fault diagnosis. In: 2021 7th International Conference on Condition Monitoring of Machinery in Non-Stationary Operations (CMMNO) (2021). In press
17. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 27–30 June 2016, pp. 770–778 (2016)
18. He, K., Zhang, X., Ren, S., Sun, J.: Identity mappings in deep residual networks. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9908, pp. 630–645. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46493-0_38