A Deep Learning Method Based on Convolutional Neural Network for Automatic Modulation Classification of Wireless Signals

Yu Xu[1], Dezhi Li[1], Zhenyong Wang[1,2(✉)], Gongliang Liu[1], and Haibo Lv[1]

[1] School of Electronics and Information Engineering,
Harbin Institute of Technology, Harbin, Heilongjiang, China
{xu_yu,lidezhi,ZYWang,liugl,elitelv}@hit.edu.cn
[2] Shenzhen Academy of Aerospace Technology, Shenzhen, Guangdong, China

Abstract. Automatic modulation classification (AMC) plays an important role in many fields to identify the modulation type of wireless signals. In this paper, we introduce deep learning to signal recognition. Based on architecture analysis of the convolutional neural network (CNN), we used real signal data generated by instruments as dataset, and proposed an improved CNN architecture to achieve compatible recognition accuracy of modulation classification. According to various conditions of signal noise ratio (SNR), we test the proposed CNN architecture with the real sampled signals. Experiments results show that the high-layer network is not necessary for modulation recognition with high SNR signals. The proposed CNN architecture has higher average classification accuracy than RESNET and is more compatible for modulation classification of signals with lower SNR.

Keywords: Modulation classification · Deep learning
Convolutional neural network · Wireless signal

1 Introduction

Automatic modulation classification is aiming to detect the modulation type of received signals in order to recover signals by demodulation. The dominant approach of signal modulation recognition can be categorized as likelihood-based methods and feature-based methods [1]. Most likelihood-based classifiers require parameter estimation, while feature-based methods can be free from parameter estimation and achieve high popularity in recent years. Generally, feature-based methods consist of two steps: feature extraction and classifier, which can provide classification decisions according to some particular criterion.

Although the feature-based methods have shown great advantages in classification, two problems still remain: one is the difficulty of Manually Feature Extraction, the other is Noise Covering. Most conventional feature-based methods cannot utilize full feature information when the performance of feature-based methods relies on the

quality of the extracted features. Moreover, for modulation classification, manually feature extraction is complicated and difficult for comprehensive modulation types of wireless signals. For the Noise Covering problem, if the SNR is very low, the features we can extract are so limited that we can't get a satisfied performance for automatic modulation classification.

Deep learning is a fascinating field and has achieved a series of state-of-the-art results in different domains. However, Deep learning also has been tried for modulation classification in some related researches. In paper [2], a modulation classification method based on stacked de-noising sparse auto-encoder (SDAE) is investigated, which can extract modulation features automatically, and classify input signals based on the extracted features to get compatible results. Also based on deep learning algorithms, the stacked sparse auto-encoders to extract features from ambiguity function (AF) images of signals are proposed to discriminate digital modulated signals [3]. After that, the obtained features are fed back into a Softmax regression classifier in order to recognize 7 popular modulations including ASK, PSK, QAM, FSK, MSK, LFM and OFDM. In paper [4], deep belief network (DBN) is applied for pattern recognition and classification. Compared with those likelihood-based methods and feature-based methods, they all show great success of high recognition accuracy in various SNR conditions.

The main idea of this paper is to provide a stacked convolutional neural network of deep learning architecture for modulation classification based on extracted features of wireless signals automatically. The rest of the paper is organized as follows. In Sect. 2, principles of Convolutional Neural Network are investigated. Based on real sampled data of wireless signals, an improved CNN architecture is trained and proposed in Sect. 3. In Sect. 4, experiments are included to compare average classification accuracy with RESNET [5] under conditions of various SNR. Finally, conclusions are drawn in Sect. 5.

2   Principle of Convolutional Neural Network

The goal of a neural network is to approximate a function f^*. For a classifier, function $y = f^*(x)$ maps an input x to a category y. A neural network defines a mapping criterion $y = f(x, \theta)$ and obtains the value of the parameters θ that result in the optimal approximation function of the true mapping function.

Convolutional Neural Network is a powerful architecture of artificial neural network, which is popular because of state-of-the-art achievements in computer vision processing and natural language processing.

2.1   Architecture of Convolutional Neural Network

CNN process consists of two components: convolutional layers and pooling layers. Convolutional layers are comprised of filter kernel and feature maps. The filter kernels have weighted inputs and generate an output value like a neuron. The feature map is the output of one filter kernel applied to the previous layer. A given filter kernel is drawn across the entire previous layer and moved one point at a time, which depends on the

stride. Each position results in activation of the neuron and generates an output to form the feature map, as illustrated in Fig. 1.

The pooling layer down-samples the feature map of previous layers. Pooling layers follow a sequence of convolutional layers to consolidate the learned features in the previous feature map. Therefore, pooling may be considered as a technique to compress and generalize feature representations, so as to generally reduce the model overfitting phenomena. In Fig. 2, the max pooling process is illustrated with pool width of 3 and stride of 2.
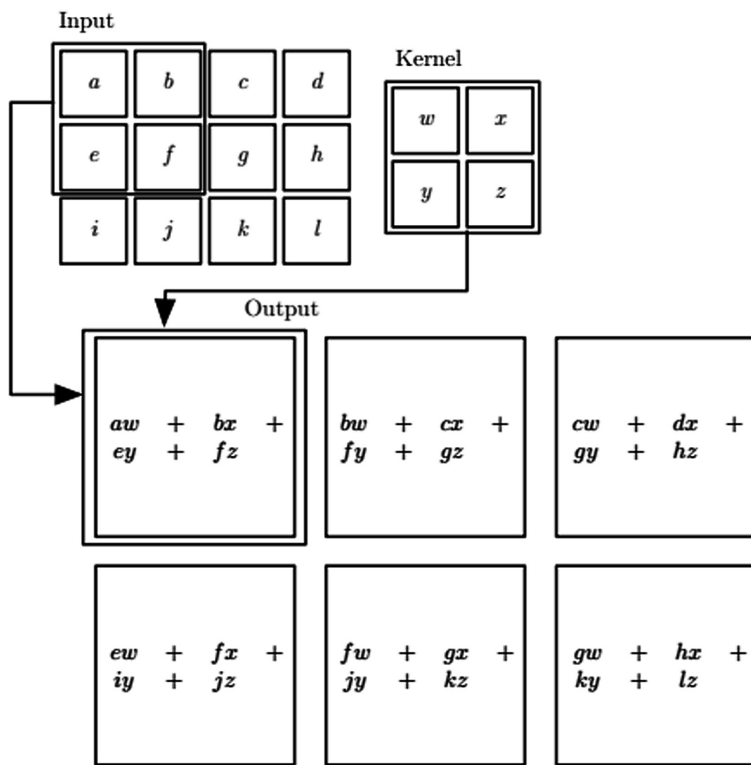

Fig. 1. 2-D convolution operation of CNN

2.2 Training Process Based on Stochastic Gradient Descent

Stochastic Gradient Descent (SGD) is typical and preferred training algorithm for neural networks. One row of data is inputted into the network at a time. The network activates neurons forward to produce an output value finally. Then the output value is compared to the expected output value to generate an error value. The error is backward propagated through the network, in which the weights of layer are updated one after another, according to the contributed amount to the error. The process is repeated for all of the examples in the training data to get a trained network of the intended goal.
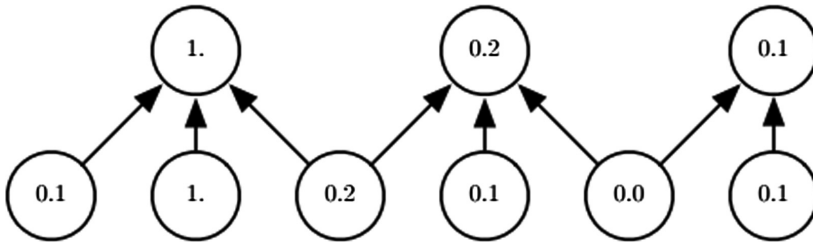
Fig. 2. Max pooling process of CNN

The weights in the network can be updated from the calculated errors for each training example, which can result in fast but also chaotic changes to the network. On the other hand, the errors can be saved up across all of the training examples and the network can be updated at the end.

For computational efficiencies, it is necessary to define the batch size of datasets. The batch size is often reduced to a small number of tens or hundreds of examples before updates. The amount that weights are updated is controlled by a configuration parameter called the learning rate, which controls the each step of network weights updating for a given error.

The cost function is important to design a deep neural network. In most cases, the parametric model defines a distribution $p(y|x; \theta)$ and the principle of maximum likelihood is used to train the model. The cost function is defined as the cross-entropy between the distribution of training data and the model predictions, as shown in Formula (1).

$$J(\theta) = -\log p_{\text{model}}(y|x) \tag{1}$$

The Softmax functions are often used as the output of a classifier, which represent the probability distribution over n different classes. The Softmax function is given by

$$soft\max(z)_i = \frac{\exp(z_i)}{\sum_j \exp(z_j)} \tag{2}$$

The exponential Softmax function works very well when the Softmax function is trained to output a target value y based on maximum log-likelihood. In this case, the Softmax function in this paper is defined in terms of exponential function:

$$\log soft\max(z)_i = z_i - \log \sum_j \exp(z_j) \tag{3}$$

When the training process maximizes the log-likelihood, the first term z_i is encouraged to be increased, while the second term is punished to be decreased. The negative log-likelihood cost function always strongly penalizes the most inactive prediction.

3   The Improved CNN for Modulation Classification

To meet the requirements of modulation classification, our network architectures are mainly inspired by ALEXNET [6], as shown in Fig. 3.
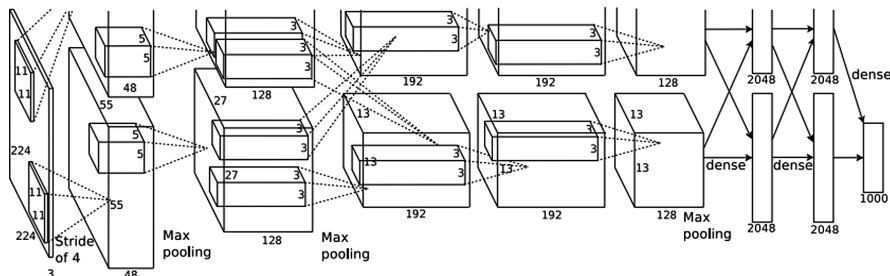

Fig. 3. The architecture of ALEXNET

3.1   Signals Data Sampled and Process

Because digital modulation has better immunity performances to interference, which is mostly discussed in the literatures for modulation classification. Here, it is assumed that there is a single carrier-transmitted signal in additive white Gaussian noise (AWGN) channel. The modulation types include 2ASK, BPSK, QPSK, 8PSK and 16QAM.

The signal data are produced by vector signal generator SMU200A. The sampling rate is 1 GHz. All the signal data of different modulation types have the same carrier frequency of 100 MHz and bandwidth of 25 MHz. Every sample has 2000 raw points and there are 25000 samples in total, 5000 samples for each modulation type. The only preprocess is to rescale the amplitude to the range of −2 V to 2 V. The spectrum map of sampled BPSK signal is shown in the Fig. 4.

For most classification and regression process, there is still possibility to get results even with small random noise added to the input. However, neural networks are proved not robust to noise [7]. One way to improve the robustness of neural networks is simply to do training process with input random noise data. So in training procedure to improve the robustness, training data of same SNR are included, which are also used to test the performance of proposed method in different SNR conditions.

When the network layers are not deep, it is not likely to encounter the problems like vanishing/exploding gradients [8, 9]. The principle of maximum likelihood is taken as the cost function, which means the cross-entropy between the training data and the prediction of the model is regarded as the cost function. The weights are initialized with Gaussian distribution initializers, which have zero means and unit variances. The SGD is involved with a mini-batch size of 256. The weight decay is 0.0001 and the momentum is 0.9. The learning rate starts from 0.1. When there are errors plateaus occur, the learning rate descends at rate of 10 times.
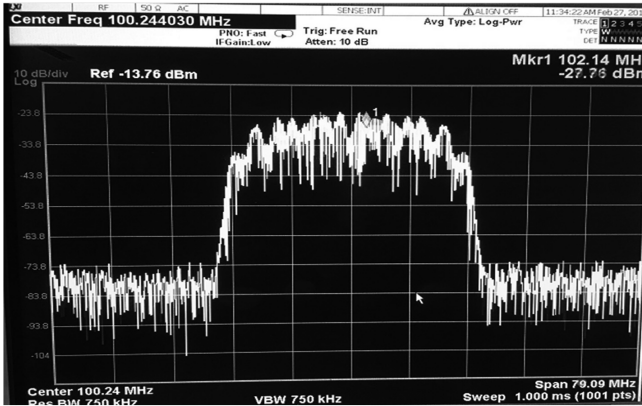
Fig. 4. The spectrum map of BPSK signal

As for the testing process, it is typically to use a simple separation of the same sampled data into training and testing datasets. In experiments, 80% data of the sampled signal is assigned to training dataset and 20% data of the sampled signal is assigned to testing dataset. Finally when the training is halt, we get the accuracy through inputting the testing datasets and statistical the accuracy.

3.2   The Improved CNN Architecture

It is found that the removal of the fully-connected layers of ALEXNET will reduce the amount of weight parameters and get little impact of the recognition accuracy performance. In this paper, the large kernel size is designed for better performances and acceptable complexity. Moreover, after investigating the deep neural network with more than 30 layers, it is found there are over-fitting problems. It is possible to apply a shallow neural network to compete modulation recognition for signals with reasonable SNR.

Based on the analysis above, the number of input neuron is set to 2000, which means every sample has 2000 raw points. The improved CNN is proposed with 3 convolutional layers, and each convolutional layer is followed by a max pooling layer. At the end of the CNN network, a 5-way fully-connected layer with Softmax is used to output the probability of 5 kinds of signal modulations classification. The convolutional layers have filter kernels with length of 40. 64 filter kernels are used in both the input layer and the second layer. For the third layer, the filter kernels are increased to length of 128. The max pooling layers perform down-sampling with stride of 2 and pool width of 3 to get overlapping pooling. We do not use the any regularization like dropout [10]. So, the improved CNN consists of 4 weighted layers, as shown in Fig. 5.
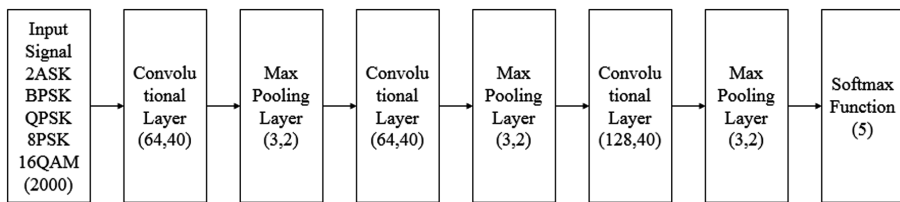
Fig. 5. The improved CNN structure

4   Experiments and Results Analysis

In order to evaluate recognition performances of deep neural networks and shallow neural networks, a 32-layers RESNET and the improve CNN with 4 layers are compared under condition of SNR = 0 dB.

Table 1. The accuracy comparison between the improved CNN and RESNET (SNR = 0 dB)

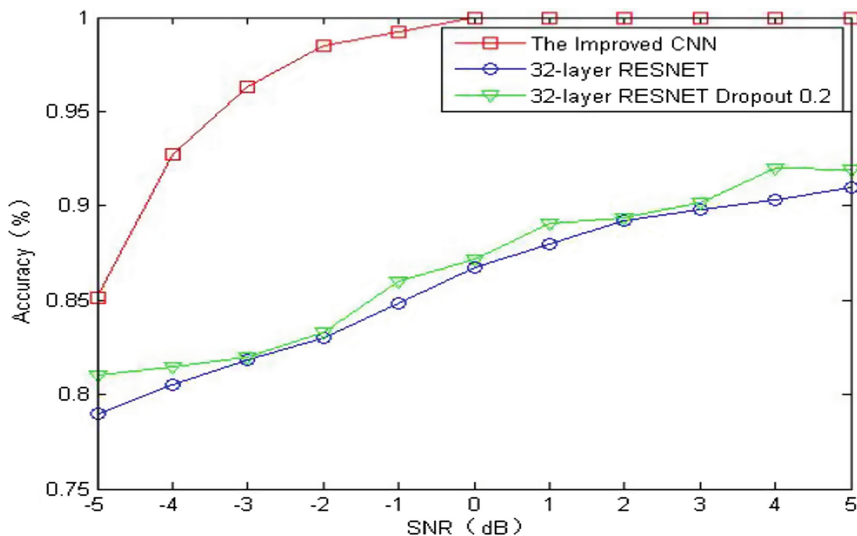|	The improved CNN	32 layers RESNET
Training accuracy	Approximate 100%	Approximate 100%
Testing accuracy	Approximate 100%	86.7%

Fig. 6. Recognition accuracy comparison

As shown in Table 1, although both have similar training accuracy, the improved 4-layer CNN has better testing accuracy than the 32-layer RESNET. Because of overfitting problem, the 32-layer RESNET may be unnecessarily large for automatic

modulation classification of SNR = 0 dB signals. For SNR = 0 dB signals, the modulation feature is distinct to be extracted for recognition, which can be completed by a shallow neural network. When the SNR of received signals is very weak, a deep neural network may provide stronger power to distinguish signal from noise.

As shown in Fig. 6, recognition accuracy performances of the improved CNN are compared with RESNET network in different SNR conditions. According to the results, the improved CNN has better recognition accuracy. But the descent speed in RESNET network is more stable than the improved CNN with shallow layers as the SNR drops to a low extent. The RESNET network is more powerful in extracting features when the SNR is low enough which leads to better robustness to noise. The RESNET network with dropout layer of 0.2 has little improvement, but more training time and training data are required, which is not recommend for applications of real-time signal detections.

5   Conclusions

The main purpose of this paper is to design a features extraction method based on convolutional neural network for automatic modulation classification of wireless signals. Real signal data generated by instruments are involved to train a CNN network and test recognition accuracy performances in different SNR conditions. It is found that the deep layer network architecture is not necessary for high SNR signals due to the overfitting problem, yet the shallow layer network architecture is more competent. By removing the fully-connected layer of CNN, the network topology is simplified to reduce complexity of training. According to the test results, the proposed improved CNN has better recognition accuracy than RESNET, which is attracting for real-time wireless signal detections.

References

1. Hazza, A., Shoaib, M., Alshebeili, S.A., Fahad, A.: An overview of feature-based methods for digital modulation classification. In: 2013 1st International Conference on Communications, Signal Processing and Their Applications, pp. 1–6. IEEE Press, New York (2013)
2. Zhu, X., Fujii, T.: A modulation classification method in cognitive radios system using stacked denoising sparse autoencoder. In: 2017 IEEE Radio and Wireless Symposium, pp. 218–220. IEEE Press, New York (2017)
3. Dai, A., Zhang, H., Sun, H.: Automatic modulation classification using stacked sparse auto-encoders. In: 13th IEEE International Conference on Signal Processing, pp. 248–252. IEEE Press, New York (2017)
4. Mendis, G.J., Wei, J., Madanayake, A.: Deep learning-based automated modulation classification for cognitive radio. In: 2016 IEEE International Conference on Communication Systems, pp. 1–6. IEEE Press, New York (2016)

5. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 29th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, pp. 770–778. IEEE Press, New York (2016)
6. Krizhevsky A., Sutskever I., Hinton G.E.: ImageNet classification with deep convolutional neural networks. In: 26th Annual Conference on Neural Information Processing Systems, vol. 25, no. 2, pp. 1097–1105. ACM, New York (2012)
7. Tang, Y., Eliasmith, C.: Deep networks for robust visual recognition. In: 27th International Conference on Machine Learning, pp. 1055–1062. ACM, New York (2010)
8. Bengio, Y., Simard, P., Frasconi, P.: Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. **5**, 157–166 (1994). IEEE Press, New York
9. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: 13th International Conference on Artificial Intelligence and Statistics, AISTATS 2010, pp. 249–256. Microtome Publishing, Brookline (2010)
10. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**, 1929–1958 (2014). Microtome Publishing, Brookline