
Automatic Modulation Classification Using Hybrid Convolutional Neural Network

1st Xiaofei Zhang¹, 2nd Anping Li¹, 3rd Lu Chai², 4th Xiaoying Ma¹, 5th Meiyong Wei¹
{zhangxf@src.org.cn¹, lianping@src.org.cn¹, chaibyr@bupt.edu.cn²,
maxiaoying@src.org.cn¹, weimeiyong@src.org.cn¹}

The state radio monitoring center¹,
Key Laboratory of Universal Wireless Communications, Ministry of Education,
Beijing University of Posts and Telecommunications, Beijing, P.R.China, 100876²

Abstract. Automatic modulation classification (AMC) plays an essential role in signal demodulation and interference identification. In this paper, we propose a novel AMC method using the Hybrid Convolutional Neural Network (HCNN), which combines with two different convolutional neural networks (CNNs) jointly using various signal features. In the former CNN, spectral correlation features (SCFs) are generated as network input, to classify FSK and BPSK. In the latter CNN, the Attention-based Densely Convolutional Neural Network (AD-CNN), which is trained using regular constellation images (RCs), is proposed to identify the modulation formats that are hardly recognized by the former CNN, such as QPSK, 16-QAM and 64-QAM. The simulation results demonstrate that HCNN displays superior classification performance than existing AMC methods with lower computational complexity.

Keywords: AMC, Deep convolutional neural network, Hybrid convolutional neural networks, Spectral correlation features, Regular constellation images.

1 Introduction

With the rapid development of wireless communications, new paradigms of telecommunications emerge currently [1]. However, the shortage of wireless spectrum is becoming an outstanding problem. To monitor spectrum resource tension and increase spectrum efficiency, the radio spectral sensing is introduced in commercial and civil applications [2]. With the task of ensuring the proper functioning of radio spectral sensing, automatic modulation classification (AMC) is utilized essentially as the link between signal receiving and demodulation to identify the signal format corrupted by noise [3].

Generally, AMC can be divided into two methodologies, *i.e.*, likelihood-based methods (LB) and feature-based methods (FB) [4]. LB methods have a better integrated performance of precision. However, LB methods need prior knowledge of statistical information of the signals, which increases the computational complexity in practice [5]. Comparatively, FB methods deal with modulation classification straightforwardly without prior information [6]. Various features are widely utilized in FB methods including high-order statistics [7], cyclostationary spectrum, *etc.* Kim *et al.* implemented the maximum of cyclostationary spectrum over frequency to reduce the computational complexity [8]. However, cyclostationary-based AMC lacks the ability to

deal with the classification between M-ary phase shift keying (M-PSK) and M-ary quadrature amplitude modulation (M-QAM). Huang *et al.* proposed grid constellation matrix and fully convolutional network for classification through a short training process. However, it is limited to classify Frequency-shift keying (FSK) using this method [9]. On this basis, this paper focuses on improving classification accuracy of FB methods.

The convolutional neural network with strong representative ability can be applied in AMC for higher classification accuracy [10]. Based on this, the authors proposed a low complexity deep neural network (DNN) to identify the signal modulation formats [11], but it hardly extracts high-dimensional features since the lack of the convolutional layers. In [12], convolutional neural network is utilized in AMC to extract features from constellation images. However, CNN with constellation images have limitations when the classification is conducted among FSK and M-QAM.

In this paper, a novel AMC method named Hybrid Convolutional Neural Network (HCNN), is proposed involving Deep Convolutional Neural Network (DCNN) and Attention-based Densely Convolutional Neural Network (AD-CNN). HCNN can effectively classify seven widely-used modulation formats with high classification accuracy and low computational complexity. Several contributions of this paper are listed as follows.

(1) Considering the various of signal features, the signal spectral correlation features (SCFs) and regular constellation images (RCs) are represented as the inputs of HCNN respectively. With SCFs and RCs, signal high-dimensional features are considered sufficiently.

(2) Deeper network structure and dropout layers, which can extract high-dimensional features directly and avoid overfitting, are exploited to obtain better performance in DCNN.

(3) AD-CNN combines the densely convolutional neural network and the attention block. It enhances the ability of feature extraction ulteriorly and reduces the network parameters simultaneously.

(4) Concretely, DCNN provides the classification ability among 2FSK, 4FSK, MSK and BPSK. Meanwhile, AD-CNN is utilized to improve the classification accuracy among QPSK, 16-QAM and 64-QAM. Simulation results show that higher classification accuracy is achieved comparing to existing AMC methods with lower computational complexity.

The remainder of this paper is organized as follows. Section II provides an overview of the signal model and the proposed SCFs and RCs. In Section III, we explain the network structure of HCNN. Section V exhibits simulation results and conclusion of this paper is given in Section VI.

2 Signal model and features extraction

2.1 Signal model

Assume that radio frequency signals are received by the single-antenna receiver. The signal transformed to baseband can be formatted as follows,

$$y(n) = he^{j(2\pi f_0 n + \theta)} x(n) + w(n), n = 1, \dots, N \quad (1)$$

where h denotes Rayleigh fading channel coefficients, f_0 and θ_0 denote the frequency and phase offset, respectively. N denotes the signal sample length. $x(n)$ denotes the $n - th$ complex sample in received segment signal. $w(n)$ denotes the additive white gaussian noise (AWGN) with mean zero and variance σ^2 . Meanwhile, the received signal-to-noise ratio (SNR) is defined as h^2/σ^2 .

2.2 Spectral correlation features

Cyclostationary signal $x(t)$ is a kind of random signal whose statistical characteristic parameters change periodically, *i.e.*,

$$R_x(t, \tau) = E\{x(t)x^*(t + \tau)\} \quad (2)$$

where its autocorrelation coefficient $R_x(t, \tau)$ and mean value $E[x(t)]$ both change with t in a period T for any value of τ . Compute the Fourier series of $R_x(t, \tau)$, then we get the cyclic autocorrelation coefficient. Considering the signal cyclic ergodicity, can be formatted as follows.

$$R_x^\alpha(\tau) = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})e^{-j2\pi\alpha t} dt \quad (3)$$

To derive a proper signal processing and surpass the influence of AWGN, we obtain the signal cyclic spectrum, which is the Fourier series of $R_x^\alpha(\tau)$ and can be defined as follows,

$$S_T^\alpha(t, f) = \frac{1}{T} X_T(t, f + \frac{\alpha}{2})X_T^*(t, f - \frac{\alpha}{2}) \quad (4)$$

where T denotes the signal segment length, α denotes the cyclic frequency, and X_T denotes finite time Fourier transform.

In a real scenario, signal samples are always be divided into K . Then we average the cyclic spectrum of each segment, which can be formatted as follows.

$$S_T^\alpha(t, f) = \frac{1}{K} \sum_{k=1}^K S_{T_k}^\alpha(t, f) \quad (5)$$

Considering various modulation formats have various signal cyclic spectrums, signal cyclic spectrum can be exploited for AMC. By normalizing the 3-dimensional cyclic spectrum along the Z axis, the 3-dimensional cyclic spectrum can be normalized to the 2-dimensional cyclic spectrum in the X-Y plane. Therefore, we extract the SCFs from signal cyclic spectrum, which are related to the carrier frequency f and symbol rate R_s . Some examples of SCFs under different modulation formats are presented in **Figure 1**. The figure shows that SCFs among 2FSK, 4FSK, MSK, BPSK have distinct patterns, while QPSK, 16-QAM, 64-QAM represented in the same form as SCFs.

Generally, SCFs have two advantages. Firstly, different SCFs are generated by different modulation formats so as to classify modulation. Secondly, SCFs have the robustness to AWGN, which could be exploited for modulation classification appropriately at low SNR.

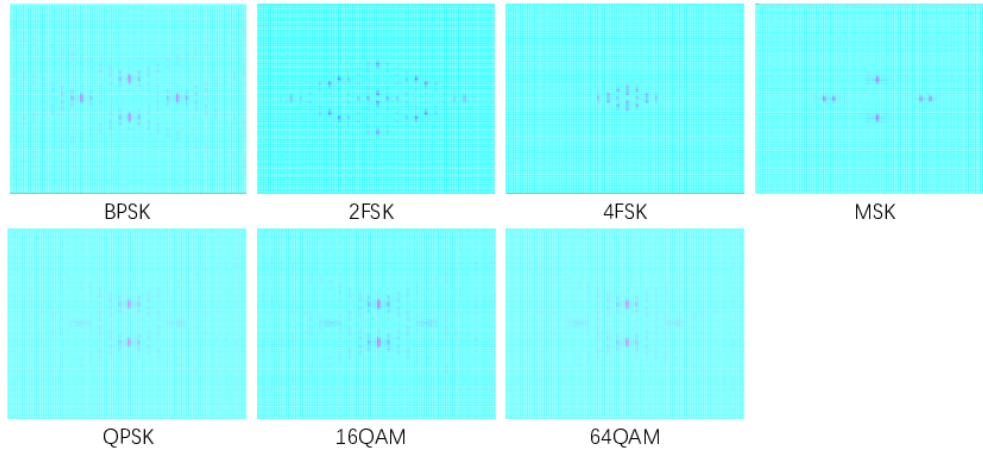


Fig. 1. Spectral Correlation Features for seven modulation formats.

2.3 Constellation Features

To further classify modulation formats of QPSK, 16-QAM, 64-QAM, received complex signal symbols are transformed into RCs as the input of the network. We plot the real and imaginary parts of the signals by transforming the signal samples into 2-dimensional scatter diagrams as RCs. Practically, RCs are widely used to learn effective representations from signals.

A generation diagram is conducted in **Figure 2** to illustrate RCs of three widely-used modulation formats (*i.e.* QPSK, 16-QAM, 64-QAM) at different SNRs. For different modulation formats under the same SNR, RCs are used to represent signal minutiae features. Generally, RCs provide a relatively new approach to extend its representation discrimination.

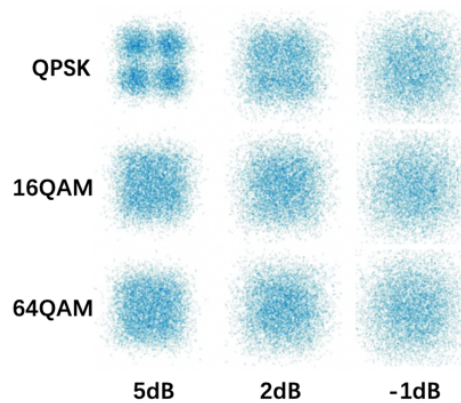


Fig. 2. Regular Constellation Images for three modulation formats under different SNRs.

3 HCNN based classification method

3.1 Structure of HCNN

Our proposed HCNN consists of two CNNs, so as to classify seven modulation formats of BPSK, QPSK, 2FSK, 4FSK, MSK, 16-QAM, 64-QAM. As shown in **Figure 3**, the former CNN, DCNN, is designed to classify modulation formats except QPSK, 16-QAM, 64-QAM, since the characterization limitation of SCFs. The latter CNN, AD-CNN, trained using signal RCs, is designed to classify QPSK, 16-QAM, 64-QAM.

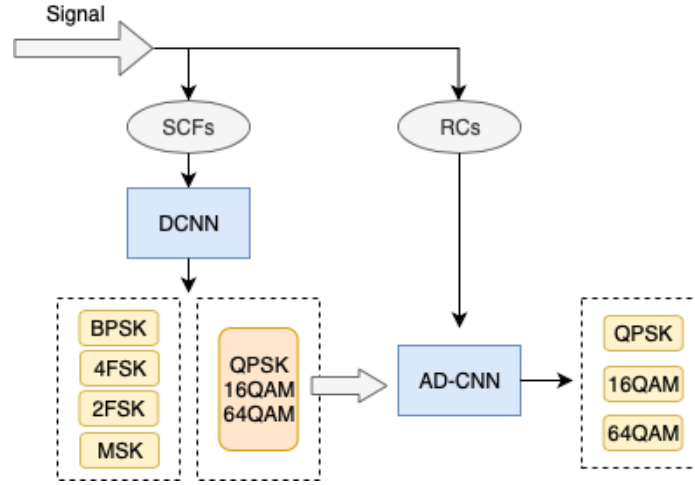


Fig. 3. The structure of HCNN.

3.2 Structure of DCNN

DCNN includes three convolutional blocks, one fusion block and three fully-connected blocks. Each convolutional block consists of four convolutional layers followed by batch normalization. In first three convolutional layers, convolutional kernel size decrease to 3×3 and stride step set as 1 in order to realize sparse connectivity between layers. Moreover, ReLU activation function is utilized to increase sparsity and nonlinearity. Specially, the fourth layer, whose kernel size is 2×2 and stride step is 2, replaces max-pooling layer to prevent feature loss in down-sampling. Besides, the channels are incremental among convolutional blocks, intending to get sufficient specific receptive fields. The fusion block unifies the form of feature representations for the fully-connected blocks by convolutional layer and fusion layer. In fully-connected blocks, there are 3 fully connected layers whose channel numbers are 256, 64, 4, respectively, reducing output channels and improving computational efficiency. The dropout is set as 0.6 to solve overfitting problem [14] and increase robustness to model mismatches. Additionally, the output channel number of the last hidden layer is limited to 2, which indicates that the dimension of the deep feature is reduced to 2 and beneficial to two-dimensional visualization. Finally, with Softmax [15], the output layer could sum all neurons to one and map multiple scalars to probability distributions.

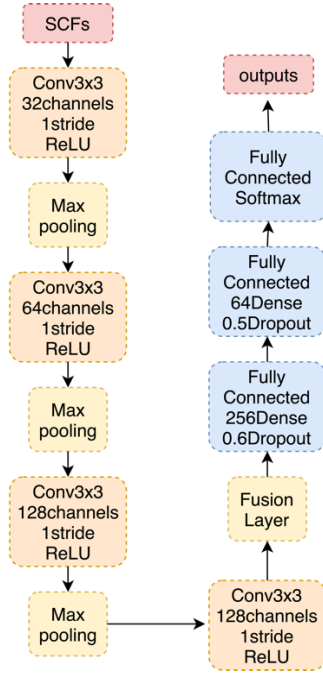


Fig. 4. The structure of DCNN.

3.3 Structure of AD-CNN

AD-CNN uses RCs to solve the classification among QPSK, 16-QAM and 64-QAM. Considering the similarity of RCs between 16-QAM and 64-QAM in low SNRs, AD-CNN is involved essentially to optimize the CNN structure so as to increase classification accuracy. As shown in **Figure 5**, AD-CNN consists of four parts, attention blocks, convolutional layers, max- pooling layers and fully-connected blocks.

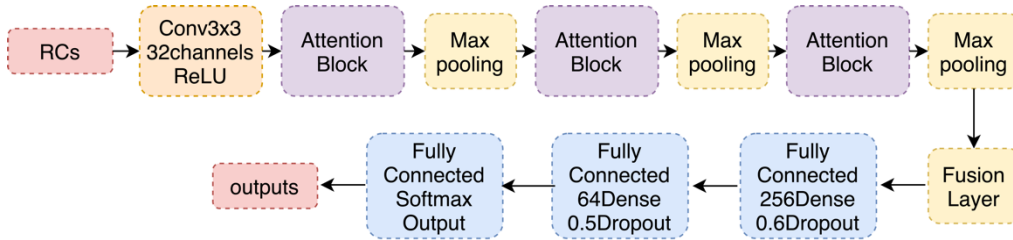


Fig. 5. The structure of AD-CNN.

Initially, attention block is composed of dense convolutional block and attention layer. For dense convolutional block, it has direct connections among nonadjacent layers, which is inspired by ResNet [16] and shown in **Figure 6**. To improve the feature utilization, each layer in the dense convolutional block receives the composition of preceding layers' output, meanwhile, as the input for subsequent layers. Incorporating both accuracy and efficiency, dense convolutional block brings two advantages. Firstly, it solves the problem of network gradient disappearance

and explosion. Secondly, it reduces the parameters that high-dimensional features can be extracted with low computational complexity.

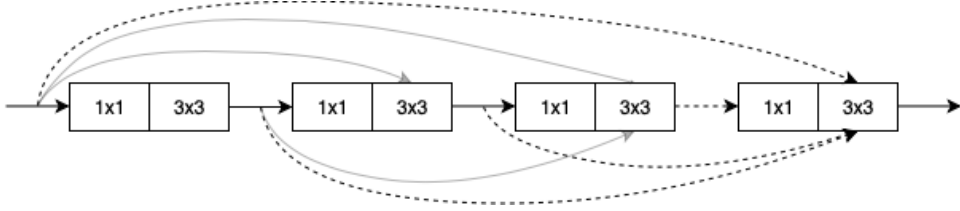


Fig. 6. The structure of dense convolutional block.

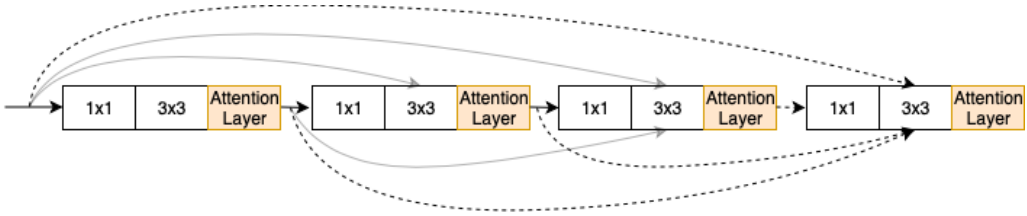


Fig. 7. The structure of attention block.

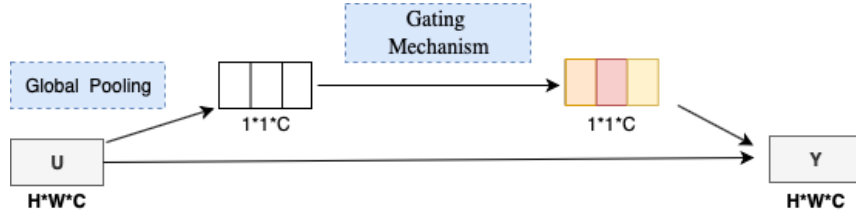


Fig. 8. The structure of attention layer.

Assume the network includes L layers, on which adopt a non-linear transformation $H_l(\cdot)$. The output of the l -th layer can be denoted as x_l . Therefore, the input of l -th layer, which contains the features from all the previous layers, can be denoted as follows,

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (6)$$

where x_0, x_1, \dots, x_{l-1} refer to the connection of the function diagram generated by layers.

Besides, attention layer is added after every two 1×1 convolution layer and 3×3 convolution layer in dense convolutional block, which is pictured in **Figure 7**. For attention layer, it could recalibrate the feature channel weights by calculating the interdependence among channels and improve significant performance of the most advanced architecture. As shown in **Figure 8**, there are four steps for the attention layers to recalibrate the features. Firstly, in order to construct the interdependence among channels, we convert the feature map $U \in R^{H \times W \times C}$ on the spatial dimension $H \times W$ to the channel descriptor. Next, the feature map U is shaped to $1 \times 1 \times C$ by a global average pooling layer. Thirdly, the gating mechanism composed of full

connection layers and sigmoid activation function is used to learn the nonlinear interaction among multiple channels. The feature map is shaped to $1 \times 1 \times C$ and then decoded back to $1 \times 1 \times C$ as the reweighed vector V_i with the gating mechanism. Lastly, the output of attention block Y can be formatted as follows,

$$Y = \sum_i^C U_i \cdot V_i \quad (7)$$

where U_i is the U of each channel and multiplied with reweighed factor V_i in vector V .

Generally, several optimizations of AD-CNN are listed as follows.

(1) In AD-CNN, feature reuse and recalibration are considered simultaneously. On the one hand, dense convolutional block achieves feature reuse and reduces the number of parameters remarkably. On the other hand, attention layer linked with convolution could reweight calibration for each feature channel. It helps to extract effective features and improve the classification accuracy.

(2) It is creative to make some optimization in the structure of AD-CNN so as to have an combination of classification accuracy and computational complexity. For instance, max-pooling layers in dense convolutional block are removed to protect low-level features. 1×1 convolutional layers are involved to protect global features and reduce feature redundancy.

4 Simulations

In this section, simulations are operated to illustrate the superiority of HCNN. The signal candidate modulation format set is $M = \{2\text{FSK}, 4\text{FSK}, \text{MSK}, \text{BPSK}, \text{QPSK}, 16\text{-QAM}, 64\text{-QAM}\}$. To reduce the computational complexity and improve the generalization ability, the normalized pixel size of SCFs and RCs are down sampled to 80×80 . SCFs and RCs are both obtained by received signals with different pre-defined sample lengths, *i.e.*, 1024, 2048, 4096, 9600. In the training stage, 40000 SCFs and RCs are involved respectively for each modulation at a given SNR as training data. In the classification stage, 10000 SCFs and RCs in different assumed conditions are used to calculate the classification accuracy P_{cc} .

In **Figure 9**, we compare the performance of HCNN under different settings of sample lengths. For same sample length, the HCNN performs consistently within the range of SNR, proving the practicability of HCNN. Moreover, better classification performance by employing more signal samples, proving the asymptotic behavior of HCNN.

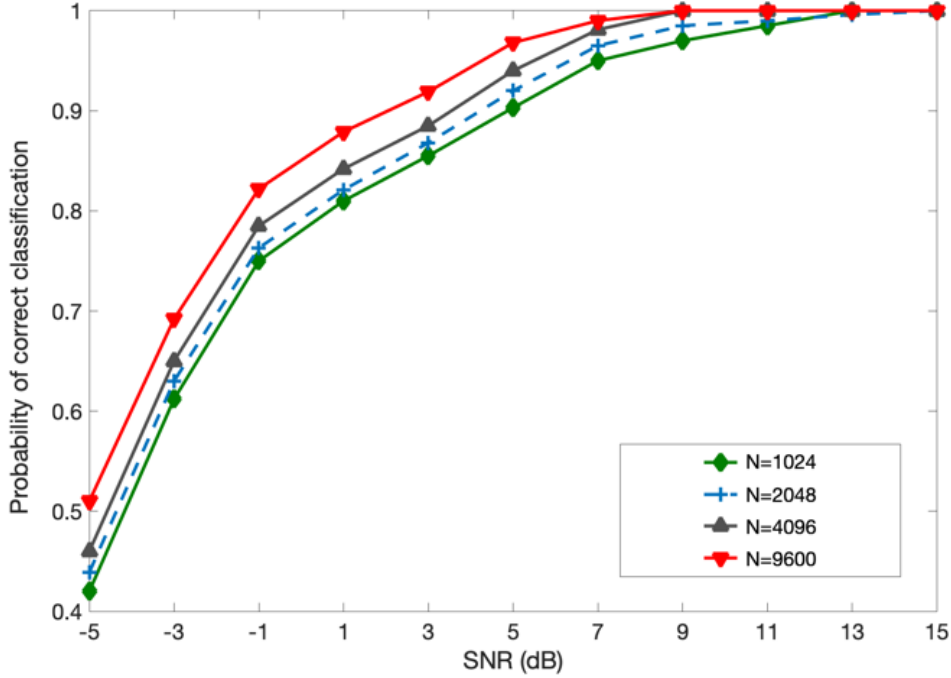


Fig. 9. The performance of HCNN under different settings of sample lengths.

Figure 10 is an overall methods comparison among HCNN proposed, HCNN with both DCNN, CNN with RCs [6] and CNN with SCFs [8] under different SNRs, and the sample length is set as 4096. The result verifies that proposed HCNN outperforms the others for the entire SNR range. Specifically, HCNN proposed yields 3 dB gains over HCNN with both DCNN at 95% P_{CC} , illustrating the superiority of the HCNN. However, the accuracy is only up to 0.63 when using a single kind of signal feature, indicating the wide range of classification modulation formats of HCNN. The reasons are from two aspects. Initially, AD-CNN in HCNN can extract more discriminative features from input features. Additionally, it is illustrated that single-feature limits the types of classification modulation. More importantly, the per epoch training time of proposed HCNN and HCNN with both DCNN are 90 seconds (s), 710s, respectively, using the same dataset and equipment (a GTX1080 GPU). It shows the lower computational complexity of HCNN.

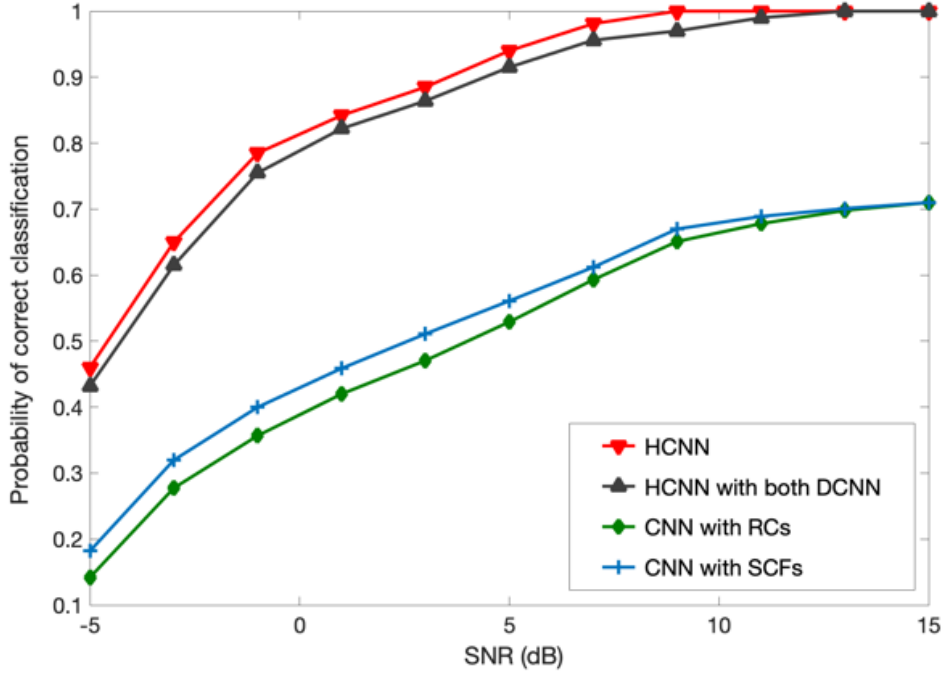


Fig. 10. The performance comparison among HCNN and other AMC methods.

5 Conclusion

In this paper, a novel AMC method named HCNN is proposed. First of all, the HCNN combines with DCNN and AD-CNN to make classification decisions of seven modulation formats with high accuracy. DCNN uses deep convolutional neural network and replaces the pooling operation by convolutional layer to reserve discriminative features. Especially, AD-CNN is designed to consider feature reuse and recalibrationare simultaneously, in order to improve performance and reduce computational complexity. Besides, SCFs and RCs are conducted respectively to expand the modulation scopes of classification. Simulation results verify the accuracy superiority and low computational complexity of HCNN.

Acknowledgments.

This work was supported in part by the National Key Research and Development Program of China under Grant (2018YFF0301202, 2019YFB1804400), in part by the National Natural Science Foundation of China under Grant (61801052, 61227801), and in part by the Beijing Natural Science Foundation under Grants (4202046).

References

- [1] E. Pateromichelakis and K. Samdanis, "A graph coloring based inter-slice resource management for 5G dynamic TDD RANs," in Proc. IEEE Int. Conf. Commun. (ICC), MO, USA, pp. 1-6 (2018)

-
- [2] Y. Ma, Y. Gao, Y. C. Liang, and S. Cui, "Reliable and efficient sub-Nyquist wideband spectrum sensing in cooperative cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2750–2762 (2016)
- [3] O. A. Dobre, A. Abdi, Y. Bar-Ness and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137-156 (2007)
- [4] Z. Zhu and A. K. Nandi, *Automatic modulation classification: Principles, Algorithms and Applications*. Hoboken, NJ, USA: Wiley (2005)
- [5] T. V. R. O. C`amara, A. D. L. Lima, B. M. M. Lima and A. I. R. Fontes, "Automatic modulation classification architectures based on cyclostationary features in impulsive environments," *IEEE Access.*, vol. 7, no. 1, pp. 138512-138527 (2019)
- [6] S. Huang, L. Chai and Z. N. Li, "Automatic modulation classification using compressive convolutional neural network," *IEEE Access.*, vol. 7, no. 1, pp. 79636-79643 (2019)
- [7] S. Huang, Y. Yao, Y. Xiao and Z. Feng, "Cumulant based maximum likelihood classification for overlapped signals," *Electr. Letters.*, vol. 52, no. 21, pp. 1761-1763 (2016)
- [8] K. Kim, I. A. Akbar, K. K. Bae, "Cyclostationary approaches to signal detection and classification in cognitive radio," in *Proc. IEEE Int. Sym-posium Dynamic Spectrum Access Netw. (DySPAN)*, Dublin, Ireland, pp. 212-215 (2007)
- [9] S. Huang, Y. Jiang, Y. Gao, Z. Feng and P. Zhang, "Automatic modulation classification using contrastive fully convolutional network," *IEEE Wireless Commun. Letters.* (2019)
- [10] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074-4077 (2019)
- [11] P. T. D. Boer, D. P. Kroese and S. Mannor, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, vol. 133, no. 1, pp. 19-67 (2005)
- [12] T. J. Shea, J. Corgan and T. C. Clancy, "Convolutional radio modulation recognition networks," *Commun. Comput. and Inf. Sci.*, vol. 629, no. 1, pp. 213- 226 (2016)
- [13] D. N. Godard, "Self-recovering equalization and carrier-tracking in two dimensional data communication systems," *IEEE Transactions on Communications.*, vol. 28, no. 11, pp. 1867-1875 (1980)
- [14] T. J. Shea, J. Corgan and T. C. Clancy, "Convolutional radio modulation recognition networks," *Commun. Comput. and Inf. Sci.*, vol. 629, no. 1, pp. 213- 226 (2016)
- [15] Y. Sun, X. Wang and X. Tang, "Deep learning face representation from predicting 10,000 classes", in *Proc. IEEE Conf. Comput. Vision and Pattern Recognition. (CVPR)*, Ohio, USA, pp. 1891-1898 (2014)
- [16] K. He, X. Zhang and S. Ren, "Deep residual learning for image recognition," in *2016 IEEE Conf. Comput. Vision and Pattern Recognition.(CVPR)* (2016)