The Performance Analysis of Signal Recognition Using Attention Based CNN Method

ZAN YIN\textsuperscript{1,2}, BO CHEN\textsuperscript{1}, (Senior Member, IEEE), WEIMIN ZHEN\textsuperscript{2}, CHAOJIE WANG\textsuperscript{1}, AND TING ZHANG\textsuperscript{1}

\textsuperscript{1}National Laboratory of Radar Signal Processing, Xi\'an University, Xi\’an 710000, China
\textsuperscript{2}China Research Institute of Radiowave Propagation, Qingdao 266000, China

Corresponding author: Zan Yin (lantianauto2003@126.com)

This work was supported by the National Key Research and Development Program of China under Project 2018YFB0505100.

\textbf{ABSTRACT} Modulation recognition has always been an important task in the development of the cognitive radio. At present, there are two main application methods for signal data, namely, directly using the signal sequence and using some conversions such as constellation diagram. In this paper, the converted contour stella images are adopted as data source for research. The deep learning method has been proposed, which is called Image-based CNN with Attention Model (ICAM). ICAM is based on Residual Neural Network (ResNet). To evaluate the performance of ICAM, we generate a dataset which contains contour stella images involving of 8 kinds of modulation types with signal to noise ratio (SNR) from \(-6dB\) to \(20dB\). Compared with other state-of-the-art image-based methods, which are including those using constellation diagram and contour stellar images, ICAM makes more satisfactory performance. Besides, the Grad-Cam technology is applied to visualize and prove the effectiveness of ICAM.

\textbf{INDEX TERMS} Modulation recognition, CNN, attention mechanism, contour stella images, I/Q data.

\section{I. INTRODUCTION}

Automatic Modulation Recognition (AMR) technology is an important part of the field of Cognitive Radio (CR) \cite{1}. AMR is usually regarded as the foundation of other wireless communication problems, such as parameter estimation, signal demodulation, and spectrum management \cite{2}. Parameter estimation can also be used for feature extraction of equipment for security identification \cite{3}. Moreover, it has a large range of use in civilian and military applications, especially in dynamic spectrum management and interference identification.

Modulation aims to improve the effectiveness and reliability of long-distance communication \cite{4}. Amplitude, frequency, phase can be the candidates for modulation contents. Modulation recognition then is to figure out which kind of the modulation type is used for the received wireless signal \cite{5}. Generally, In-phase and Quadrature (IQ) signal data can be obtained by the receiver such as GNU Radio or USRP. And it mainly includes three methods to deal with the IQ data in AMR \cite{6}. 1. Convert I/Q signals into image formats such as constellation diagram, time-frequency diagram, and spectrogram; 2. Do research directly on the paired I / Q sequence data; 3. Extract features for further research based on expert knowledge. There is a trend in AMR applied to image processing and I/Q sequence rather than crafting expert features \cite{7}. Because extracting features exists a significant challenge that it performs well on specialized tasks but often lacks flexibility. On the contrary, the image itself can show its visible feature, for example, the spatial information in the constellation diagram, which provides great convenience for feature extraction. The use of I/Q sequence data is like solving the problem from the source, which is natural and direct. In this paper, we will explore the potential of image-based methods.

Deep learning has made great success in the field of image processing, natural language processing \textit{et al.} Deep neural networks use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. It can automatically optimize the extracted features for minimizing classification error. Among them, Convolution Neural Network (CNN) is good at extracting features of spatial data and shows good shift-invariant properties, which is similar to the filters in traditional communication. In this paper, CNN is the candidate model separately to deal with the converted signal image and signal sequence. To maximize the potential of the
model, the attention mechanism is introduced into the CNN model to extract more valuable information.

The overall contributions of this study are as follows:

1) To obtain the image data, the I/Q signal is first converted into Contour Stella Image which is the constellation diagram with density information. We then introduce a plug-and-play attention module into the CNN model to improve the ability to extract features precisely. The proposed recognition structure is called Image-based CNN with the Attention Model (ICAM).

2) We compare the ICAM with other state-of-the-art image-based methods, including those using constellation diagram and contour stellar images. Besides, Grad-Cam technology is applied to visualize and prove the effectiveness of ICAM.

The rest of the paper is organized as follows. Section II introduces the related work. Section III describes the design concept and construction of the image-based CNN models. Section IV shows the experimental results obtained for different datasets and compares our approach with state-of-the-art methods. Section V provides a discussion and conclusions.

II. RELATED WORKS

Modulation classification is a major issue in the communications system. Many scholars have presented different methods to deal with different application cases. Here, related works are about the approaches applying deep learning on the converted signal images and I/Q data.

In 2017, Yao et al. converted the raw modulated signals into a constellation diagram that has a grid-like topology [8]. And then AlexNet was adopted for classification compared with SVM and traditional cumulant based methods. Due to the limitation of image resolution and loss of original information, their method can’t always outperform the compared approaches. Later in [9], Yao et al. continued their work using GoogleNet and AlexNet. In order to solve the problem of image resolution loss, they have developed several methods to represent modulated signals in data formats with grid-like topologies for the CNN, such as constellation diagram, gray image, enhanced gray image, and 3-channel image. The size of selected area and image resolution are under research. The experiment results showed that the DL-based approaches can automatically extract features and obtain a superior recognition performance than traditional cumulant and machine learning based approaches. In [10], Yao et al. applied their CNN based methods to solve the problem of modulation recognition in varying SNR environment. In their work, multiple labeling technology was adopted, where both SNR and modulation type are known for each signal in training phase. And in inferring stage, the SNR was first estimated and then the modulation type could be inferred later.

In 2017, Yun et al. proposed a conversion algorithm for complex signals [11], named contour stellar image, which has different colorful areas in constellation diagram. AlexNet, GoogLeNet, ResNet, and VGG were separately adopted on their dataset and each model achieved similar performance over different SNR. In 2018, Yun et al. continued their previous work and proposed a data augmentation method using auxiliary classifier generative adversarial networks [12]. Several measures were adopted to avoid unconvergence problems and the proposed method obtained 0.1%-6% accuracy improvement compared with the baseline. Besides, in [13], Yun et al. presented a semi-supervised strategy based on Generative Adversarial Networks to exploit unlabeled data. Also, experiments were conducted on contour stella image dataset. Their approach could handle three different sources of training data including real images with labels, real images without labels, and images from the generator.

Guan et al. proposed a deep learning-based modulation recognition method which contains two CNNs [14]. Dropout was applied to replace the pooling operation to obtain better performance. They have explored both I/Q samples and contour stella images. The contour stella images were supplement for the I/Q samples due to the bad performance on the QAM signals. In [15], Zeng et al. looked into the potential of spectrogram CNN (SCNN) applied on the spectrogram images. They first transformed the 1-D radio signals into spectrogram diagrams using short-time discrete Fourier transform. They also adopted a Gaussian filter to reduce noise. The experiment was conducted on the RadioML2016.10a dataset and the result showed the SCNN with noise reduction worked better than some of existing models especially at high SNR.

As for the I/Q samples, in [16], O’Shea et al. first presented the RadioML2016.10a dataset using GNU Radio and this dataset has been the benchmark dataset for training and evaluating the performance of modulation recognition methods. Besides, they proposed a CNN based method for modulation recognition. They evaluated the influence of Dropout and compared their CNN model with other expert-based machine learning strategies such as Expert-KNN. Their work and dataset become baseline for the I/Q based methods. Latter in 2017, O’Shea et al. applied Convolutional Long short-term Deep Neural Networks (CLDNNs) on the task of modulation recognition [17]. They tried to find the best number of filters and filter size for RF modulation recognition and evaluated the influence of network depth and filter size. They concluded the features purely CNN architectures can learn will affect the performance but not the network depth.

In [18], Yin et al. made a performance comparison of traditional high-order cumulant-based neural network and deep neural network (DNN) using RadioML2016.10a dataset. They introduced a short-time VGG which contains only three residual blocks. They created a dataset containing 19 types of digital modulation signals in MATLAB. Experiments were conducted on short-time VGG and short-time ResNet. The result were said to obtain a better advantage than O’Shea models. In [6], they proposed a Fully Dense Neural Network using residual blocks. Their model consisted of three modules: the Residual block, the Transition block, and the Signal Attention block. Among them, the Transition block adopted a one-dimensional pooling layer to preserve signal information. And the attention block was like location-based
method, which used a fully-connected layer to obtain the weights. Besides, they also picked 11 kinds of modulation signals with SNR of 18 for visualization. In [19], a data-driven model for modulation classification based on long short-term memory network was proposed. The intuition of using LSTM model for classification was based on the fact that different modulation types show different amplitude and phase characteristics, and the model can effectively learn these temporal correlations. In addition, this paper expanded the dataset to evaluate the sample rate. Experiment results showed that the model performed better in the case of time domain amplitude and phase information than that of only time domain IQ samples. In [20], Yun et al. proposed a new filter-level pruning technique based on activation maximization (AM) that omits the less important convolutional filter. This method takes into account the resource constraints of deep learning deployment equipment. Compared to other methods, the convolutional neural network pruned via the AM method achieves equal or higher classification accuracy.

### III. IMAGE BASED MODULATION RECOGNITION

#### A. MOTIVATION OF THE ICAM
As we all know, the convolution neural network is a hot and strong tool in the field of Computer Vision (CV). It has its excellent capabilities of image processing, such as image classification, object recognition, and style transfer et. al. CNN is also suitable for modulation recognition. In addition, there exists the concept of space and channel in the convolution neural network. Besides, not every zone in the space is helpful for the extraction of the final information. Similarly, the importance of each channel is also different. Attention mechanism promises to make good use of different space and channel.

#### B. DATA PREPROCESSING
Modulation recognition can be regarded as a problem of N-class classification problem. Supposing a base-band time series signal $s(n)$, the path loss or constant gain $c$, and an additive Gaussian white noise $g(n)$, the simple expression of received signal can be written as:

$$r(n) = s(n) * c + g(n)$$

(1)

Analytically, the expression of received sampled signal can be re-written as:

$$r(t) = Ae^{j(2\pi f_0 n + \theta_0)} + g(n)$$

(2)

where $A$ is the amplitude coefficient, $f_0$ is the frequency offset, $\theta_0$ is the phase offset, $n$ is the symbol index, and $g(n)$ is the complex additive white Gaussian noise with mean 0 and power $\sigma_g^2$.

After the received signals are obtained, the I/Q parts are first plotted on the image to form a constellation diagram. Each point in the image represents a sample in signal wave. Different modulation type has its unique shape of constellation which leads to different cluster of samples. So different area has different number of samples. Based on above, a density square window is used to slide on the constellation diagram. During the sliding, count how many data points are in the square. And so the density $\rho(i,j)$ of each window can be calculated as:

$$\rho(i,j) = \frac{\sum_{m=1}^{m_2} \sum_{j=1}^{n_2} dots(i,j)}{\sum_{m=1}^{m_1} \sum_{j=1}^{n_1} dots(i,j)}$$

(3)

where $(m_1, n_1)$ and $(m_2, n_2)$ separately represent the coordinates of the upper-left and lower-right vertices of the current window. $(W_0, H_0)$ and $(W_1, H_1)$ are the coordinates of the upper-left and lower-right vertices of the whole constellation diagram. And the $dots(i,j)$ is a function to determine whether point $(i,j)$ should be counted, which is as follow:

$$dots(i,j) = \begin{cases} 
1 & \text{if sample exists in } (i,j) \\
0 & \text{if no sample exists in } (i,j)
\end{cases}$$

(4)

Later, to give different value of density unique identification, a color bar is used as shown in Fig.1. The color bar is continuous and roughly contains three areas, the low density area, the medium area, and the high density area. The larger the density is, the brighter the color is.

From the point of view of signal processing, the contour stellar images can describe in-depth statistical information. Both constellation diagram and contour stellar images at high signal-to-noise ratio can well represent the statistical information of modulation signals, such as Gaussian noise, Non-coherent single frequency interference, Phase noise in the signal waveform. However, due to the severe interference of the noise to the constellation diagram at low SNR, the statistical information of the modulated signal under the constellation diagram cannot be displayed well. Simultaneously, in above-mentioned way, contour star images are obtained. Fig.2 and Fig.3 show the contour stellar images of different modulation type at SNR of 14dB and 4dB, respectively. It is easy to find that the shapes of modulation types in contour stellar images are as the same as those in the constellation diagrams. With the density representation, the feature is more significant than that in the constellation diagram.

#### C. NETWORK STRUCTURE

1) ATTENTION MODULE
The idea of attention mechanism takes inspiration from the behaviour of human. Attention happens, to some extent, when we focus mainly on some local regions of an image or some special words in one sentence. Attention mechanism helps to do more with limited resources.

![FIGURE 1. The color bar for the transfer of density.](image-url)
In this paper, we adopt the Convolutional Block Attention Module (CBAM) proposed by Sanghyun [21]. CBAM is said to be a plug-and-play block for convolution neural network. It works after the convolution layer and helps the convolution layer to make good use of every spatial and channel information. There are two concepts in CBAM, the channel attention and spatial attention.

The channel attention is to compress the feature map in the spatial dimension. Average-pooling and Max-pooling are both used to aggregate spatial information of feature maps. And the pooling results are separately fed into the shared fully-connected layers. Element-wise summation is then applied to merge the average-pooled features and max-pooled features. Use the sigmoid to obtain the channel attention value latter. Finally, the output is obtained by the product of original feature map and channel attention value. The whole process of channel attention can be summarized as:

$$M_C(F) = \sigma(W_1(W_0(F^c_{avg})) + W_1(F^c_{max}))$$  (5)

where $F$ is the input feature map, $W_0$ and $W_1$ separately represent the weight matrix of hidden layer and fully-connected layer, $F^c_{avg}$ and $F^c_{max}$ are the pooling contexts.

The spatial attention module compresses the channels and performs average pooling and maximum pooling in the channel dimensions. Then the previously extracted feature maps (the number of channels are both equal to 1) are combined to obtain a 2-channel feature map. In summary, the spatial attention works as follow:

$$M_S(F) = \sigma(f_7 \times 7([F^s_{avg}, F^s_{max}])))$$  (6)

where $F^s_{avg}$ and $F^s_{max}$ are the average-pooling context and max-pooling context. $f_7 \times 7$ stands for the kernel size of convolution.

2) IMAGE-BASED CNN WITH ATTENTION MODEL

In this subsection, our image-based modulation recognition model is built based on CBAM and ResNet18. The general pipeline of dealing with image data is shown in Fig.6. The CBAM is used as a plug-and-play module attached behind each convolutional block of ResNet18. The structure of ResNet18 in our paper is similar to the standard one as introduced in [22].

In detail, at the beginning of the ICAM, a $7 \times 7$ filter for downsampling with $stride = 2$ is first adopted in the input layer followed by a $3 \times 3$ MaxPooling with 2 strides. The $7 \times 7$ filter is used to keep most of the original information but avoid adding the number of filters. Besides, there exist four blocks and each block has the same $3 \times 3$ kernel size of the filter. As for the number of filters, the initial number is 16 and the number of filter in the next block is twice of that in the former block. Also there exists an attention module after each block to calculate the weights and
it work together with the output of block to obtain the input of next block. Finally, there exists several fully connected layers and the output size equals 8 with softmax activation function.

IV. EXPERIMENTS RESULTS
A. DATASET AND EXPERIMENTS SETTINGS
As for the modulation signals that are converted into images, we use the Matlab2017a to generate as [11] did. The dataset contains 8 different modulation types, binary phase-shift keying (BPSK), 4 amplitude-shift keying (4ASK), quadrature phase-shift keying (QPSK), offset QPSK (OQPSK), eight phase-shift keying (8PSK), 16 quadrature amplitude modulation (16QAM), 32QAM, 64QAM. The white Gaussian noise is added and the SNR of the samples ranges from $-6$dB to 20dB with a step size of 2dB. And then the preprocessing as shown in chapter 3 is applied on these data. Finally, there are 10,000 images for training and 1,000 for testing. And each image contains 10,000 pairs of I/Q samples.

We conduct 2 groups of experiments in total.
- **Experiment 1** compares the ICAM with other state-of-the-art image-based methods, including those using constellation diagram and contour stellar images.
- **Experiment 2** uses the Grad-Cam technology for qualitative analysis of the recognition performance, which visualizes the importance of spatial locations in the contour stellar images.

B. COMPARISON WITH OTHER IMAGE-BASED METHODS
We compare our approach with those described in other studies. Among them, the approach using AlexNet on constellation diagram is presented in [9]. And for comparison, AlexNet and ResNet are adopted in [11] to use the contour stellar images for modulation recognition. We implement the baseline methods according to the descriptions provided in the appropriate papers. It should be noted that the data set used in this paper is exactly the same as that in [11], and the benchmark ResNet model used is also the same. The experiment result is shown in Fig.7.

In general, the recognition accuracy of the four methods all show an upward trend as the signal-to-noise ratio increases. This is mainly due to the fact that there is less noise under high signal-to-noise ratio. Even if the same AlexNet is used, the method based on contour stellar images is better than the method based on constellation diagram over every SNR. This verifies that the features contained in the contour stellar images that retains density information are more meaningful than the constellation diagram. Besides, for the same contour stellar images dataset, except that ResNet is slightly better at low signal-to-noise ratios, AlexNet and ResNet perform similarly. Our ICAM directly applies attention module after each convolution layer of ResNet. It can be found that the performance of ICAM is slightly better than the general ResNet, which is the same as expected. Especially, at SNR of $-6$dB, the recognition accuracy of the ICAM is about 3% higher than that of the general ResNet18 model. The good performance of ICAM at low SNR shows that the attention mechanism can really help extract effective features from difficult-to-recognize data that is polluted by noise. At 0dB, the accuracy rate of ICAM
has reached over 99%. For further proof, we have selected recognition results on the testing dataset at several SNR and presented them in the form of confusion matrices as shown in Fig.8.

Overall, when the signal-to-noise ratio is 14 dB and 20 dB, all signal modulation types are classified correctly. When the SNR is less than 0db, 4ASK, 16QAM, 32QAm and 64QAM can still be distinguished. However, it is difficult to distinguish PSK-based modulation types, such as QPSK, OQPSK and 8PSK. Especially, 8psk signal is easier to be divided into QPSK and OQPSK. This may be due to the fact that the QAM signals have both amplitude characteristics and phase information, while the PSK signals only have differences in phase characteristics. The key to distinguishing PSK signals is the phase information. The low signal-to-noise ratio brings strong randomness and blurs the phase characteristics, so the recognition ability for PSK signals becomes weak.

C. VISUALIZATION USING GRAD-CAM

For the qualitative analysis, the Grad-Cam is applied on the ICAM using the contour stellar images. Grad-Cam is a visualization approach which can calculate the importance of the spatial locations in convolutional layers [23]. It is hailed as a visual interpreter for CNN. By observing the regions that network has considered as important for predicting a class, we attempt to look at how this network is making good use of features.

We apply the Grad-Cam technology on the combination of last convolution layer and attention module. The data for experiment are contour stellar images of each modulation type at 14dB. The reason why the signals at 14dB are selected for the visualization is that the shape of the contour stellar diagram at 14dB is clear enough to recognize by human. In this way, if the highlighted part of the heat map (that is, the hotter area) is similar to the area recognized by human, it can be proved that the unique features in the constellation diagram are extracted.

Fig.9 shows the experiment results. The upper part of Fig.9 shows the contour stellar diagrams of the different modulation signals and the corresponding heat maps are at the lower part. It can be seen that in the heatmap of each modulation signal, the regions with prominent features, that is, the areas of interest of the ICAM are consistent with the areas of human interest, which illustrates the effectiveness of the model. For example, the BPSK is the simplest form of phase shift keying. As we all know, it uses two phases which are separated by 180°. In Fig.9(b), they are shown on the real axis, at 0° and 180°. It is very clear to find that the bright zone of corresponding heatmap is located at the center of the those two clusters. Another case is 8PSK. The shape of contour stellar diagram is distinguishable and there are clusters of sample points in 8 directions with the similar distance to the origin. In the corresponding heatmap, the shapes of bright areas are not that regular and some areas even overlap each other. However, it is also not hard to
find eight clusters in 8 direction with similar distance to the origin.

V. DISCUSSION AND CONCLUSION
In this paper, the deep learning frameworks based on convolution neural network and attention mechanism is constructed for modulation recognition. The ICAM can process the contour stella images. For the image-based method, the experiment results showed that the attention will improve the performance of CNN model. And with the help of Grad-Cam, visualization results showed that ICAM can capture the significant areas in the image effectively as human does. In future work, we will apply more deep learning methods to automatic modulation recognition and further improve the accuracy of modulation recognition at low SNR.

REFERENCES
[1] T. Liu, Y. Guan, and Y. Lin, “Research on modulation recognition with ensemble learning,” EURASIP J. Wireless Commun. Netw., vol. 2017, no. 1, p. 179, Dec. 2017.
[2] C. Weber, T. Felhauer, and M. Peter, “Automatic modulation classification technique for radio monitoring,” Electron. Lett., vol. 51, no. 10, pp. 794–796, May 2015.
[3] Z. Zhang, X. Guo, and Y. Lin, “Trust management method of D2D communication based on RF fingerprint identification,” IEEE Access, vol. 6, pp. 66082–66087, 2018.
[4] Y. Lin, X. Zhu, Z. Zheng, Z. Dou, and R. Zhou, “The individual identification method of wireless device based on dimensionality reduction and machine learning,” J. Supercomput., vol. 75, no. 6, pp. 3010–3027, Jun. 2019.
[5] A. Tsakmalis, S. Chatzinotas, and B. Ottersten, “Centralized power control in cognitive radio networks using modulation and coding classification feedback,” IEEE Trans. Cognit. Commun. Netw., vol. 2, no. 3, pp. 223–237, Sep. 2016.
[6] M. Du, Q. Yu, S. Fei, C. Wang, X. Gong, and R. Luo, “Fully dense neural network for the automatic modulation recognition,” 2019, arXiv:1912.03449. [Online]. Available: http://arxiv.org/abs/1912.03449

[7] T. O’Shea and J. Hoydis, “An introduction to deep learning for the physical layer,” IEEE Trans. Cognit. Commun. Netw., vol. 3, no. 4, pp. 563–575, Dec. 2017.

[8] S. Peng, H. Jiang, H. Wang, H. Alwaheed, and Y.-D. Yao, “Modulation classification using convolutional neural network based deep learning model,” in Proc. 26th Wireless Opt. Commun. Conf. (WOCC), Apr. 2017, pp. 1–5.

[9] S. Peng, H. Jiang, H. Wang, H. Alwaheed, Y. Zhou, M. M. Sebdani, and Y.-D. Yao, “Modulation classification based on signal constellation diagrams and deep learning,” IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 3, pp. 718–727, Mar. 2019.

[10] X. Xie, Y. Ni, S. Peng, and Y.-D. Yao, “Deep learning based automatic modulation classification for varying SNR environment,” in Proc. 28th Wireless Opt. Commun. Conf. (WOCC), May 2019, pp. 1–5.

[11] Y. Lin, Y. Tu, Z. Dou, and Z. Wu, “The application of deep learning in communication signal modulation recognition,” in Proc. IEEE/CIC Int. Conf. Commun. China (ICCC), Oct. 2017, pp. 1–5.

[12] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, “Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks,” IEEE Access, vol. 6, pp. 15713–15722, 2018.

[13] Y. Tu, Y. Lin, J. Wang, and J.-U. Kim, “Semi-supervised learning with generative adversarial networks on digital signal modulation classification,” Comput. Mater. Continua, vol. 55, no. 2, pp. 243–254, 2018.

[14] Y. Wang, M. Liu, J. Yang, and G. Gai, “Data-driven deep learning for automatic modulation recognition in cognitive radios,” IEEE Trans. Veh. Technol., vol. 68, no. 4, pp. 4074–4077, Apr. 2019.

[15] Y. Zeng, M. Zhang, F. Han, Y. Gong, and J. Zhang, “Spectrum analysis and convolutional neural network for automatic modulation recognition,” IEEE Wireless Commun. Lett., vol. 8, no. 3, pp. 929–932, Jun. 2019.

[16] T. J. O’Shea, J. Corgan, and T. C. Clancy, “Convolutional radio modulation recognition networks,” in Proc. Int. Conf. Eng. Appl. Neural Netw. Cham, Switzerland: Springer, 2016, pp. 213–226.

[17] N. E. West and T. O’Shea, “Deep architectures for modulation recognition,” in Proc. IEEE Int. Symp. Dyn. Spectr. Access. Netw. (DySPAN), Mar. 2017, pp. 1–6.

[18] R. Yin, J. Huang, and Z. Fei, “Short-time modulation classification of complex wireless communication signal based on deep neural network,” in Proc. 24th Asia–Pacific Conf. Commun. (APCC), Nov. 2018, pp. 520–524.

[19] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, “Deep learning models for wireless signal classification with distributed low-cost spectrum sensors,” IEEE Trans. Cognit. Commun. Netw., vol. 4, no. 3, pp. 433–445, Sep. 2018.

[20] Y. Lin, Y. Tu, and Z. Dou, “An improved neural network pruning technology for automatic modulation classification in edge devices,” IEEE Trans. Veh. Technol., vol. 69, no. 5, pp. 5703–5706, May 2020.

[21] S. Woo, J. Park, J.-Y. Lee, and I. So Kweon, “CBAM: Convolutional block attention module,” in Proc. Eur. Conf. Comput. Vis. (ECCV), 2018, pp. 3–19.

[22] X. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.

[23] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 618–626.

**ZAN YIN** was born in Ningyang, Tai’an, Shandong, China, in 1985. He received the B.S. and M.S. degrees in automation and electronic engineering from Xidian University, Xi’an, China, in 2007 and 2010, respectively, where he is currently pursuing the Ph.D. degree with the National Laboratory of Radar Signal Processing. From 2012 to 2013, he was a member of the 29th Chinese Antarctic Team with the Great Wall Station, Antarctica, where he is responsible for satellite communication and radio wave observation and research. He is also a Senior Engineer in GNSS interference detection with the China Research Institute of Radiowave Propagation, Qingdao, China. His research interests include deep learning, signal modulation recognition, and interference detection.

**BO CHEN** (Senior Member, IEEE) was born in Zhengzhou, Henan, China, in 1979. He received the B.S., M.S., and Ph.D. degrees in electronic engineering from Xidian University, Xi’an, China, in 2003, 2006, and 2008, respectively. From 2008 to 2012, he was a Postdoctoral Fellow, a Research Scientist, and a Senior Research Scientist with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA. Since 2013, he has been a Professor with the National Laboratory for Radar Signal Processing, Xidian University. His current research interests include statistical machine learning, statistical signal processing, and radar automatic target detection and recognition. He received the Honorable Mention for 2010 National Excellent Doctoral Dissertation Award and is selected into One Thousand Young Talent Program, in 2014.

**WEIMIN ZHEN** was born in Hebei, China, in 1963. He received the B.S. degree in radio physics from Peking University, Beijing, China, in 1983, and the M.S. degree in radio wave propagation from the China Research Institute of Radiowave Propagation (CIRIRP), Qingdao, China, in 1990. He is currently a Research Fellow and Senior Engineer of CIRIRP. He is also a Doctoral Preceptor of Wuhan University and an Enterprise Tutor of Xidian University. He has also joined the scientific research in the Antarctic and Arctic area for ionospheric scintillation observation and research in upper atmosphere physics. He has participated in many important national and international projects. His research interests include electromagnetic environment, ionospheric radiowave propagation, and satellite navigation.

**CHAOJIE WANG** was born in Shaoxing, Zhejiang, China, in 1994. He received the B.S. degree in electronic engineering from Xidian University, Xi’an, China, in 2016, where he is currently pursuing the Ph.D. degree in signal processing. His research interests include statistical machine learning and its combinations with real world applications, including multimodal learning, natural language processing, and knowledge graphs.

**TING ZHANG** was born in Sichuan, China, in 1993. She received the B.S. degree in network engineering from the China University of Mining and Technology, in 2015, and the M.S. degree in circuits and systems from Xidian University, in 2018. From July 2018 to May 2020, she was an AI Algorithm Engineer with Vivo Company, working on reinforcement learning, model compression, and user portrait. Since June 2020, she has been working as an Assistant Researcher with the Integrated Research Institute of Information Sensing and Understanding, Xidian University. Her research interests include computer vision, machine learning, deep learning, and reinforcement learning.

***