UAV Communication Signal Recognition: A New Feature Representation and Deep-Learning Method

Lin LI, Zhiyuan DONG, Xiaorui YU, Zhiyuan REN, Zhigang ZHU, Li JIANG

School of Electronic Engineering, Xidian University, 710071 Xi'an, China

lilin@xidian.edu.cn, zgzhu@xidian.edu.cn, yolanda_jiangli@163.com

Submitted April 16, 2021 / Accepted August 19, 2021

Abstract. *As the threats from unmanned aerial vehicles (UAVs) increases gradually, to recognize and classify unknown UAVs have became more and more important in both civil and military security fields. Classification of signal modulation types is one of the basic techniques for specific UAV recognition. In this paper, to represent the hidden features involved in the transmitted signals from UAVs, we propose a two-dimensional squeezing transform (TDST) to characterize the UAV communication signals in a compressed time-frequency plane. The new time-frequency representation, TDST, retains the instantaneous characteristics of the UAV signal, and is with excellent data reduction and noise suppression capabilities. The TDST plane of different modulation types are then considered as input features, and we propose a convolutional neural network (CNN) based on deep-learning to recognize the UAV signals. We design an interception system and consider 10 types of UAV signals with random initial phase, bandwidth and frequency offset. Experimental results demonstrate the effectiveness and superiority of the proposed algorithm.*

Keywords

Automatic modulation classification, unmanned aerial vehicles, squeezing transform, convolutional neural network

1. Introduction

Unmanned aerial vehicles (UAVs), which are commonly called "drones", are growing rapidly in the worldwide. With their inherent advantages, such as mobility, flexibility, adaptive cruise, etc., UAVs have been applied successfully in many fields of civil and military [1–6]. However, as what we have witnessed, UAVs can deliberately or accidently violate the social or national security. For instance, they can act as a carrier for transfering explosive payloads, or violating the periphery of security sensitive areas. To deal with these threats, there is a rapid demand for the technologies that can timely detect and recognize unknown UAVs. Identifying the modulation type of the intercepted communication signal is of great significance for signal demodulation and specific UAV recognition. The current works on automatic modulation classification (AMC) mainly focus on feature extraction. Such methods extract the effective representation of the signal and then perform modulation recognition. For example, A new AMC scheme is proposed using Extreme Learning Machine (ELM) as a classifier, which use the Local Binary Pattern (LBP) to extract the histogram features [7]. With the rapid development of deep learning in recent years, automatic modulation recognition have been continuously attracting wide attention. In [8], an automatic modulation recognition framework is proposed for detecting radio signals in communication systems. This framework combines in-phase, quadrature and fourth-order statistics of the modulation signal by deep convolutional neural networks (CNN) and long short-term memory networks (LSTM). Additionally, more deep-learning architectures are utilized to achieve automatic recognition of communication signals [9]. Especially, a CNN approach is introduced for a robust AMC that classifies the received signals under heavy noises [10]. O'Shea et al. [11] considered a rigorous baseline method using higher order moments and strong boosted gradient tree classification. In [12], the authors developed several methods to represent the modulated signal with CNN grid topology, including the gray image and the enhanced gray image of constellation diagram. After this, some deep models, such as AlexNet and GoogLeNet, are used to achieve effective modulation recognition. Based on multi-domain features and fusion strategies, [13] proposes a new multi-branch asymmetric convolution squeezing and excitation (ACSE) network to achieve better performance under low Signal-to-noise ratio (SNR).

In this paper, we combine the squeezed time-frequency analysis and the deep learning methods to propose an automatic modulation classification framework for UAV signal classification. First, we propose a new time-frequency representation with a two-dimensional squeezing transform, called TDST. Then, motivated by Visual Geometry Group Network-16 (VGG-16), we propose a CNN model suitable for UAV modulation signals, thereby identifying UAV signals effectively. We also give the size change process when the time-frequency spectrum passes through the proposed network. Finally, we compare the recognition accuracy of our proposed method with the other two methods under various SNRs. Experimental results demonstrate that our method achieves superior performance.

2. Proposed Method

2.1 Interception of UAV Communication Signals

Like most of existing communication signal reconnaissance systems, we build a UAV signal interception and recognition system as shown in Fig. 1.

First, the unknown UAV signals are received by an omni-directional antenna with frequency ranging from 400 MHz to 8 GHz, which are transmitted to the receiver. Then through the sweep-frequency receiver, the received signals are transformed to zero intermediate frequency signals with the down converter. Subsequently, the zero intermediate frequency signals are sampled by a 50 MHz analog-to-digital converter (ADC) and transmitted to the graphics processing unit (GPU) system by the peripheral component interconnect express (PCI-E) interface. The GPU further processes the sampling results to obtain signal feature and various parameters, and tests the modulation classification method. Finally, the obtained signal parameters and classification results are transmitted to the upper computer for display. It should be noted that, to evaluate the classification performance, we will use a RF signal generator to simulate the received UAV signals with 10 types of modulation signals and transmit them to the receiver directly.

2.2 Two-dimensional Squeezing Transform

The TDST, in essence, is the squeezing transform of continuous wavelet transform (CWT) in the direction of time and frequency respectively. Different from the synchrosqueezing transform (SST) methods [14–17], which focus on component recovery, the TDST obtains a clear time-frequency representation by suppressing noise and preserving the instantaneous characteristics of the signal. These lay the groundwork for the subsequent classification process.

A function $\psi(t) \in L_2(\mathbb{R})$ is called a continuous wavelet (or an admissible wavelet) if it satisfies the admissible condition:

$$0 < C_\psi = \int_{-\infty}^{\infty} \left|\widehat{\psi}(\xi)\right|^2 \frac{\mathrm{d}\xi}{|\xi|} < \infty \tag{1}$$

where ξ is the frequency, $\widehat{\psi}(\xi)$ is the Fourier transform of $\psi(t)$, defined by

$$\widehat{\psi}(\xi) = \int_{-\infty}^{\infty} \psi(t)\exp(-\mathrm{j}2\pi\xi t)\mathrm{d}t. \tag{2}$$

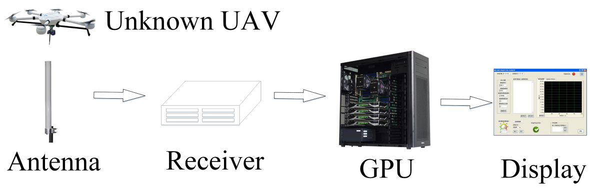

Fig. 1. Interception system of UAV signals.

Denote $\psi_{a,b}(t) = \frac{1}{a}\psi\left(\frac{t-b}{a}\right)$. The continuous wavelet transform (CWT) of a signal $x(t) \in L_2(\mathbb{R})$ with a continuous wavelet ψ is defined by

$$W_x(a,b) = \langle x, \psi_{a,b}\rangle = \int_{-\infty}^{\infty} x(t)\frac{1}{a}\overline{\psi\left(\frac{t-b}{a}\right)}\mathrm{d}t. \tag{3}$$

The variables a and b are called the scale and time variables respectively.

In this paper, we consider the continuous wavelets:

$$\psi_\sigma(t) = \frac{1}{\sigma}\overline{g\left(\frac{t}{\sigma}\right)}\exp(\mathrm{j}2\pi\mu t) - \frac{1}{\sigma}\overline{g\left(\frac{t}{\sigma}\right)}c_\sigma(\mu) \tag{4}$$

where $\sigma > 0$ is the window width of wavelet $\psi_\sigma(t)$, $\mu > 0$, g is a function in $L_2(\mathbb{R})$ with certain decaying order as $t \to \infty$, and $c_\sigma(\mu)$ is a constant such that $\widehat{\psi}_\sigma(0) = 0$. If $\widehat{g}(0) \neq 0$, we let $c_\sigma(\mu) = \overline{\widehat{g}(\sigma\mu)}/\overline{\widehat{g}(0)}$. For example, if $g(t)$ is given by

$$g(t) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right), \tag{5}$$

then ψ_σ is Morlet's wavelet.

The choice of the parameter σ for the wavelet ψ_σ affects the representation of the CWT. In this paper, we propose an ensemble CWT to suppress noises, defined by

$$\widetilde{W}_x(a,b) = \int_{I_\sigma}\int_{-\infty}^{\infty} x(t)\frac{1}{a}\overline{\psi_\sigma\left(\frac{t-b}{a}\right)}\mathrm{d}t\mathrm{d}\sigma \tag{6}$$

where $I_\sigma = [\sigma_0, \sigma_1] \subset \mathbb{R}_+$ is a given interval of σ. The values of σ are different for different modulation signal parameters. In order to adapt to more modulation signal types, we take the average in the simulation.

Then we define the new phase transform

$$\widetilde{\omega}_x(a,b) = \mathrm{Re}\left\{\frac{\partial_b\left(\widetilde{W}_x(a,b)\right)}{\mathrm{j}2\pi\widetilde{W}_x(a,b)}\right\} \tag{7}$$

where $\widetilde{W}_x(a,b) \neq 0$. The modified SST is defined by

$$T_x^{\mathrm{m}}(\xi,b) = \int_{\{\widetilde{W}_x(a,b)\neq 0\}} \widetilde{W}_x(a,b)\delta\big(\widetilde{\omega}_x(a,b) - \xi\big)\frac{\mathrm{d}a}{a} \tag{8}$$

where ξ is the frequency variable.

For non-stationary and frequency modulation signals, the second-order SST is introduced to improve the concentration furthermore. However, the UAV signals are usually stationary locally and suitable to use the first order SST as defined by (8). Take a simple 2 Phase Shift Keying (2PSK) signal for example,

$$s(t) = \exp(\mathrm{j}2\pi\omega_0 t) g\left(\frac{t}{T_g}\right) - \exp(\mathrm{j}2\pi\omega_0 t) g\left(\frac{t}{T_g} - 1\right) \quad (9)$$

where $g(t)$ is gate function with $g(t) = 1$ only for $t \in [0, 1]$, T_g is the width of $g(t)$.

Then the modified SST T_s^{m} of $s(t)$ is extremely sparse in view of $\sigma \ll T_g$. In other words, $|T_s^{\mathrm{m}}(\xi, b)|$ is constant except for $b \in [T_g - \triangle_{\psi_{\sigma_1}}, T_g + \triangle_{\psi_{\sigma_1}}]$, where $\triangle_{\psi_{\sigma_1}}$ is the duration of ψ_{σ_1}.

We propose a squeezing operator along the time axis b of $T_x^{\mathrm{m}}(\xi, b)$ as follows

$$\mathcal{T}(a, k\triangle_h) = \frac{1}{\hbar} \int_{\mathbb{R}} \frac{1}{\lambda} \widetilde{T}_x^{\mathrm{m}}(\xi, b) h\left(\frac{b - k\triangle_h}{\lambda}\right) \mathrm{d}b \quad (10)$$

where h is an admissible window with $\hbar = \int_{\mathbb{R}} h(t)\mathrm{d}t$ and duration $\triangle_h$. In order to discretize (10) to compute the features, the scale a is discretized as $a_j = 2^{j}_{n_\mathrm{v}}\triangle t$, where n_v is the number of voice.

Observe that (10) is a low-pass filter capitalized on the local stationary of the UAV signals. Then $\mathcal{T}(a, k\triangle_h)$ in (10) is called the two-dimensional squeezing transform.

Remark: the TDST algorithm is essentially a compression transformation of continuous wavelet transform both in the time and frequency directions. This new feature representation can effectively mine the hidden features in UAV transmitted signals. The experimental results to be presented in Sec. 3 also show that the combination of wavelet transform and squeezing transform can obtain a clear feature representation of the signal, which is conducive to subsequent classification.

2.3 The Proposed CNN Model

Deep learning (DL), as a novel development method, can extract more meaningful features through its hierarchical learning process. DL-based methods can automatically learn distinctive representations of high-dimensional data [18], [19].

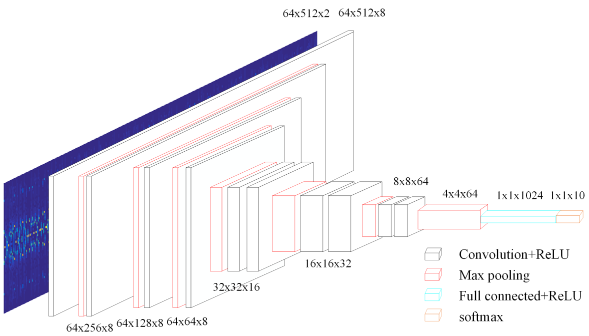

Fig. 2. The schematic diagram of the proposed CNN architecture.

Layer (type)	Output shape	Parameters number
Conv2- Pool2	64, 256, 8	144
Conv2- Pool2	64, 128, 8	576
Conv2- Pool2	64, 64, 8	576
Conv2- Pool2	32, 32, 8	576
Conv2- Conv2 - Pool2	16, 16, 16	2 304
Conv2- Conv2 - Pool2	8, 8, 32	9 216
Conv2- Conv2 - Pool2	4, 4, 64	36 864
Full Connected	1 024	262 144
Full Connected	128	131 072
Softmax	10	1 280

Tab. 1. Simulation parameters.

To adapt to the classification of UAV signals, we build a special CNN architecture which is shown in Fig. 2. The motivation for choosing the CNN is that it can automatically extract the hidden features involved inside of the signal (including the signal feature representation through TDST in this paper). We can see that, the matrix size of the input model after TDST is 64×512×2, consisting of real and imaginary parts. Then we separate the real and imaginary parts of the complex value, and form a dual channel mode to input the CNN architecture. In each channel, the convolutional kernels of CNN perform arithmetic operations with real values. Similar to VGG-16 [20], our CNN structure consists of 19 layers, including convolutional layer, pooling layer, fully connected layer and softmax layer. Table 1 shows the detailed model configuration.

In Fig. 2, the first four convolutional kernels is set to 8. The next 8 convolutional layers are divided into 3 types, each of which has the same 2 convolutional layers. The number of convolution kernels included in each convolutional layer is 16, 32 and 64, respectively. We use a convolution kernel with the size of 3×3 to extract distinctive features. Compared with the 7×7 convolution kernel, the convolution kernels with the size of 3×3 not only increase the CNN depth but also reduce the amount of parameters. We also introduce a pooling kernel with the size of 2×2 to reduce the parameters. To avoid network overfitting, dropout operator with a ratio of 0.5 is introduced to follow the convolutional layer. A sliding average and L2 regularization are introduced to the model. Rectified Linear Unit (ReLU) is selected as an activation function. As mentioned in [9] and [11], our proposed CNN is built on the VGG-16 according to the actual properties of the dataset with two fully connected layers too. Specifically, we set the number of neurons in the last two fully connected layers to 1024 and 128 (both of fully connected layers' dimensions of VGG-16 are 4096) to adapt to the proposed CNN structure, respectively. Finally, the last layer of the network is connected to a softmax classifier to obtain the predicted scores of modulation categories. The loss function and optimizer of the model are the cross entropy error function and the Adam respectively. The cross entropy error function is as follows:

$$loss = -\sum_{i=1}^{n} y_i \log\left(\widehat{y}_i\right) + (1 - y_i) \log\left(1 - \widehat{y}_i\right) \quad (11)$$

where $\widehat{y}_i$ represents the prediction and y_i represents the groundtruth label.

2.4 Benchmarks

Time-Domain based method: As a comparison, this method uses the same structure proposed in Sec. 2.3. The dataset used in this method is the modulated signal. The input of the network is still dual-channel input. Finally, we input the test set to predict the category.

ANN based method: As a comparison, this method uses an artificial neural network (ANN) based on characteristics. The ANN used in this method is a back propagation (BP) artificial neural network. In addition, we introduce multiple signal features as the input of ANN, including signal instantaneous frequency, multiple high-order cumulants, etc.

3. Experiments

3.1 Dataset

In this experiment, we consider to classify 10 common UAV modulation signals, including 2 Frequency-Shift Keying (2FSK), 4FSK, 8FSK, 2PSK, 4PSK, 8PSK, 16 Absolute Phase Shift Keying (16APSK), 16 Quadracture Amplitude Modulation (16QAM), 32QAM and Gaussian-filtered Minimum Shift Keying (GMSK). As mentioned in Sec. 2.1, here we use a RF signal generator to simulate different UAV signals and transmit them to the receiver directly. For each modulation type, 3 200 samples are available, where the SNR varies from –10 dB to 5 dB with 200 samples. The size of each signal sample is 256×256 and it includes real and imaginary parts. The real and imaginary parts of each signal sample are stored in two folders separately. The sampling frequency we use is 50 MHz. The signal bandwidth and frequency offset are 3–6 MHz and 0–200 kHz respectively.

3.2 Experiment Procedure and Model Training

For the modulated signal, we calculate the corresponding TDST. The frequency is discretized to 64 bins and the time axis is squeezed to 512 points. Note that here we just use a constant σ to calculate the TDST. For the TDST of each sample, the input size is 64×256×2. Then, half the samples are selected for training and the rest are for testing. The training set and testing set are selected equally for each SNR and modulation type.

For hyper-parameters, when the initial value of exponential decay learning rate is set to 0.005, the best recognition performance of the method is selected for the final prediction. The batch size in the training is set to 128. These hyper-parameters configuration have been used for the final modulation signal classification. As the two comparison methods mentioned in Sec. 2.4, CNN model based on time-domain dataset and ANN based on features use the same original dataset and experimental conditions. For the built CNN model based on time-domain dataset, the hyper-parameters are the same as them proposed in this paper. The GeForce RTX 2080 Ti GPU is used as the main hardware execution environment for CNN implementation, training and testing. The deep learning framework used in this article is Tensorflow, which uses Python to achieve various functions.

3.3 Modulation Classification Performance

As shown in Fig. 3, we give the convolutional features and TDST results of 2FSK, 2PSK and 16QAM. It can be clearly seen that after TDST, the three kinds of signals contain almost no features in the time-frequency domain and have obvious discrimination. At the same time, the corresponding deep representations are also quite different. They reflect the rationality of the proposed CNN architecture.

Figure 4 shows the classification performance of 10 signals. In general, when the SNR is greater than –5 dB, the accuracies of all modulation types are over 80%. When the SNR is greater than 0 dB, the accuracies exceed 90%. Even if the SNR is –10 dB, the accuracies of 4PSK, 8PSK and GMSK are around 70%. We also observe that the proposed method has high recognition accuracies for FSK signals. For 16QAM and 32QAM that are difficult to distinguish, our method can also keep the accuracies above 90% in the SNR range of –10 dB to 5 dB.

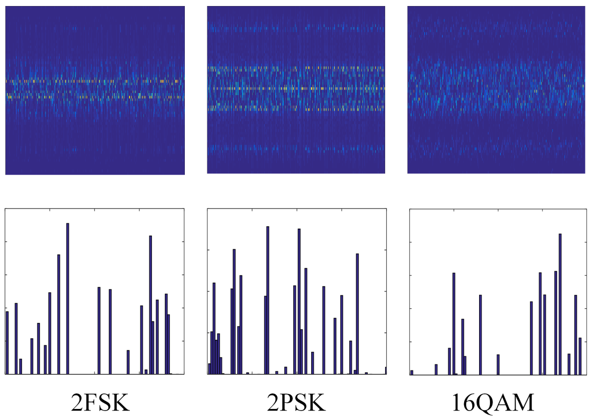

Fig. 3. TDST and corresponding deep representation.

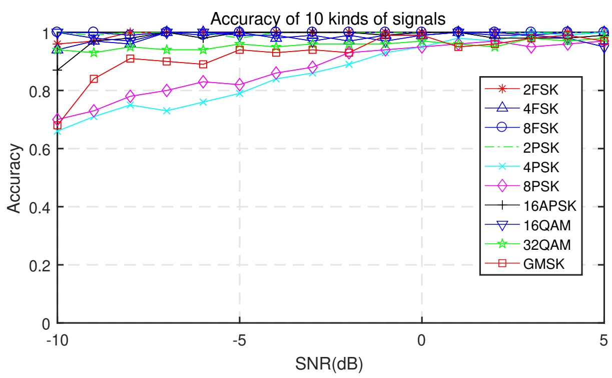

Fig. 4. Accuracy variation of the model with SNRs.

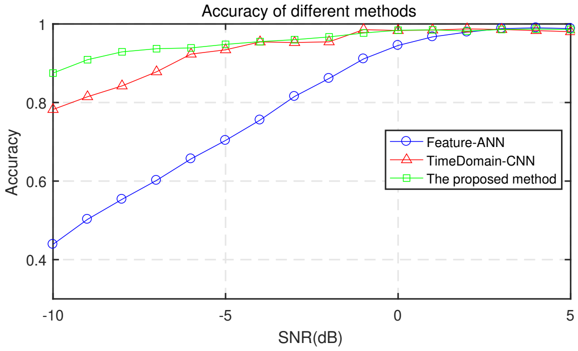

Fig. 5. Average accuracy comparisons of three different methods under different SNRs.

Figure 5 shows the average recognition accuracy variations of models with SNRs. It is observed that the CNN-based methods significantly outperform the traditional handcrafted one. Although the feature-based ANN model shows an overall upward trend, the average recognition rate at low SNR is 10% to 20% lower than that of the other two methods. However, when the SNR is less than –5 dB, the TDST-based deep learning model can maintain a high average accuracy. Instead, the average accuracy of the CNN model based on time-domain information has dropped significantly. It can be seen that our proposed method significantly exceeds the CNN model based on pure time-domain information, which proves the validity of TDST representation.

Based on the above two result figures, the method proposed in this paper provides a better classification performance. Even when the SNR is low, TDST can still guarantee that the average classification accuracy is close to 90%.

4. Conclusions Proposing

This paper proposes a new automatic modulation classification based on time-frequency analysis. We first propose a new two-dimensional compression transform (TDST) to represent UAV communication signals. Then, we build a special CNN architecture for identifying UAV signals. Finally, they are concatenated into a whole architecture for UAV modulation classification. Experimental results demonstrate that compared with the feature-based artificial neural network (ANN) method and the time-domain signal-based CNN method, the proposed method achieves superior recognition performance. Future work will focus on extracting hidden features based on multiple types of deep networks and using the obtained features to jointly identify more types of signals.

Acknowledgments

Thanks to the National Nature Science Foundation of China (Grant No. 62071349 and No. 61803294), Natural Science Foundation of Shaanxi Province, China (Grant No. 2020JQ-684) and the Fundamental Research Funds for the Central Universities (Grant No. XJS210211) for funding.

References

[1] KALEEM, Z., REHMANI, M. H., AHMED, E., et al. Amateur drone surveillance: Applications, architectures, enabling technologies, and public safety issues: Part 1. *IEEE Communications Magazine*, 2018, vol. 56, no. 1, p. 14–15. DOI: 10.1109/MCOM.2018.8255731

[2] VASHIST, S., JAIN, S. Location-aware network of drones for consumer applications: Supporting efficient management between multiple drones. *IEEE Consumer Electronics Magazine*, 2019, vol. 58, no. 3, p. 68–73. DOI: 10.1109/MCE.2019.2892279

[3] MOZAFFARI, M., SAAD, W., BENNIS, M., et al. A tutorial on UAVs for wireless networks: Applications, challenges, and open problems. *IEEE Communications Surveys and Tutorials*, 2019, vol. 21, no. 3, p. 2334–2360. DOI: 10.1109/COMST.2019.2902862

[4] BASAK, S., SCHEERS, B. Passive radio system for realtime drone detection and DOA estimation. In *International Conference on Military Communications and Information Systems (ICMCIS)*. Warsaw (Poland), 2018, p. 1–6. DOI: 10.1109/ICMCIS.2018.8398721

[5] SOLOMITCKII, D., GAPEYENKO, M., SEMKIN, V., et al. Technologies for efficient amateur drone detection in 5G millimeter-wave cellular infrastructure. *IEEE Communications Magazine*, 2018, vol. 56, no. 1, p. 43–50. DOI: 10.1109/MCOM.2017.1700450

[6] LI, W., DOU, Z., QI, L., et al. Wavelet transform based modulation classification for 5G and UAV communication in multipath fading channel. *Physical Communication*, 2019, vol. 34, p. 272–282. DOI: 10.1016/j.phycom.2018.12.019

[7] GUNER, A., ALCIN, O. F., SENGUR, A. Automatic digital modulation classification using extreme learning machine with local binary pattern histogram features. *Measurement*, 2019, vol. 145, p. 214–225. DOI: 10.1016/j.measurement.2019.05.061

[8] ZHANG, M., ZENG, Y., HAN, Z., et al. Automatic modulation recognition using deep learning architectures. In *IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. Kalamata (Greece), 2018, p. 1–5. DOI: 10.1109/SPAWC.2018.8446021

[9] RAMJEE, S., JU, S., YANG, D., et al. Fast deep learning for automatic modulation classification. *arXiv:1901.05850*, 2019, p. 1–29.

[10] ZHOU, S., YIN, Z., WU, Z., et al. A robust modulation classification method using convolutional neural networks. *EURASIP Journal on Advances in Signal Processing*, 2019, vol. 21, p. 1–16. DOI: 10.1186/s13634-019-0616-6

[11] O'SHEA, T. J., ROY, T., CLANCY, T. C. Over-the-air deep learning based radio signal classification. *IEEE Journal of Selected Topics in Signal Processing*, 2018, vol. 12, no. 1, p. 168–179. DOI: 10.1109/JSTSP.2018.2797022

[12] PENG, S., JIANG, H., WANG, H., et al. Modulation classification based on signal constellation diagrams and deep learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, vol. 30, no. 3, p. 718–727. DOI: 10.1109/TNNLS.2018.2850703

[13] WEI, S., QU, Q., WANG, M., et al. Automatic modulation recognition for radar signals via multi-branch ACSE networks. *IEEE Access*, 2020, vol. 8, p. 94923–94935. DOI: 10.1109/ACCESS.2020.2995203

[14] DAUBECHIES, I., LU, J., WU, H. T. Synchrosqueezed wavelet transforms: an empirical mode decompositionlike tool. *Applied and Computational Harmonic Analysis*, 2011, vol. 30, no. 2, p. 243–261. DOI: 10.1016/j.acha.2010.08.002

[15] MEIGNEN, S., OBERLIN, T., MCLAUGHLIN, S. A new algorithm for multicomponent signals analysis based on synchrosqueezing: With an application to signal sampling and denoising. *IEEE Transactions on Signal Processing*, 2012, vol. 60, no. 11, p. 5787–5798. DOI: 10.1109/TSP.2012.2212891

[16] AUGER, F., FLANDRIN, P., LIN, Y. Time-frequency reassignment and synchrosqueezing: An overview. *IEEE Signal Processing Magazine*, 2013, vol. 30, no. 6, p. 32–41. DOI: 10.1109/MSP.2013.2265316

[17] LI, L., CAI, H., HAN, H., et al. Adaptive short-time Fourier transform and synchrosqueezing transform for non-stationary signal separation. *Signal Processing*, 2020, vol. 166, p. 107231.1–107231.15. DOI: 10.1016/j.sigpro.2019.07.024

[18] SAINATH, T. N., VINYALS, O., SENIOR, A., et al. Convolutional, long short-term Memory, fully connected deep neural networks. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. South Brisbane (QLD, Australia), 2015, p. 4580–4584. DOI: 10.1109/ICASSP.2015.7178838

[19] SEPAS-MOGHADDAM, A., ETEMAD, A., PEREIRA F., et al. Facial emotion recognition using light field images with deep attention-based bidirectional LSTM. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Barcelona (Spain), 2020, p. 3367–3371. DOI: 10.1109/ICASSP40776.2020.9053919

[20] SIMONYAN, K., ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*, 2014.

About the Authors . . .

Lin LI was born in Xianyang, China. He is an Associate Professor in the School of Electronic Engineering, Xidian University. His research interests include non-stationary signal processing, signal spectrum estimation, adaptive signal separation, time-frequency analysis, signal reconstruction, passive radio signal detection and recognition.

Zhiyuan DONG was born in Luoyang, China. He is currently working towards the M.S. degree in the School of Electronic Engineering, Xidian University. His research interests include signal detection and recognition, automatic modulation classification and deep learning.

Xiaorui YU was born in Yantai, China. She is currently working towards the M.S. degree in the School of Electronic Engineering, Xidian University. Her research interests include remote sensing signal denoising, non-stationary signal time-frequency analysis and signal reconstruction.

Zhiyuan REN was born in Yantai, China. He is currently working towards the M.S. degree in the School of Electronic Engineering, Xidian University. His research interests include machine learning, modulation signal feature fusion and recognition.

Zhigang ZHU (corresponding author) was born in Yantai, China. He received his Ph.D. degree from Pattern Recognition and Intelligent System, Xidian University in 2020. His research interests include pattern recognition, deep learning and radar emitter recognition.

Li JIANG was born in Shangluo, China. She received her Ph.D. degree from Signal and Information Processing, Xidian University in 2016. Her research interests include signal detection and estimation, mechanical signal processing, seismic signal processing, etc.