Specific Emitter Identification via Convolutional Neural Networks

Lida Ding, Shilian Wang, Fanggang Wang, *Senior Member, IEEE*, and Wei Zhang

***Abstract*— Specific emitter identification (SEI) is a technique that distinguishes between unique emitters using the external feature measurements from their transmit signals, primarily radio frequency fingerprints. The SEI has been widely adopted for military and civilian spectrum management applications. We propose a deep-learning-based SEI approach that uses the features of the received steady-state signals. In particular, the bispectrum of the received signal is calculated as a unique feature. Then, we use a supervised dimensionality reduction method to significantly reduce the dimensions of the bispectrum. Finally, a convolutional neural network is adopted to identify specific emitters using the compressed bispectrum. This approach essentially extracts overall feature information hidden in the original signals, which can then be used to improve identification performance. Results from both the simulations and the software radio experiments are provided. A signal acquisition system is designed to collect steady-state signals from multiple universal software radio peripherals. Both the simulations and the experiments validate our conclusion that the proposed approach outperforms other existing schemes in the literature.**

***Index Terms*— Bispectrum, convolutional neural network, deep learning, specific emitter identification, USRP.**

I. Introduction

SPECIFIC emitter identification (SEI) is a process of discriminating individual emitters by comparing radio frequency fingerprints (RFFs) in the received signal that originated from the unique hardware imperfections. These are also known as transmitter impairments inherent to the analog components in transmitters, which are independent of information content. SEI has many applications in both military and civilian scenarios [1]–[3]. Recently, it has become increasingly important in cognitive radios and self-organized networking [4]–[7]. The RFFs of individual emitters characterize different impairments in the transmitters and are difficult to forge, which makes SEI promising [8].

Manuscript received June 11, 2018; revised August 8, 2018; accepted September 6, 2018. Date of publication September 20, 2018; date of current version December 10, 2018. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 2018JBM078, in part by the National Natural Science Foundation under Grants 61571034 and 61725101, in part by the Beijing Natural Science Foundation under Grants 4182051 and L172020, in part by the State Key Laboratory of Rail Traffic Control and Safety under Grant RCS2018ZT016, in part by the Key Laboratory of Universal Wireless Communications (BUPT), Ministry of Education, P. R. China, under Grant KFKT-2018102, in part by the Major Projects of Beijing Municipal Science and Technology Commission under Grant Z181100003218010, in part by National Key Research and Development Program under Grant 2016YFE0200900 and in part by Nokia. The associate editor coordinating the review of this letter and approving it for publication was T. Ngatched. *(Corresponding authors: Shilian Wang; Fanggang Wang.)*

L. Ding, S. Wang, and W. Zhang are with the School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China (e-mail: wangsl@nudt.edu.cn).

F. Wang is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China (e-mail: wangfg@bjtu.edu.cn).

Digital Object Identifier 10.1109/LCOMM.2018.2871465

In the literature, SEI techniques can be categorized into two classes, i.e., approaches using transient signals and approaches using steady-state signals. Transient-based approaches generally offer good identification performance and extract the RFF features by detecting received transient signals [9], [10]. However, transient signals are difficult to capture, especially in noncooperative communications circumstances. Accurate and consistent detection of the start and end points of the transient signals is critical for correct identification. However, in contrast to the approaches using transient signals, the use of steady-state signals provides more statistically stable RFF features. They provide an alternative perspective in which the received signals are transformed into a new signal domain and the unique and stable fingerprint features can be extracted for identification. Higher-order spectrums and time-frequency transforms are widely used, such as integral bispectrum, wavelet, and Hilbert-Huang transform (HHT) [11]–[14]. However, wavelet-based approaches heavily rely on the selection of basic wavelet functions. Bispectrum-based methods and HHT-based methods both have problems with generally high feature dimensions. Some methods focus on extracting low-dimensional features from the high-dimensional ones. For example, three HHT-based methods that use the steady-state signals have been proposed in [13]. Local integral bispectrum with Fisher discriminants have also proved to be useful in [11]. However, these methods merely utilize certain dimension-reduced features. In contrast, deep learning could be a potential approach for extracting universal features from the received signals in order to further improve identification performance.

Deep learning has recently been shown to be effective in various applications, particularly for identification. Multiple hidden layers with non-linear logical functions enable the learning of higher-level information hidden in the data. This letter proposes a novel deep-learning based method for specific emitter identification. First, the bispectrum dimension of the received signals is reduced by a supervised dimensionality reduction method in [15], reducing the influence of redundant information for identification. Then, a convolutional neural network is built for learning in order to classify the compressed bispectrum. This approach is shown to be more effective than conventional bispectrum-based, HHT-based, nonlinear based and constellation-based methods, especially in some challenging situations such as a case with multiple emitters with low signal to noise ratio (SNR). Since the main computational load of the proposed scheme is developed for offline training, the identification process is computationally efficient with parallel computing of neural networks, once the model is well trained. Furthermore, the performance validation of the RFF approaches occurs mainly via simulations in past literature. This makes it difficult to reproduce because the identification performance is closely related to the RFF modeling of the emitters and the channel mod-

els adopted in the simulation. In our evaluation, we use simulations and experiments with real measurements from the USRPs.

II. System Model

In general, multiple RFFs are involved in different processes when generating transmit signals, among which the power amplifying (PA) process is the main contributor. The inherent nonlinearity and memory effects of power amplifier induce severe nonlinear distortion on the transmit signals, which can serve as an RFF.

A memory polynomial model is adopted that consists of multiple delays and nonlinear functions [16], which is expressed as

$$y(n) = \sum_{q=0}^{Q} \sum_{l=1}^{L} h_q p_{2l-1} |x(n-q)|^{2(l-1)} x(n-q) \tag{1}$$

where $y(\cdot)$ is the output signal; h_q, p_{2l-1}, Q, and L are the memory coefficients, nonlinear coefficients, memory depth and order of the polynomial, respectively; and $x(\cdot)$ is the input signal to the power amplifier, which can be written as

$$x(n) = s(n) e^{j2\pi n f T} \tag{2}$$

with $s(\cdot)$ as the baseband modulated signal, T as the sampling period, and f as the carrier frequency. This model is a modification of the discrete Volterra series. By truncating the Volterra series, the number of parameters is reduced.

In an additive white gaussian noise (AWGN) channel, the discrete-time received signal at time n can be expressed as

$$r(n) = y(n) + c(n) \tag{3}$$

where $c(n)$ is the AWGN. At the receiver, the RFF features extracted from $r(n)$ are used to identify different emitters.

III. CNN-Based SEI Approach

We propose a CNN-based SEI approach using the compressed bispectrum as the feature. The main procedures are as follows: The bispectrum of the received signals is estimated first, and its dimension is then reduced by a projection method in [15]. After that, the CNN is adopted to identify the emitters by their compressed bispectrum.

A. Estimation of Bispectrum

The bispectrum of the received signal $r(n)$ is estimated using a nonparametric method. First $r(n)$ is divided into Γ segments, each containing Δ observation samples. The third-order cumulant of $r(n)$ can be estimated by

$$\hat{c}_{3r}(\tau_1, \tau_2) = \frac{1}{\Gamma} \sum_{i=1}^{\Gamma} \chi^{(\gamma)}(\tau_1, \tau_2) \tag{4}$$

where $\chi^{(\gamma)}(\tau_1, \tau_2)$ is the third-order cumulant estimation of each signal segment. Then, the estimated bispectrum of $r(n)$ can be expressed as

$$\hat{B}(\omega_1, \omega_2) = \sum_{\tau_1=-\delta}^{\delta} \sum_{\tau_2=-\delta}^{\delta} \hat{c}_{3r}(\tau_1, \tau_2) \omega(\tau_1, \tau_2) e^{-j(\omega_1 \tau_1 + \omega_2 \tau_1)} \tag{5}$$

where $\delta < \Delta - 1$, and $\omega(\tau_1, \tau_2)$ is the hexagonal window function.

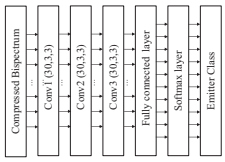

Fig. 1. Convolutional neural network structure for SEI.

B. Dimension Reduction of Bispectrum

The objective function of bispectrum dimension reduction can defined as (6) referring to [15]

$$\min_{\boldsymbol{U}_1, \ldots, \boldsymbol{U}_M} \frac{\sum_{i,j} \|\boldsymbol{U}_i - \boldsymbol{U}_j\|^2 s_{ij}}{\sum_{i,j} \|\boldsymbol{U}_i - \boldsymbol{U}_j\|^2 d_{ij}} \tag{6}$$

where the set $\{\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_M\}$ forms the compressed bispectrum projection space of the original selected bispectrum set $\{\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_M\}$. $\boldsymbol{S}$ and $\boldsymbol{D}$ are the weight matrix of the bispectrum, from the same or different emitter, respectively. If $\boldsymbol{V}_i$, $\boldsymbol{V}_j$ are from different emitters, then $s_{ij} = 0$, otherwise

$$s_{ij} = \frac{1}{\varepsilon} e^{-\|\boldsymbol{V}_i - \boldsymbol{V}_j\|^2}. \tag{7}$$

If $\boldsymbol{V}_i$, $\boldsymbol{V}_j$ are from the same emitter, then $d_{ij} = 0$, otherwise

$$d_{ij} = \frac{1}{\varepsilon} e^{-\|\boldsymbol{V}_i - \boldsymbol{V}_j\|^2} \tag{8}$$

where the parameter ε in (7) and (8) is a positive real number. The objective function in (6) can be converted into

$$\min_{\boldsymbol{w}} \frac{\boldsymbol{w}^{\mathrm{T}} \boldsymbol{V} (\boldsymbol{\Phi}_1 \otimes \boldsymbol{I}_n) \boldsymbol{V}^{\mathrm{T}} \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{T}} \boldsymbol{V} (\boldsymbol{\Phi}_2 \otimes \boldsymbol{I}_n) \boldsymbol{V}^{\mathrm{T}} \boldsymbol{w}} \tag{9}$$

where $\boldsymbol{\Phi}_1 = \boldsymbol{O} - \boldsymbol{S}$, $\boldsymbol{\Phi}_2 = \boldsymbol{F} - \boldsymbol{D}$, $o_{ii} = \sum_j s_{ij}$, $f_{ii} = \sum_j d_{ij}$; $\boldsymbol{V} = [\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_M]$ consists of the matrices of the selected bispectrum from all emitters. $\otimes$ is the Kronecker product.

Let $\boldsymbol{w}^{\mathrm{T}} \boldsymbol{V} (\boldsymbol{\Phi}_2 \otimes \boldsymbol{I}_n) \boldsymbol{V}^{\mathrm{T}} \boldsymbol{w} = \theta$, where θ is a nonzero constant, then the Lagrangian function can be defined as

$$G(\boldsymbol{w}, \lambda) = \boldsymbol{w}^{\mathrm{T}} \boldsymbol{V} (\boldsymbol{\Phi}_1 \otimes \boldsymbol{I}_n) \boldsymbol{V}^{\mathrm{T}} \boldsymbol{w} + \lambda (u - \boldsymbol{w}^{\mathrm{T}} \boldsymbol{V} (\boldsymbol{\Phi}_2 \otimes \boldsymbol{I}_n) \boldsymbol{V}^{\mathrm{T}} \boldsymbol{w}). \tag{10}$$

The solution to minimize (9) occurs when $\frac{\partial G(\boldsymbol{w}, \lambda)}{\partial \boldsymbol{w}} = 0$. Therefore, the optimal solution in (10) can be calculated, which is the eigenvector corresponding to the minimum eigenvalue of the following equation:

$$\boldsymbol{V} (\boldsymbol{\Phi}_1 \otimes \boldsymbol{I}_n) \boldsymbol{V}^{\mathrm{T}} \boldsymbol{w} = \lambda \boldsymbol{V} (\boldsymbol{\Phi}_2 \otimes \boldsymbol{I}_n) \boldsymbol{V}^{\mathrm{T}} \boldsymbol{w}. \tag{11}$$

Let $\lambda_1 < \lambda_2 < \cdots < \lambda_d$ be the d eigenvalues, and the corresponding eigenvectors are $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_d$, respectively. Then, we can calculate the projection matrix $\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_d]$. The compressed bispectrum $\boldsymbol{U}_i$ can be calculated as

$$\boldsymbol{U}_i = \boldsymbol{V}_i \boldsymbol{W}. \tag{12}$$

C. The CNN-Based Approach

The compressed bispectrum $\bar{\boldsymbol{B}}$ is fed into a CNN model to learn to identify each emitter. The CNN structure for SEI is shown in Figure 1, which contains three convolutional layers and a single dense layer before the softmax classifier. Each hidden layer has a ReLU activation function and dropout rate of 20% [17]. The filter size of every convolutional layer is

Algorithm 1 The CNN-Based SEI Approach

Training Procedure: Let $\boldsymbol{B}_i$ be the bispectrum of the ith training signal sequence, where $i = 1, .., N_{\rm t}$, with $N_{\rm t}$ being the total number of the training sequences from the K emitters.

1: Estimate $\boldsymbol{B}_i$ using (4) and (5), where $\Delta = 125$, and the two-dimensional fast Fourier transform size is 128;
2: Select N_d bispectrum from the training sequences of each emitter, then get the $\{\boldsymbol{B}_{k_0,i_0}\} = \{\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_{K*N_d}\}$, where $k_0 = 1, .., K$, $i_0 = 1, .., N_d$, and $N_d \geq 20$. Then, calculate the projection matrix $\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{10}]$ using (6)-(11), and $\varepsilon = 1$ in (6)-(7)[1];
3: Compute the compressed bispectrum $\bar{\boldsymbol{B}}_i$ using (12); [2]
4: Let $\{\bar{\boldsymbol{B}}_i, \ell_k\}$ be the set of training data, where $\ell_k \in \{1, .., K\}$ is the emitter class label. Then, start training the CNN model with the data, i.e., to obtain and save the optimal network weights.

Identification Procedure: Let $\boldsymbol{B}_j$ be the bispectrum of the jth testing sequence, where $j = 1, .., N_{\rm s}$, with $N_{\rm s}$ being the total number of the testing sequences of the K emitters.

1: Estimate $\boldsymbol{B}_j$ using (4) and (5) like the training procedure;
2: Compute the compressed bispectrum $\bar{\boldsymbol{B}}_j$ using (12), where the projection matrix $\boldsymbol{W}$ is the one calculated in the training procedure[2];
3: Utilize the trained CNN to identify the testing sequence.

3×3, which is commonly used in CNN models. The filter number of each layer is 30. In practice, the layer number and filter number are selected through simulations and experiments. To simplify the process of parameter selection, we assume that every layer has the same number of filters. The outputs of the last convolutional layer are flattened and then they are fed into the fully connected layer, which has 128 neurons. The final dense softmax layer maps the features into one output emitter class. The cross-entropy loss function is adopted to measure the error probability in discrete identification tasks. We use the Adam optimizer [18], i.e., a first-order gradient-based optimizer with a learning rate of 0.001.

IV. Performance Evaluation

In this section, we evaluate the identification performance of the proposed algorithm, compared with the square integral bispectrum (SIB) algorithm [11], the HHT-based algorithm [13], the nonlinear-based method [19] and the constellation-based method [20].

A. Transmit Signal Generation

A signal acquisition system in Figure 2 is designed to collect signals from five USRPs including one N210 (UBX-40), one E310 and three B210. The radio frequency (RF) is 800 MHz, and the signal from each USRP is sampled at a rate

[1] $\boldsymbol{B}_{k_0,i_0}$ is the bispectrum of selected training sequence i_0, which is from emitter k_0. The lager N_d is, the heavier computation burden of computing the projection matrix $\boldsymbol{W}$ is, which is not necessary. In the simulations and experiments of this letter, $N_d = 20$. The number of eigenvectors in $\boldsymbol{W}$ is ascertained by experiments. In general, first ten eigenvectors can contain most of the information useful for identification.

[2] $\bar{\boldsymbol{B}}_i$ and $\bar{\boldsymbol{B}}_j$ are the compressed bispectrum of the ith training sequence and the jth testing sequence, respectively.

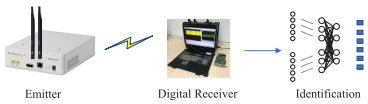

Fig. 2. Signal acquisition system for SEI.

TABLE I
POWER AMPLIFIER COEFFICIENTS

Emitter	p_1	p_3	p_5	h_0	h_1	h_2
Λ1	1.00	-0.36	-0.036	1.00	-0.36	-0.036
Λ2	1.00	-0.27	-0.027	1.00	-0.27	-0.027
Λ3	1.00	-0.18	-0.018	1.00	-0.18	-0.018
Λ4	1.00	-0.09	-0.009	1.00	-0.09	-0.009
Λ5	1.00	0.00	0.000	1.00	0.00	0.000
Λ6	1.00	0.09	0.009	1.00	0.09	0.009
Λ7	1.00	0.18	0.018	1.00	0.18	0.018
Λ8	1.00	0.27	0.027	1.00	0.27	0.027
Λ9	1.00	0.36	0.036	1.00	0.36	0.036
Λ10	1.00	0.45	0.045	1.00	0.45	0.045

TABLE II
AVERAGE IDENTIFICATION ACCURACY WITH DIFFERENT STRUCTURES

#Layers	#Filters	#Emitter:5	#Emitter:10	#USRP:5
1	20	84.77%	81.18%	68.47%
1	30	84.97%	82.17%	66.59%
1	40	84.51%	81.50%	65.75%
2	20	85.50%	83.72%	73.69%
2	30	85.95%	83.99%	72.85%
2	40	85.48%	83.68%	72.96%
3	20	86.69%	85.26%	75.01%
3	30	87.11%	85.56%	75.69%
3	40	86.83%	84.84%	74.91%

of 8 Mbps with the oversampling rate set to 8, ie., the symbol rate is 1 Mbps. The modulation format is quadrature phase shift keying (QPSK) followed by a square-root raised cosine roll-off filter with the factor of 0.35. The RF signals are first converted to the intermediate frequency, (IF) 70 MHz, then collected by a wideband receiver at a sampling rate of 50 MHz. Different batches of data are measured independently with SNR from 0 to 30 dB. The numbers of the training and testing sequences are both 300 from one emitter at each SNR. Each sequence has 1000 sampling points, i.e., 20 symbols.

Meanwhile, we consider ten emitters using the model in (1)-(3) in the simulation. The order of the polynomial and the memory in (1) are $Q = 2, L = 3$. The memory coefficients and the nonlinear coefficients are listed in Table I. The modulation format of the emitter is also QPSK, and the carrier frequency is set to 2 GHz. The duration of each sequence is 100 μs and the sampling rate is 10 GHz, and each observed sequence also contains 1000 sampling points. The numbers of the training and testing sequences are both 300 from one emitter at each SNR, which ranges from 0 to 30 dB. For the case of 5 emitters, we use the labels Λ1, Λ2, ..., Λ5, and when there are 10 emitters, we use the labels Λ1, Λ2, ..., Λ10.

B. Identification Experiments

The average identification accuracy of the CNN models with varying parameters are listed in Table II. It can be seen that the deeper layer significantly improves the identification performance, while the number of filters should be moderate, which fluctuates the identification accuracy within about 1% when it ranges from 20 to 40.

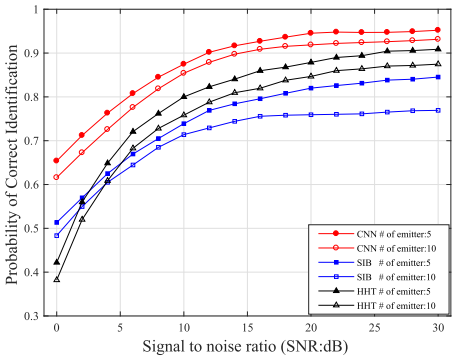

Fig. 3. Identification performance of the simulations.

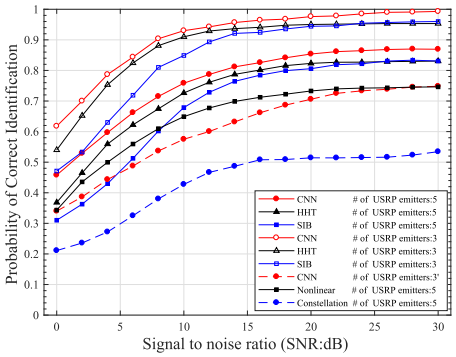

Fig. 4. Identification performance of the experiments.

The simulation and experimental results are shown in Figure 3, 4 and 5. The CNN-based algorithm outperforms other existing schemes in the literature and it achieves a gain of about 3 dB for $P_c \geq 0.8$ or $P_c \geq 0.9$ when identifying five USRPs or three different type of USRPs (labeled 3), respectively. Furthermore, the identification performance is limited by a large confusion among three B210 (labeled 3'). As the length of measurements increases, the estimation accuracy of bispectrum and the identification performance of CNN are improved. The results show that the CNN-based algorithm can identify real emitters using steady-state signals, while the performance requires further improvement when identifying emitters of the same type.

V. Conclusion

In this letter, we proposed a promising SEI approach based on the compressed bispectrum of the received signals using the CNN. The original RFF, such as bispectrum which has high dimensionality, can cause dimension curse and misidentification, while the compressed RFF reduces the influence on identification of redundant information. The CNN learns high-level feature hidden in the data, improving identification performance. Both the simulation results and software radio experiments show that this proposed method outperforms the existing SIB-based and HHT-based methods. This is an instructive example of transferring the special emitter identification problem into an image identification problem in which the deep learning technology can be used efficiently. This method provides a convenient and effective solution for radio spectrum protection and other wireless secure communications.

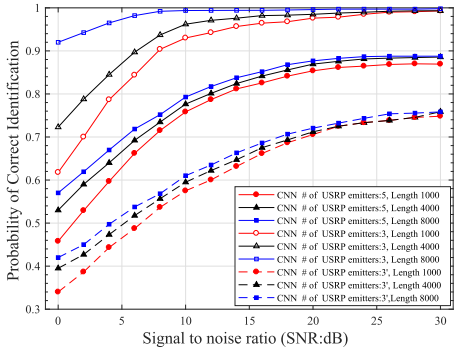

Fig. 5. Identification performance of different lengths of measurements.

References

[1] R. A. Poisel, *Introduction to Communication Electronic Warfare Systems*. Norwood, MA, USA: Artech House, 2002.
[2] B. Danev, H. Luecken, S. Capkun, and K. El Defrawy, "Attacks on physical-layer identification," in *Proc. ACM WiSec*, 2010, pp. 89–98.
[3] Y. A. Eldemerdash, O. A. Dobre, O. Üreten, and T. Yensen, "Identification of cellular networks for intelligent radio measurements," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 8, pp. 2204–2211, Aug. 2017.
[4] Z. Zhang, K. Long, and J. Wang, "Self-organization paradigms and optimization approaches for cognitive radio technologies: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 36–42, Apr. 2013.
[5] K. Kim, C. M. Spooner, I. Akbar, and J. H. Reed, "Specific emitter identification for cognitive radio with application to IEEE 802.11," in *Proc. IEEE GLOBECOM*, Nov./Dec. 2008, pp. 1–5.
[6] O. A. Dobre, "Signal identification for emerging intelligent radios: Classical problems and new challenges," *IEEE Instrum. Meas. Mag.*, vol. 18, no. 2, pp. 11–18, Apr. 2015.
[7] Y. A. Eldemerdash, O. A. Dobre, and M. Öner, "Signal identification for multiple-antenna wireless systems: Achievements and challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1524–1551, 3rd Quart., 2016.
[8] Y. Huang and H. Zheng, "Theoretical performance analysis of radio frequency fingerprinting under receiver distortions," *Wireless Commun. Mobile Comput.*, vol. 15, no. 5, pp. 823–833, 2015.
[9] B. Danev and S. Capkun, "Transient-based identification of wireless sensor nodes," in *Proc. IPSN*, 2009, pp. 25–36.
[10] Y. Yuan, Z. Huang, H. Wu, and X. Wang, "Specific emitter identification based on Hilbert–Huang transform-based time–frequency–energy distribution features," *IET Commun.*, vol. 8, no. 13, pp. 2404–2412, 2014.
[11] S. Xu, "On the identification technique of individual transmitter based on signal prints," Ph.D. dissertation, Dept. Elect. Eng., Huazhong Univ. Sci. Technol., Wuhan, China, 2007.
[12] C. Bertoncini, K. Rudd, B. Nousain, and M. Hinders, "Wavelet fingerprinting of radio-frequency identification (RFID) tags," *IEEE Trans. Ind. Electron.*, vol. 59, no. 12, pp. 4843–4850, Dec. 2012.
[13] J. Zhang, F. Wang, O. A. Dobre, and Z. Zhong, "Specific emitter identification via Hilbert–Huang transform in single-hop and relaying scenarios," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 6, pp. 1192–1205, Jun. 2016.
[14] J. Zhang, F. Wang, Z. Zhong, and O. Dobre, "Novel Hilbert spectrum-based specific emitter identification for single-hop and relaying scenarios," in *Proc. IEEE GLOBECOM*, Dec. 2015, pp. 1–6.
[15] Y. Xu, G. Feng, and Y. Zhao, "One improvement to two-dimensional locality preserving projection method for use with face recognition," *Neurocomputing*, vol. 73, no. 13, pp. 245–249, 2009.
[16] A. S. Sappal, "Simplified memory polynomial modelling of power amplifier," in *Proc. IEMCON*, Oct. 2015, pp. 1–7.
[17] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
[18] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, 2015.
[19] S. Deng, Z. Huang, X. Wang, and G. Huang, "Radio frequency fingerprint extraction based on multidimension permutation entropy," *Int. J. Antennas Propag.*, vol. 2017, Jul. 2017, Art. no. 1538728.
[20] F. Zhuo, Y. Huang, and J. Chen, "Radio frequency fingerprint extraction of radio emitter based on I/Q imbalance," in *Proc. Int. Congr. Inf. Commun. Technol. (ICICT)*, 2017, pp. 472–477.