# Semantic Learning for Analysis of Overlapping LPI Radar Signals

Kuiyu Chen, Lipo Wang, *Senior Member, IEEE*, Jingyi Zhang, *Student Member, IEEE*, Si Chen, *Member, IEEE*, and Shuning Zhang

***Abstract*— The increasingly complex radio environment may cause the received low probability of intercept (LPI) radar signals to overlap in time–frequency domains. Analyzing overlapping LPI radar signals requires identifying the modulation type and estimating the parameters of each component. Prior research performs overlapping signal analysis as a multistage task, where each stage is designed to perform a part of the task. The multistage system will increase the calculation burden and cannot be optimized as a whole. Instead, this article proposes a novel framework for analyzing overlapping signals in a single stage. Specifically, we develop a joint semantic learning deep convolutional neural network (JSLCNN) that jointly learns three tasks, i.e., feature restoration, modulation classification, and parameter regression. Since the whole cognitive pipeline is a single network, it can be optimized end-to-end directly on cognitive performance. To verify the validity of the proposed JSLCNN, numerous comparative experiments are carried out in terms of modulation recognition and parameter estimation of overlapping signals. Experimental results demonstrate that the JSLCNN has desirable extensibility for identifying unseen signal combinations and robustness against unknown jamming. Furthermore, we show that the JSLCNN outperforms other existing approaches in generic real-time parameter estimation for LPI radar signals.**

***Index Terms*— Feature restoration, modulation classification, overlapping low probability of intercept (LPI) radar signals, parameter regression, semantic learning.**

## I. Introduction

BLIND analysis of intercepted signals is a significant technology performed by intelligent instruments in both military and commercial applications, such as autonomous driving, wireless communication, and electronic surveillance [1], [2], [3]. Its main tasks include signal modulation identification and signal parameter estimation, which can be utilized by an intelligent jammer to create an optimal jamming attack [4]. As the number of radiation sources in the modern electronic battlefield increases, the electromagnetic environment has become more complicated [5]. Multiple low probability of intercept (LPI) radar emitters may have to share the same bandwidth, causing the received signals to overlap in both time and frequency domains. Hence, the blind analysis of overlapping LPI signals has been a hot but fairly challenging topic.

Existing research builds multistage systems for overlapping signals, where each stage is designed to perform a part of the task and each stage feeds its output to the input of the next stage. These multistage systems generally consist of three parts, i.e., blind signal separation (BSS), automatic modulation classification (AMC), and parameter estimation. BSS techniques [6], [7], [8], [9] focus on separating the overlapping signals. Via BSS, the dilemma of overlapping signals analysis is transformed into the study of multiple single-component signals. Next, AMC algorithms [10], [11], [12], [13], [14], [15] are employed to recognize the modulation of these separated single-component signals. Finally, the corresponding parameter estimation methods [16], [17], [18] [19], [20], [21], [22] are adopted to calculate crucial parameters based on the AMC result.

However, some hurdles hamper the widespread deployment of these multistage methods. First, BSS techniques highly depend on prior information, such as the number of sources. In addition, some electromagnetic equipment usually has only one antenna. The underdetermined condition exacerbates difficulties in acquiring prior knowledge. Second, these parameter estimation methods are elaborately designed for a specific type of signal. The analysis system has to store these parameter estimation methods accordingly, which increases the burden of the hardware. Third, BSS, AMC, and parameter estimations are conducted separately, which are not particularly optimized as a whole. Thus, the analysis performance of these methods is relatively poor.

With great revolutions in deep learning, some studies attempt to eliminate the dependence on BSS. In [23] and [24], AMC methods of overlapping signals based on multiclass learning (MCL) [25] were developed. Nevertheless, each combination of overlapping signals is regarded as a new class, which inherently hinders the extensibility of the model.

Manuscript received 2 October 2022; revised 17 December 2022; accepted 13 January 2023. Date of publication 13 February 2023; date of current version 16 February 2023. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62271261 and Grant 61971226, and in part by the Natural Science Foundation of Jiangsu Province for Excellent Young Scholars under Grant BK20200075 and Grant BK20220941. The Associate Editor coordinating the review process was Dr. Michael Gadringer. *(Corresponding author: Si Chen.)*

Kuiyu Chen is with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China, and also with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: kuiyu_chen@163.com).

Lipo Wang is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ELPWang@ntu.edu.sg).

Jingyi Zhang, Si Chen, and Shuning Zhang are with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: jingyizhang_njust@163.com, csnjust@163.com, zsnnjust@163.com).

Digital Object Identifier 10.1109/TIM.2023.3242013

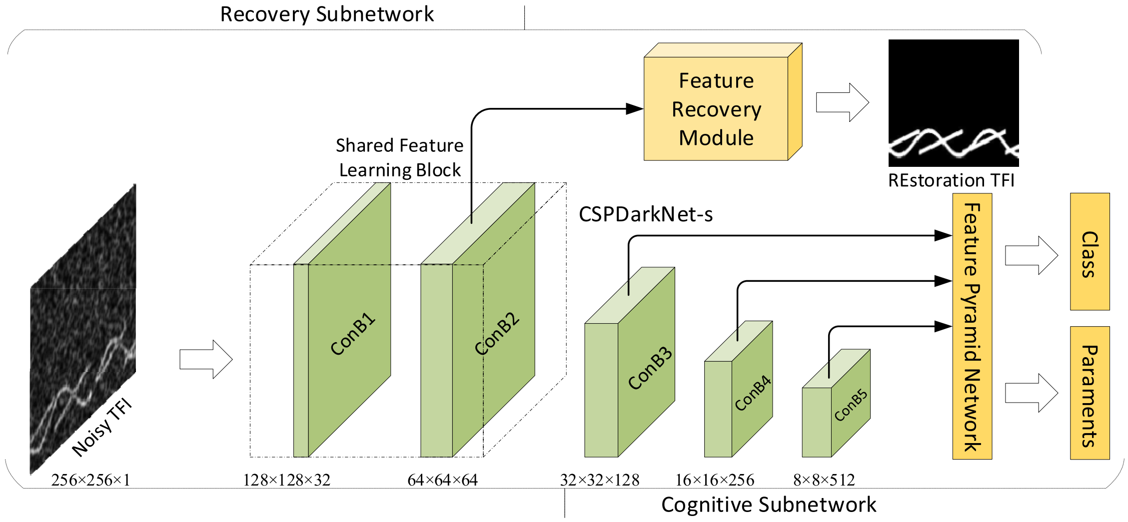


Fig. 1. Proposed JSLCNN architecture includes two submodules: an analysis subnetwork responsible for modulation recognition and parameter estimation, and a recovery subnetwork.

An alternative but closely related line of research focused on multilabel learning (MLL) [26]. MLL maps each component in input to the corresponding label. In [27], [28], [29], and [30], they have given a significant performance boost based on MLL. However, there are still some limits to MLL. First, when the overlapping signal is composed of signals with the same class, such as a signal overlapped by two linear frequency modulation (LFM) signals, MLL can only identify the existence of an LFM signal in the overlapping signal. It cannot confirm the number of LFM signals. In other words, MLL recognizes a signal composed of two overlapping LFM signals and a single-component LFM signal as the same class. This is obviously contrary to the actual situation. Second, these studies only concentrate on modulation recognition of overlapping signals while ignoring the subsequent parameter estimation.

On the other hand, some researchers focus on parameter estimation of overlapping LPI radar signals. In [31] and [32], inverse Radon transform (IRT) was adopted to estimate the parameters of sinusoidal frequency modulation (SFM) signals from time–frequency images (TFIs). However, a prerequisite for utilizing IRT is that the signal energy distribution must be in the center of the TFIs. For overlapping signals intercepted by the receiver, it is challenging to estimate the center frequency of the SFM signal in advance and move it to the middle of the TFI. In [9] and [32], singular value decomposition (SVD) was utilized to estimate the modulation frequency of SFM signals and maximum likelihood estimation (MLE) was adopted to calculate the modulation index and carrier frequency. Then, the estimated parameters were used to reconstruct the component. After removing the reconstructed part from the overlapping signal, the same procedure was repeated to estimate the remaining SFM components. However, multidimensional search is computationally expensive, which hinders practical applications. Besides, the periodicity of the peaks in SVD limits its performance on components with relatively large modulation frequencies. In [34], the Teager–Huang transform (THT) and the Hough transform were combined to compute the parameters of multicomponent LFM signals. THT offers a clear time–frequency expression and avoids cross-terms but is sensitive to noise. In [35] and [36], fractional Fourier transform (FrFT) and optimized search algorithms were utilized to calculate the chirp rates of overlapping LFM signals in the FrFT domain with a small computational cost. Because of the chirp kernel of FrFT, these methods can hardly generalize to signals with other modulation modes. Moreover, all techniques above depend on the assumption that the multicomponent signal is overlapped by signals with the same modulation modes. The estimation performance will decline obviously if signals with other modulation modes are introduced. Therefore, a generic parameter estimation approach for more complex overlapping LPI radar signals needs to be proposed urgently.

Considering all the aforementioned problems, this article develops a novel analysis method for overlapping LPI signals, called joint semantic learning deep convolutional neural network (JSLCNN). A signal analysis is reframed as a semantic learning task, straight from a short-time Fourier transform (STFT) [53] image to parameter regression and modulation classification. Meanwhile, we propose a feature restoration subnetwork for the analysis system to enhance the clarity of STFT in noisy backgrounds. The JSLCNN is trained end-to-end to simultaneously learn about feature restoration, parameter regression, and modulation classification. Benefitting from the joint optimization strategy, clean features filtered by the feature restoration subnetwork can be shared to learn better parameter regression and modulation classification, boosting the analysis performance of the JSLCNN in a noisy environment.

The main contributions of this study are listed as follows.

1) This work avoids the dependence on BSS and develops a single network to simultaneously predict signal modulation mode and estimate associated parameters. The end-to-end model can be optimized as a whole.
2) The feature-based regression model is exploited to calculate the parameters of LPI radar signals, which shares the same features as the modulation recognition module. Thus, the proposed framework is extremely fast.

3) The joint restoration subnetwork cooperates learning with the analysis system to better study the overlapping LPI signal under intensely noisy environments without increasing the computational burden.

The remainder of this article is organized as follows. Section II provides a brief description of the overlapping signal model and semantic learning. Section III introduces the proposed JSLCNN in detail. In Section IV, numerous comparative experiments are conducted on modulation recognition and parameter estimation of overlapping signals to validate the superiority of the JSLCNN. Finally, the conclusion is presented in Section V.

## II. Problem Formulation

This work reformulates the analysis of an overlapping LPI radar signal as a semantic learning task, aiming to simultaneously identify the modulation mode and estimate the parameter of each constituent.

### A. Overlapping LPI Radar Signals

In a complicated electromagnetic environment, multiple radiation sources may work simultaneously. These transmitted signals will overlap in time and frequency domains when they are intercepted by the receiver. Assume that the receiver intercepts $K$ standalone radio frequency (RF) LPI radar signals. The intercepted RF signal is downconverted into an intermediate frequency (IF) signal and then sampled at sampling frequency $f_s$ to generate digital $y[n]$ as

$$y[n] = \sum_{i=1}^{K} A_i s_i[n] + w[n] \tag{1}$$

where $s_i[n]$ represents the radar signal from the $i$th independent radar emitter, $A_i$ denotes the amplitude of $s_i[n]$, and $w[n]$ is the additive white Gaussian noise.

This work takes typical continuous-wave (CW) LPI radar signals [52] as an example. There are frequency modulation signals, i.e., LFM and SFM signals, and phase code signals, i.e., binary phase shift keying (BPSK) and Frank signals. The related signals are detailed in Table I.

### B. Semantic Learning for Overlapping Signals

Semantics is the primary carrier of knowledge and information. Semantic learning [37] converts input data into understandable expression, depicting concrete visual-to-cognitive issues. Instead of straightforwardly describing the whole input like traditional AMC tasks, semantic learning can focus on the exciting local object and can provide a high-level illustration. Thus, a semantic learning-based method is an intuitive and promising solution to simultaneously implement modulation recognition and parameter estimation for overlapping signals.

## III. Methodology

This research explores an overlapping LPI radar signals analysis from a new perspective and proposes a one-stage strategy to simultaneously classify the signal module and predict the signal parameters. The structure diagram of the proposed method is shown in Fig. 1.

TABLE I
Expression of Relevant Radar Signals

| Modulation | Expression |
|---|---|
| LFM | $s[n] = Ae^{j[2\pi((f_c - k_m T_m/2)n + \frac{k_m n^2}{2}) + \varphi_0]}\ 0 < n \le T_m$ |
| SFM | $s[n] = Ae^{j[2\pi f_c n + m_f \sin(2\pi n/T_m + \varphi_0)]}\ 0 < n \le T_m$ |
| BPSK | $s[n] = Ae^{j(2\pi f_c n + \varphi[n] + \varphi_0)}\ 0 < n \le T_m$<br>$\varphi[n] = \sum_{i=1}^{I} \phi_i \cdot r_{T_s}(n - iT_s)$ |
| Frank | $s[n] = Ae^{j(2\pi f_c n + \varphi[n] + \varphi_0)}\ 0 < n \le T_m$<br>$\varphi[n] = \sum_{i=1}^{N_P^2} \phi_{p,q} \cdot r_{T_s}(n - iT_s)$<br>$\phi_{p,q} = \frac{2\pi}{N_P}(p-1)(q-1)\quad 1 \le p,q \le N_P$ |

Here $A$, $f_c$, $T_m$, $T_s$, and $\varphi_0$ denote the amplitude, the center frequency, the modulation period, the width of the symbol, and the initial phase, respectively. $j$ is the imaginary unit. $k_m$ and $m_f$ indicate the modulation slope and the frequency modulation index. $\phi$ is phase encoding sequence and Barker code is applied in BPSK. $r_{T_s}[n] = u[n + T_s/2] - u[n - T_s/2]$, $u[n]$ is the step function. $I$ describes the length of Barker code. $N_P$ denotes the number of phases.

### A. Time–Frequency Analysis

Time–frequency analysis [53], a generalization and refinement of Fourier analysis, provides the conjoint distribution of the signal in both time and frequency domains. It has a wide range of applications among signals with frequency-varying characteristics, such as speech signals, electroencephalogram signals, and radar signals.

Common time–frequency analysis tools mainly involve STFT, Wigner–Ville distribution transform (WVD) [54], and Choi–William distribution transform (CWD) [55]. Given the computational complexity and cross-terms in nonlinear time–frequency transform, this article adopts STFT as a time–frequency analysis tool. The STFT, a special windowed Fourier transform, is calculated as

$$\text{STFT}[m, l] = \sum_{n=0}^{N-1} y[n]h[n-m]e^{-j2\pi nk/N} \tag{2}$$

where $h[n]$ indicates the windows function and $\text{STFT}[m, l]$ depicts the frequency change of $y[n]$ along with time $m$.

### B. Classification of Overlapping Signal

Current signal recognition methods [10], [11], [12], [13], [14], [15] repurpose classifiers to fulfill identification. To recognize the modulation mode, these approaches apply a classifier for a whole input (TFI, IQ samples, or other feature domains) and map it to a label space (one-hot vector). Although MLL [27], [28], [29], [30] is capable of overlapping signals identification, it is a coarse recognition and can hardly recognize overlapping signals with the same modulation class.

This work reframes overlapping signals recognition as a semantic-based task to identify various subimages in TFI. As shown in Fig. 2, the entire TFI of overlapping signals

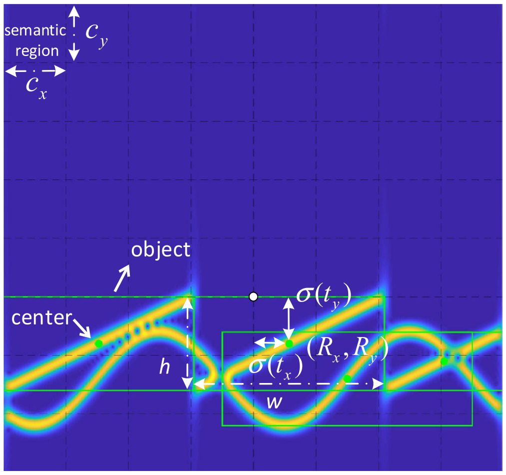


Fig. 2. Illustration for semantic-based identification and regression.

is decomposed into multiple semantic regions by stacking convolutional blocks (CBs), and the fine-grained identification is realized on the instance level. The energy distribution of a signal with a complete modulation period in TFI is defined as an object. Each semantic region is responsible for identifying the object whose center falls into the cell. Generally, when objects in the overlapping signal fall in different regions, these regions can classify these objects simultaneously. Even though these objects fall in the same grid, multiple detection heads in the JSLCNN can realize the identification of these objects. Binary cross-entropy loss [27] is utilized to split a multiclass classification into one binary classification problem per class. Each binary classification learns class-specific discriminative features that individually distinguish between matches and nonmatches for each class.

*1) Generic Parameter Representation:* For parameter estimation of the radar signal, there has always been a lack of a generalized method. Existing methods usually focus on a specific type of signal modulation. To estimate the parameter of different signals, the system has to integrate different parameter estimation methods, which will increase the complexity of the system. Thus, a parameter estimation technique providing a generic parameter expression is urgently needed.

A time–frequency spectrum can provide the energy distribution of the signal in both time and frequency domains. Although Hough translation [16], [17] and Radon translation [31], [32] have been utilized to perform parameter estimation based on the mathematical properties of geometric shapes in the time–frequency spectrum, they are still hard to generalize to other signals. Besides, the expensive computational cost limits their further applications. This work develops a novel parameter estimation method by approximating the boundary [56] of the energy distribution in the time–frequency spectrum. First, the parameters that need to be estimated in various signals can be generalized as $f_c$, $T_m$, and bandwidth $B$ (for frequency modulation signals). If the energy distribution of a signal component with a complete modulation period can be approximated by a rectangular box with width $w$, height $h$, and center coordinates $(R_x, R_y)$, the relationship between signal parameters and the rectangular box can be expressed as

$$f_c = R_y \times f_s / H_{\mathrm{TFI}} \tag{3}$$

$$T_m = w / f_s \tag{4}$$

$$\begin{aligned} B &= T_m \times k = 2 \times m_f / T_m \\ &= h \times f / H_{\mathrm{TFI}} \end{aligned} \tag{5}$$

where $H_{\mathrm{TFI}}$ is the height of the input TFI.

*2) Deep Regression Model:* Regression analysis [38], a predictive modeling technique, describes the relationship between independent variables and a response, dependent, or target variable. This technique has made remarkable achievements in the market analysis [39], stock forecasting [40], and inferring gene expression [41]. For parameter estimation of radar signals, a regression analysis can predict possible parameters by establishing a predictive model between TFIs and signal parameters. To the best of our knowledge, there are still few studies utilizing regression analysis for the task of parameter estimation.

Since the value of predicted parameters is usually much greater than one, to facilitate network optimization, this work establishes a regression model to predict offsets in terms of the top-left corner of the semantic area, the height, and width of the TFI, instead of directly predicting $w$, $h$, $R_x$, and $R_y$. These predicted values can be defined as

$$\sigma(R_x) = R_x / c_x - \lfloor R_x / c_x \rfloor \tag{6}$$

$$\sigma(R_y) = R_y / c_y - \lfloor R_y / c_y \rfloor \tag{7}$$

$$\sigma(w) = w / W_{\mathrm{TFI}} \tag{8}$$

$$\sigma(h) = h / H_{\mathrm{TFI}} \tag{9}$$

where $\sigma()$ represents the predicted offsets value, $\lfloor . \rfloor$ indicates the round down function, $c_x = c_y$ is the width of the semantic region, and $W_{\mathrm{TFI}}$ denotes the width of the input TFI.

Optimizing these regression outputs is a tricky dilemma. Mean squared error (mse) is an excellent evaluation metric for regression tasks, and work [42] utilizes a similar metric called sum of squared error (SSE) for boundary regression. However, these metrics cause the model to overfit the regression object in the training sample. To alleviate this issue, the intersection over union (IoU) loss function [43] is introduced to regress the four offsets as a whole unit as follows:

$$L_{\mathrm{reg}} = -\ln(I/U) \tag{10}$$

where $I$ and $U$ indicate the intersection and union between the predicted and actual boundaries, respectively. Benefitting from IoU loss, boundary regression can be optimized jointly rather than four independent regressions. Thus, the estimation model converges more efficiently and predicts more accurately.

*3) Supplementary Hierarchical Classification for Parameter Estimation of Phase Code Signals:* The regression model realizes the parameter estimation of frequency modulation signals. For phase code signals, $f_c$ and $T_m$ can be fulfilled by boundary regression, while $T_s$ is still an unresolved problem. $T_s$ can be calculated as

$$T_s = T_m / \mathrm{num}_s \tag{11}$$

where $\mathrm{num}_s$ indicates the number of phase codes or frequency steps.

If the estimation of $\mathrm{num}_s$ is realized, $T_s$ can be smoothly solved. Thus, a hierarchical classification network is designed

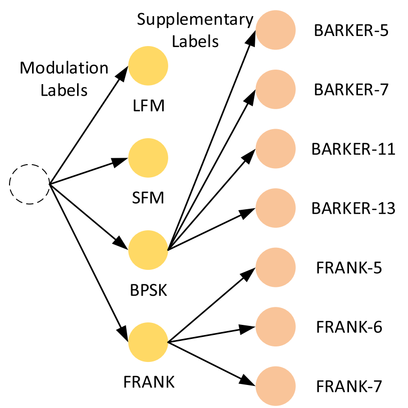


Fig. 3. Hierarchical label tree of LPI radar signals. Here, BARKER-5 denotes the BPSK signal modulated by Barker Code 5. FRANK-5 represents the FRANK signal with five frequency steps.

to predict the phase codes of BPSK signals and the frequency steps of Frank signals. Fig. 3 shows the hierarchical label tree, which structures concepts of classes, their subclasses, and how they relate. In the label tree, "BARKER-5" and "BARKER -7" are both hyponyms of "BPSK." These hyponyms are displayed as leaves and clustered into main categories. The loss function of a hierarchical classification network is a weighted summation of the modulation class and supplementary subclass, which is defined as

$$L_{\text{cla}} = \sum_{d=1}^{D} \sum_{i=1}^{C} L_{di} \tag{12}$$

where $D = 2$ indicates the number of levels in the label tree and $C = 4$ denotes the number of modulations considered in this work. $L_{di}$ represents the binary classification loss of $i$th class on the $d$th level tree.

### C. Joint Optimization of Classification and Regression in One Stage

In the proposed analysis networks, the efficient and powerful CSPDarkNet [44] is adopted as the shared backbone feature extraction network for classification and parameter estimation. As shown in Fig. 1, CSPDarkNet consists of five shallow-to-deep CBs. Due to the hierarchical downsampling structure of CNN, semantically stronger features are spatially coarser. Specifically, shallow low-level feature maps are not semantically rich enough to be directly used for classification, while they contain abundant shape and edge information for boundary regression. Deep high-level features are more trustworthy for identification but have low resolution. Thus, the feature pyramid network (FPN) [45] follows the backbone to create an architecture where the semantically richer features are merged with the features from previous layers by continuous upsampling, cross-layer fusion mechanisms, and top-down connections. Finally, decoupled head branches predict modulation and estimate parameters.

Fig. 4 shows the structure of FPN and the decoupled head. FPN constructs three-scale feature pyramids consisting of features with the shape of $8 \times 8 \times 512$, $16 \times 16 \times 256$, and $32 \times 32 \times 128$, based on feature maps from ConB3, ConB4, and ConB5, respectively. In other words, the original TFI with the shape of $256 \times 256 \times 1$ is divided into $32 \times 32$, $16 \times 16$, and $8 \times 8$ semantic units. Via FPN, different semantic units all contain the semantic information of the original deepest feature map and edge information from the original shallow layers. Next, $1 \times 1$ convolutions in the decoupled head are applied to reduce the number of feature channels to 128. The subsequent combination of $3 \times 3$ convolutions and $1 \times 1$ convolutions map features to classification. Another parallel branch map features to five values, i.e., $\sigma(R_x)$, $\sigma(R_y)$, $\sigma(w)$, $\sigma(h)$, and confidence, which denotes the probability that the cell has an object. Decoupled head maps the shared features to predict classification and regression simultaneously, which integrates modulation recognition and parameter estimation into one stage. As we will show in this article, our one-stage analysis method significantly improves computing speed and analysis performance.

For three-scale feature pyramids, the number of all potential predictions can be calculated as

$$\text{num}_P = \sum_{i=1}^{3} H_i \times W_i = 1344 \tag{13}$$

where $H_i$ and $W_i$ indicate the length and width of the $i$th scale feature pyramid, respectively.

Among these predictions, only a few contain positive samples and the vast majority are negative samples. Thus, SimOTA [44] is used to pick out all possible positive samples among all predictions. Then, the joint loss function is applied to optimize the JSLCNN as follows:

$$L_A = L_{\text{cls}}/N_{\text{pos}} + \text{reg}_{\text{weight}} * L_{\text{reg}}/N_{\text{pos}} + L_{\text{obj}} \tag{14}$$

where $L_{\text{obj}}$ denotes the loss of confidence, and binary cross-entropy loss is adopted in this optimization. $N_{\text{pos}}$ indicates the number of positive samples determined by SimOTA. $\text{reg}_{\text{weight}}$ represents a balancing term applied to weigh the regression loss over the other losses for its importance, which is set as 5 in subsequent experiments according to the literature [44].

### D. Recovery Subnetwork

Given the complicated working environment of radars, the intercepted signals may be disturbed by intense noise. Multitask learning strategies [46], [47] are adopted in this work. A restoration subnetwork is developed to repair the noisy TFI, which will simultaneously enhance signal analysis capability in low signal-to-noise ratio (SNR) conditions. As shown in Fig. 1, a shared feature learning block (SFLB) and a feature restoration block (FRB) constitute the recovery subnetwork.

Considering that spatial features promoting image restoration are mainly contained in shallow layers, whereas those of deeper layers are destroyed during continuous downsampling, we choose the first two CBs in the analysis subnet to constitute SFLB. The FRB follows the classic structure of the decoder. First, a dimension reduction module utilizes $1 \times 1$ convolution to decrease $8\times$ channels of features extracted by SFLB. Second, the bilinear interpolation technique [48] is applied to recover the feature size as the input TFI in an upsampling module. Then, the fusion of multiscale features is realized

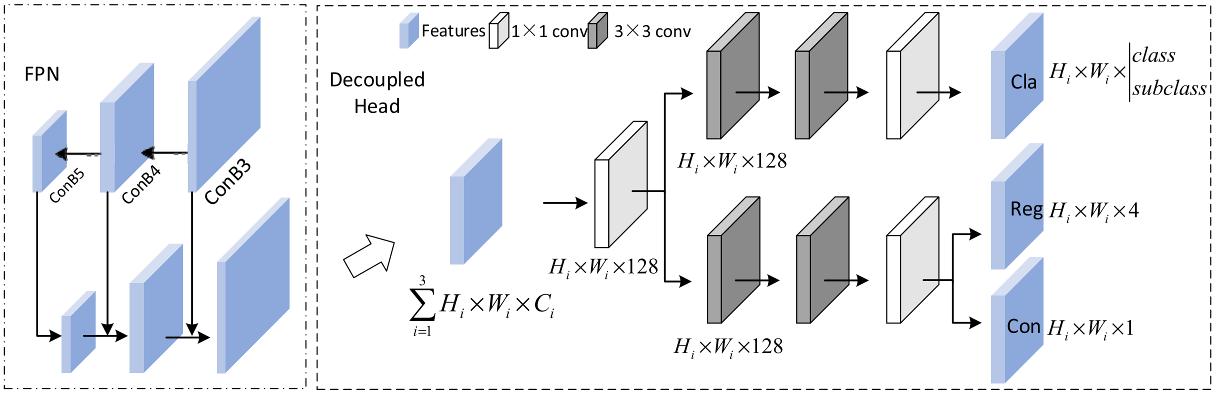


Fig. 4. Structure of FPN and the decoupled head.

by a spatial pyramid pooling (SPP) module [49]. Finally, the restored TFI is generated by another 3 × 3 convolution. The shared design is beneficial to preserving a simple structure of the recovery subnetwork. The feature maps extracted by SFLB are simultaneously transferred to the FRB for clarity optimization and subsequent CBs for modulation recognition and parameter estimation.

The recovery subnetwork removes the harmful effect of noise by the mse loss as follows:

$$L_R = \left(\widehat{T_i} - T_i\right)^2 \tag{15}$$

where $T_i$ represents the noise-free TFI label and $\widehat{T_i}$ indicates the recovery TFI by FRB.

Note that the recovery subnetwork is designed not to generate the noise-free TFI as an input of the analysis subnetwork despite the fact that it can directly produce noise-free TFI. Instead, the recovery subnetwork is developed to improve the analysis performance of the proposed model by jointly optimizing learning. In other words, the joint learning recovery subnetwork will not add any parameters or computation costs in the test phase.

TABLE II
PARAMETERS OF INVOLVED SIGNALS

| Signals | Parameter | Scope |
|---|---|---|
| LFM | center frequency $(f_c)$ | $(0.1 \sim 0.4) * f_s$ |
| | bandwidth $(B)$ | $(0.1 \sim 0.2) * f_s$ |
| | modulation period $(T_m)$ | $(0.25 \sim 0.5) * N$ |
| SFM | $f_c$ | $(0.1 \sim 0.4) * f_s$ |
| | $B$ | $(0.1 \sim 0.2) * f_s$ |
| | $T_m$ | $(0.25 \sim 0.5) * N$ |
| BPSK | $f_c$ | $(0.05 \sim 0.45) * f_s$ |
| | Barker code | [5,7,11,13] |
| | $T_m$ | $(0.25 \sim 0.5) * N$ |
| FRANK | $f_c$ | $(0.1 \sim 0.4) * f_s$ |
| | $N_p$ | [5,6,7] |
| | $T_m$ | $(0.25 \sim 0.5) * N$ |

where $N$ is the length of the signal.

## E. Dynamic Joint Learning

In the training process of the JSLCNN, the final loss function takes the recovery subnetwork and the analysis subnetwork into consideration to ensure that the restoration structure can play the role of internal guide to the whole network as follows:

$$L = A_R L_R + A_A L_A \tag{16}$$

where $A_R$ is the loss weight of the recovery subnetwork and $A_A$ symbolizes the loss weight of the analysis subnetwork.

Inspired by the branch training strategy [50], this work adjusts the loss weight of each subnetwork dynamically during the training process. The loss weight depicts the contribution of each subnetwork to the final loss function. For example, if $A_R = 0$ and $A_A = 1$, the JSLCNN converges to a single analysis network without training the recovery subnetwork. Also, loss weights of $A_R = 0.9$ and $A_A = 0.1$ express that the JSLCNN values the feature recovery but also wants to train a little bit of analysis module in experiments. During the designed dynamic joint learning (DJL) strategy, the "focus" of distribution will shift from the recovery subnetwork to the analysis subnetwork. This procedure requires the JSLCNN to extract visual enhancement features first and fine-tune parameters to further analysis. The dynamic training strategy exploits the potential of loss weight distribution to achieve better performance.

# IV. EXPERIMENTS AND DISCUSSION

In this section, to verify the superiority of the proposed JSLCNN, extensive experiments are carried out from two aspects, i.e., modulation identification and parameter estimation. During experiments, each detected component will be distinguished according to its parameters. The components with the same parameters will be generalized as a CW signal, and the parameters of this signal will be determined by the component with the largest confidence value.

*Training Data:* Four kinds of CW LPI radar signals are considered in this study, whose parameters are randomly distributed (Table II). IF sampling frequency $f_s$ is set as 400 MHz. A set of $N = 1600$ consecutive signal samples is collected for analysis. All combinations of single-, two-, and three-component signals are involved in training data. Each component has the same power. When the SNR is from $-8$ to 12 dB at intervals of 2 dB, we randomly simulate 500 samples of each combination for training data; 80% of the samples are selected randomly as the training set, and the rest is as the validation set. There are 149 600 training samples in total, and the number of validation samples is 37 400. A Hamming window with a length of 63 is adopted in the STFT analysis in subsequent processing.

*Experimental Environment:* Network computation of the JSLCNN is carried out on the RTX 3070 graphic processor and implemented in Pytorch 1.10.1. We train our model 50 epochs on training and validation sets. Adam optimizer (beta1 = 0.9, beta2 = 0.999, epsilon = $1e^{-8}$, and weight decay = $5e^{-4}$), step learning rate schedule (initial learning rate $\eta = 0.001$ and attenuation factor $\gamma = 0.92$ for each epoch), and a batch size of 32 are adopted throughout training.

*Component Filtering:* During the testing phase, the semantic regions with relatively large predicted scores are first filtered out. Predicted scores can be calculated as the product of the classification and confidence scores. The components with a prediction score below the threshold would be abandoned, and those with a prediction score above the threshold will be selected for parameter estimation. The threshold utilized in the experiment is 0.85. The threshold model reserves more space for unknown or meaningless signals, which contributes to reducing the open space risk. Then, nonmaximum suppression (NMS) algorithm [57] is adopted to pick the most representative predicted region (with the largest prediction score) within adjacent semantic regions. The IoU threshold utilized in the NMS is 0.75. These thresholds are referred to the relevant literature [42], [44]. Next, the filtered semantic region will present the parameters of signal components. In addition, one CW signal could have different periodic components in the TFIs. A pair of components with the same modulation type will be regarded as the different components of one CW signal if they meet the following conditions: 1) their center frequency, modulation period, and bandwidth (or the width of the symbol) are almost the same and 2) the phase difference between them is equal to a modulation period. Finally, the component with a high predicted score will be selected to represent this signal.

## A. Recognition Performance

*1) Evaluation Metric:* The overall accuracy (Acc) is a metric to describe a fraction of correctly recognized signals (the predicted result totally matches the actual label of the testing signal). Acc is defined as

$$\mathrm{Acc} = \frac{1}{N_t}\sum_{i=1}^{N_t}(P_i = R_i) \tag{17}$$

where $N_t$ denotes the number of testing signals, $P_i$ represents the predicted result of the $i$th testing signal, and $R_i$ is the corresponding real label.

During transmission, signals are inevitably disturbed by noise in the channel. SNR is adopted to express the quality of the received signal as follows:

$$\mathrm{SNR} = 10\log_{10}\left(\left(\sum_{i=1}^{K}\sigma_i^2\right)\Big/\sigma_n^2\right) \tag{18}$$

where $\sigma_i^2$ and $\sigma_n^2$ symbolize the variance of the $i$th signal component and the noise variance, respectively, and $K$ represents the number of radar signal components in the received signal.

Except for channel noise, signals radiated from other unconcerned sources will interfere with the concerned LPI radar signals. Thus, we define the signal-to-jamming ratio (SJR) as

$$\mathrm{SJR} = 10\log_{10}\left(\left(\sum_{i=1}^{K}\sigma_i^2\right)\Big/\sigma_j^2\right) \tag{19}$$

where $\sigma_j^2$ represents the variance of the jamming signals.

Due to the difference in transmission distances and transmitting power, each component in overlapping signals may have a different power. A power ratio (PR) is utilized to describe the difference in the energy as follows:

$$\mathrm{PR} = 10\log_{10}\left(\sigma_i^2/\sigma_r^2\right) \tag{20}$$

where $\sigma_i^2$ and $\sigma_r^2$ represent the power of the $i$th component and the power of the rest components in the overlapping signal, respectively.

*2) Baseline Methods:* To reveal the excellence of the JSLCNN, two existing overlapping signal modulation recognition approaches are employed as performance baselines.

*a) Multiclass Learning CNN (MCLCNN):* It treats each overlapping combination as a class, which is established by the same backbone as the JSLCNN [25]. Instead of FPN and a decoupled head, a global average pooling layer, two fully connected layers (fc-256 and fc-128), and a softmax layer followed by ConB5 predict the classification. The length of the output label vector is the same as the number of the overlapping combination.

*b) Multilabel Learning CNN (MLLCNN):* It maps each component in input to the corresponding label [27], [28], [29]. A baseline MLLCNN shares the same backbone as the baseline MCLCNN. Compared with the MCLCNN, the softmax layer and the categorical cross-entropy are replaced by a sigmoid layer and the binary cross-entropy in the MLLCNN. The length of the label vector is the same as the number of classes.

The experimental environment and training schedule are the same as the JSLCNN. Then, sufficient experiments are conducted to validate the superiority of the JSLCNN.

*3) Recognition Robustness Against SNR:* During transmission, channel noise will destroy details in signals, which may reduce the recognition performance of systems. This section establishes testing data consisting of all combinations when the number of components is less than or equal to three. The parameters of each component are randomly distributed (Table II), and each component has the same power. When the SNR is from $-12$ to 12 dB at intervals of 2 dB, we randomly simulate 100 samples of each combination as testing data. Fig. 5 reveals the overall recognition performance against the noise of three systems. The overall accuracy

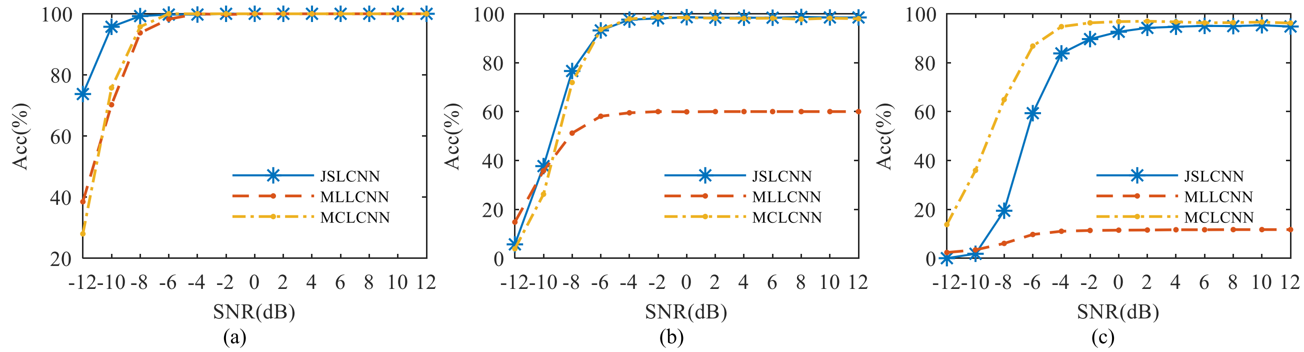


Fig. 5. Recognition performance against SNR. (a) Performance for single-component signals. (b) Performance for two-component signals. (c) Performance for three-component signals.

of all methods is enhanced with the improvement of SNR. The proposed JSLCNN can recognize single-component signals with a nearly 100% accuracy when the SNR is over $-10$ dB. The JSLCNN is also superior to baselines at all the composite SNRs in terms of single-component signals. For two-component LPI radar signals, the JSLCNN still has the best recognition performance, which achieves nearly 99% accuracy when the SNR exceeds $-6$ dB. When the number of components in the signal is more than one, the MLLCNN has the worst performance among all methods. This is because the MLLCNN has no ability to recognize signals overlapped by the same modulation. For three-component LPI radar signals, the JSLCNN can still realize modulation recognition with a nearly 95% accuracy when the SNR reaches 0 dB. The recognition performance of the MCLCNN is better than that of the JSLCNN under low SNRs. This is because of the loss of semantic information in TFI caused by intense noise and more overlapping components. Generally, the identification performance of the JSLCNN of each combination is desirable at all the SNR values.

*4) Recognition Robustness Against Power Ratio:* Due to different transmission distances and the transmitting powers of various radiation sources, each component in overlapping signals may not have the same energy. The difference in the power between each signal component will dramatically influence the recognition performance. Signal components with high energy may suppress weak ones in both time and frequency domains. Thus, the recognition system should be robust for an extensive range of PRs. Namely, the system should pay the same attention to each component with different power in an overlapping signal.

The above-defined PR is utilized to analyze the performance of systems. Due to the concern of the robustness against PR rather than the absolute accuracy, we select combinations of two-component signals instead of all combinations as testing data. In addition, to better investigate the influence of PR, the SNR is fixed at 20 dB to suppress the effect of noise. The variation range of PR is set from $-12$ to 12 dB at intervals of 3 dB. The power of the other component is fixed at the value of 1. The power of the selected component is computed by multiplying PR by the value of 1.

Fig. 6 shows the recognition curve of two-component signals at different PR values. The proposed JSLCNN and the

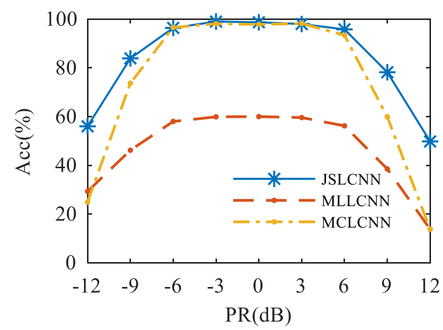


Fig. 6. Recognition robustness against PR.

MCLCNN maintain high precision when the PR is from $-3$ to 3 dB. As the power discrepancy increases, the JSLCNN performs more robustly than the MCLCNN. It can also be demonstrated that the recognition curve is approximately symmetrical about the midpoint, which indicates that the identification result is optimal when each component has the same power and decreases with the extension of the energy discrepancy. This is because the features represented by a component with relatively low energy will fade in TFI. In addition, the identification performance with PR from $-12$ to $-3$ dB is better than that with PR from 3 to 12 dB, revealing that differences between components are sensitive to noise.

*5) Recognition Performance Facing Unseen Signal Combinations:* As the number of components in the overlapping signal increases, the amounts of potential combinations increase exponentially. Training models utilizing all possible combinations are almost impossible. Thus, the recognition system should have the extensibility to recognize unseen signal combinations in training data. The proposed JSLCNN focuses on local interesting semantic information rather than global features by decomposing the TFI of an overlapping signal into multiple semantic regions. For unseen signal combinations, the JSLCNN can automatically decompose unseen mixtures into several seen signal instances. Benefiting from fine-grained semantic learning, training data utilized in the JSLCNN eliminate the need to cover all possible signal combinations.

First, unseen combinations overlapped by four components are generated. Since the combinations of all four-component signals are excessive, five cases are produced to verify the scalability of the JSLCNN. Case 1 consists of signals overlapped

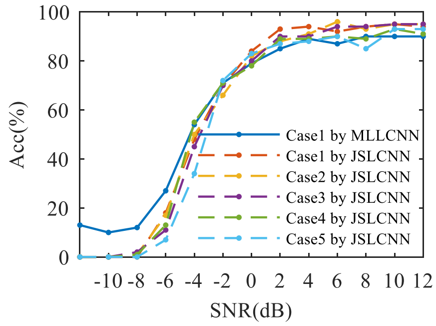


Fig. 7. Performance for unseen combinations.

by an LFM signal, an SFM signal, a BPSK signal, and a Frank signal. Case 2 includes the signals overlapped by two LFM signals, an SFM signal, and a Frank signal. Case 3 consists of signals overlapped by an LFM signal, two SFM signals, and a BPSK signal. Case 4 includes the signals overlapped by an LFM signal, two BPSK signals, and a Frank signal. Case 5 consists of signals overlapped by an SFM signal, a BPSK signal, and two Frank signals. Each component has the same power, and the parameter of these components is randomly distributed in the ranges shown in Table II. The identification results for unseen combinations are plotted in Fig. 7. Since the MCLCNN is not extensible, it is powerless against unseen combinations. For the MLLCNN, it can only recognize Case 1. The same modulated components in Cases 2–5 inherently harm the recognition of the MLLCNN. The proposed JSLCNN can identify all these unseen combinations with high accuracy, revealing the desirable extensibility of the JSLCNN.

*6) Recognition Performance Under Unknown Jamming:* In practical applications, the receiver may inevitably encounter some unknown jamming signals. Simulating all potential jamming during training is costly, time-consuming, and impossible. For a reconnaissance receiver, unknown communication signals in overlapping signals may interfere with the recognition of LPI radar signals. In traditional AMC problems, signal recognition under unknown interference has always been a hot but rather complicated topic. Unknown jamming signals will cause the feature distribution of overlapping signals to be tested to deviate from that in the training data, which will affect signal recognition to a certain extent. Thus, it is significant to analyze the recognition performance of overlapping signals under unknown jamming.

We have collected a set of actual communication signals in open space outdoors as unknown jamming to further verify the recognition performance of systems. As shown in Fig. 8, the collection system consists of a receiving antenna and an Agilent DSO81204B high-speed sampling oscilloscope with a 10-GHz sampling frequency. Fig. 9 shows the frequency spectrum of collected signals. It can be seen that signals within 0.85–1.05, 1.7–1.9, and 2.5–2.7 GHz are stronger than others, which corresponds to the spectrum situation in China [51]. Three local oscillator signals with carrier frequencies of 0.85, 1.7, and 2.5 GHz are utilized to mix with the collected signals. After passing through low-pass filters, the IF-collected signals are obtained. Then, these three signals are added to LPI radar signals as unknown jamming. When the SJR is from

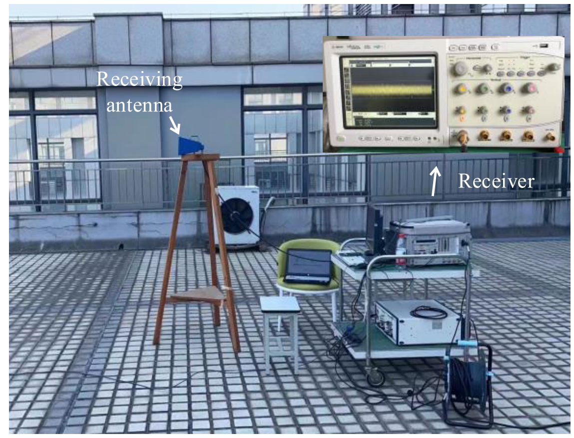


Fig. 8. Collection system of actual jamming signals.

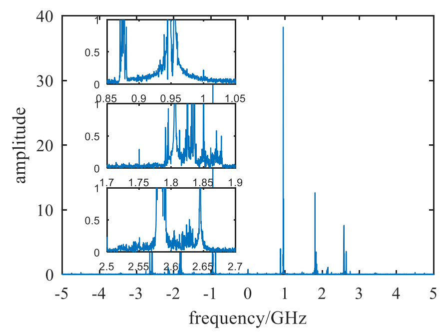


Fig. 9. Frequency spectrum of the collected signal.

−9 to 9 dB at intervals of 3 dB, we randomly simulate 100 samples of each type of overlapping combination for testing data. The SNR of testing data is 6 dB, and each component has the same energy.

Fig. 10 shows the recognition performance under measured unknown jamming. The identification results of the MCLCNN drastically deteriorate as SJR falls. The MCLCNN maps the whole feature distribution to class labels. Thus, the bias of the feature distribution caused by unknown jamming signals will remarkably affect the recognition performance. For one-component signals, extra jamming signals make its features distribution very similar to the signal overlapped by two components so that MCLCNN more likely maps it to a label representing two-component signals. Similar confusion will also occur for two-component signals. Although the feature distribution of three-component signals under jamming is similar to a signal overlapped by four components, the MCLCNN can hardly map it to an unseen combination in training data. This may explain the phenomenon that robustness against unknown jamming for three-component signals is better than others. The JSLCNN can provide more robust recognition results on signals combined by the different components compared to the MCLCNN when dealing with unknown jamming. The semantic-based JSLCNN divides the whole TFI into several semantic regions. Compared to the feature bias in the entire TFI, each region suffers from less offset, contributing to its robustness against unknown jamming. As the number of signal components increases, the energy distribution in the time–frequency domain also becomes crowded.

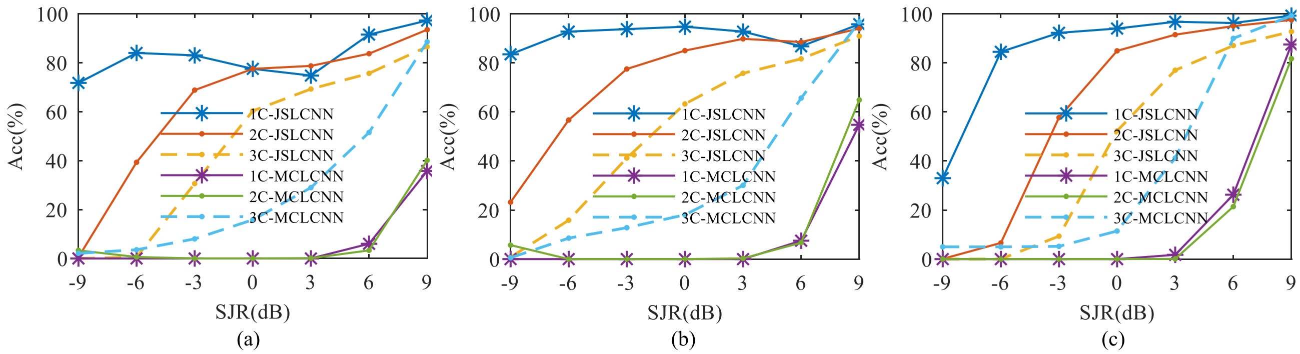


Fig. 10. Identification performance under unknown jamming. (a) With an IF jamming signal obtained by a 0.85-GHz local oscillator. (b) With an IF jamming signal obtained by a 1.7-GHz local oscillator. (c) With an IF jamming signal obtained by a 2.5-GHz local oscillator. Here, 1C, 2C, and 3C represent single-component signals, two-component signals, and three-component signals, respectively.

Jamming signals are more likely to cover a certain frequency band of the three-component signal, which inevitably destroys the detailed information of the three-component signal. Thus, the recognition performance of the JSLCNN for three-component signals is degraded. Nonetheless, identification curves of three-component signals decline gently when the SJR is over 0 dB, which further proves the practicality of the JSLCNN.

### *B. Parameter Estimation for Overlapping Radar Signals*

*1) Evaluation Metric:* To quantify the estimation performance, normalized mean squared error (NMSE) is adopted as the evaluation indicator

$$\text{NMSE} = \frac{1}{R}\sum_{i=1}^{R}(\widehat{m_i}/m_i - 1)^2 \tag{21}$$

where $\widehat{m_i}$ represents the estimated parameter for the $i$th iteration of Monte Carlo experiments and $m_i$ is the actual value of the parameter. $R = 100$ is the number of Monte Carlo experiments conducted for each value of SNR.

*2) Baseline Methods:* To demonstrate the estimation performance of the JSLCNN, several existing methods are adopted as baselines.

*Hough [16], [17]:* Hough transform is applied to estimate the parameter of straight lines in the pseudo-Wigner–Ville distribution (PWVD) image of LFM signals. The position of the maximum value in PWVD is adopted to calculate $f_c$. Hough transform theta resolution is set as a coarse 1° and then fine-tuned in 0.1° resolution to detect the chirp rates. The slice at the angle corresponding to the maximum value in the Hough transform is utilized to estimate $T_m$ and $B$.

*FrFT [35]:* This is a successive coarse-to-fine grid-search method for calculating the chirp rates of overlapping LFM signals in the FrFT domain. Fractional power is set as −1:0.01:1 first to coarse search, and then, power resolution is changed to 0.001. $T_m$ and $B$ are computed based on the distance between peaks in FrFT. $f_c$ is determined by FrFT of signal with a complete $T_m$ without offset. Offset is searched in $T_m$ by the maximum of FrFT.

*SVD-IRT [31], [32]:* This method adopts IRT to transform a 2-D sinusoidal pattern in TFI into a single point in a 2-D plane to estimate the parameter of SFM signals. SVD was utilized to compute $T_m$. The S-method is selected as a time–frequency representation. $f_c$ is obtained by the projection of TFI on the frequency axis. The energy distribution of an SFM signal is shifted to the center of TFI to perform IRT to obtain $B$.

*SVD-MLE [8], [33]:* SVD is utilized to search $T_m$ of SFM signals and adopts MLE to calculate $B$ and $f_c$. Then, the calculated parameters are applied to reconstruct the component. After removing the reconstructed component from the overlapping signal, the same procedure is repeated to estimate the remaining SFM components.

*Zhao, Atlas, and Marks (ZAM) [22]:* This is a time–frequency representation of ZAM distribution-based method to estimate the code rate and $f_c$ of BPSK signals from the negative peaks of the ZAM.

Considering that a few studies focus on parameter estimation of Frank signals and Frank signals share several similar characteristics to LFM signals [20], [21], we utilize Hough and FrFT to estimate the parameters of Frank signals as baselines.

*3) Signal Case:* To verify the validity of the proposed JSLCNN, we generate three different monocomponent signals for each type of modulation as follows.

*LFM:* The signal LFM case 1 (LFM_1) has the parameters $\{T_m, f_c, B, \varphi_0\} = \{1.25\ \mu s, 80\ \text{MHz}, 80\ \text{MHz}, \pi\}$. LFM_2 has the parameters $\{1.5\ \mu s, 100\ \text{MHz}, 40\ \text{MHz}, 2\pi/3\}$. LFM_3 has the parameters $\{1.75\ \mu s, 120\ \text{MHz}, 70\ \text{MHz}, 7\pi/10\}$.

*SFM:* SFM_1 has the parameters $\{T_m, f_c, B, \varphi_0\} = \{1.3\ \mu s, 90\ \text{MHz}, 60\ \text{MHz}, \pi/4\}$. SFM_2 has the parameters $\{1.55\ \mu s, 110\ \text{MHz}, 50\ \text{MHz}, \pi/4\}$. SFM_3 has the parameters $\{1.8\ \mu s, 130\ \text{MHz}, 70\ \text{MHz}, \pi/4\}$.

*BPSK:* BPSK_1 has the parameters $\{T_m, f_c, \text{code type}, \varphi_0\} = \{1.25\ \mu s, 70\ \text{MHz}, \text{BARKER-5}, \pi\}$. BPSK_2 has the parameters $\{1.75\ \mu s, 110\ \text{MHz}, \text{BARKER-7}, \pi\}$. BPSK_3 has the parameters $\{1.65\ \mu s, 150\ \text{MHz}, \text{BARKER-11}, \pi\}$.

*FRANK:* FRANK_1 has the parameters $\{T_m, f_c, N_p, \varphi_0\} = \{1.375\ \mu s, 110\ \text{MHz}, 5, 2\pi/5\}$. FRANK_2 has the parameters $\{1.8\ \mu s, 80\ \text{MHz}, 6, \pi/2\}$. FRANK_3 has the parameters $\{1.225\ \mu s, 50\ \text{MHz}, 7, \pi\}$.

To further illustrate the success of the JSLCNN for overlapping signals, multicomponent signals are constructed by overlapping above monocomponent signals. Case 1 is a three-component signal overlapped by LFM_1, LFM_2, and

TABLE III
ESTIMATION PERFORMANCE FOR MONOCOMPONENT SIGNALS

| Signal | Method | -6 dB | -4 dB | -2 dB | 0 dB | 2 dB | 4 dB | 6 dB | 8 dB | 10 dB |
|---|---|---|---|---|---|---|---|---|---|---|
| LFM | JSLCNN | **1.34e-4** | **8.99e-5** | **6.35e-5** | **5.29e-5** | **5.16e-5** | **4.57e-5** | **4.26e-5** | **4.17e-5** | **4.19e-5** |
| | FrFT | 9.15e-3 | 1.05e-4 | 8.15e-5 | 8.37e-5 | 7.40e-5 | 7.33e-5 | 6.87e-5 | 6.77e-5 | 6.81e-5 |
| | Hough | 3.32e-1 | 2.98e-1 | 2.57e-1 | 2.34e-1 | 3.57e-2 | 5.76e-3 | 5.96e-3 | 5.34e-3 | 5.41e-3 |
| SFM | JSLCNN | **3.60e-5** | **2.33e-5** | **1.95e-5** | **1.75e-5** | 1.68e-5 | 1.61e-5 | 1.60e-5 | 1.62e-5 | 1.64e-5 |
| | SVD-MLE | 9.80e-3 | 7.01e-3 | 3.67e-3 | 6.51e-5 | **1.36e-5** | **9.41e-7** | **8.76e-7** | **8.38e-7** | **8.25e-7** |
| | SVD-IRT | 3.15e-3 | 3.74e-5 | 3.19e-5 | 2.86e-5 | 2.82e-5 | 2.71e-5 | 2.68e-5 | 2.66e-5 | 2.55e-5 |
| BPSK | JSLCNN | **6.10e-5** | **4.68e-5** | **3.67e-5** | **3.1e-5** | **2.90e-5** | **2.77e-5** | **2.56e-5** | **2.46e-5** | **2.54e-5** |
| | ZAM | 1.48e-1 | 5.23e-2 | 8.16e-3 | 1.76e-4 | 8.98e-5 | 7.27e-5 | 6.91e-5 | 6.87e-5 | 6.68e-5 |
| FRANK | JSLCNN | **2.43e-4** | **1.50e-4** | **1.02e-4** | **1.08e-4** | **1.04e-4** | **1.00e-4** | **9.69e-5** | **9.45e-5** | **9.41e-5** |
| | FrFT | 4.23e-2 | 2.30e-2 | 1.85e-2 | 1.64e-2 | 1.25e-2 | 8.85e-3 | 4.32e-3 | 2.32e-3 | 1.96e-3 |
| | Hough | 2.91e-1 | 2.36e-1 | 1.09e-1 | 2.44e-2 | 2.28e-2 | 1.68-2 | 5.29e-3 | 9.98e-3 | 8.16e-3 |

TABLE IV
COMPARISON OF TIME COST OF EACH APPROACH

| Approach | JSLCNN | FrFT | Hough | SVD-MLE | IRT | ZAM |
|---|---|---|---|---|---|---|
| Time consuming (s) | **4.21e-2** | 7.07e-1 | 1.34e-1 | 5.33 | 4.34 | 3.87e-1 |

TABLE V
NMSE OF ESTIMATION FOR OVERLAPPING SIGNALS

| Signal | Metric | Method | -2 dB | 0 dB | 2 dB | 4 dB | 6 dB | 8 dB | 10 dB |
|---|---|---|---|---|---|---|---|---|---|
| CASE1 | Fc_NMSE | JSLCNN | **1.37e-4** | **1.02e-4** | **7.90e-5** | **7.29e-5** | **6.99e-5** | **7.72e-5** | **8.11e-5** |
| | | FrFT | 2.79e-3 | 1.56e-3 | 3.99e-4 | 2.26e-4 | 1.27e-4 | 1.32e-4 | 1.21e-4 |
| | B_NMSE | JSLCNN | **1.21e-3** | **9.69e-4** | 6.00e-4 | 4.33e-4 | 3.51e-4 | 3.48e-4 | 2.90e-4 |
| | | FrFT | 4.70e-2 | 1.88e-2 | **9.48e-5** | **7.82e-5** | **5.32e-5** | **5.65e-5** | **4.45e-5** |
| | Tm_NMSE | JSLCNN | **7.54e-4** | **4.67e-4** | 3.32e-4 | 2.64e-4 | 2.27e-4 | 1.89e-4 | 1.82e-4 |
| | | FrFT | 1.88e-2 | 4.35e-3 | **1.03e-5** | **1.29e-5** | **7.47e-6** | **6.19e-6** | **6.08e-6** |
| CASE2 | Fc_NMSE | JSLCNN | **6.44e-5** | **6.08e-5** | **5.29e-5** | **4.96e-5** | **5.15e-5** | **4.99e-5** | 5.02e-5 |
| | | SVD-MLE | 9.04e-4 | 1.09e-3 | 6.99e-4 | 2.49e-4 | 4.46e-3 | 1.43e-3 | **7.40e-7** |
| | B_NMSE | JSLCNN | **3.30e-4** | **2.49e-4** | **1.50e-4** | **1.07e-4** | **1.02e-4** | 7.12e-5 | 6.10e-5 |
| | | SVD-MLE | 2.24e-3 | 2.07e-3 | 1.46e-3 | 4.22e-4 | 2.00e-2 | **2.44e-5** | **5.21e-6** |
| | Tm_NMSE | JSLCNN | **1.37e-4** | **1.39e-4** | **1.55e-4** | **1.75e-4** | 1.88e-4 | 1.89e-4 | 1.86e-4 |
| | | SVD-MLE | 4.14e-3 | 2.65e-3 | 2.17e-4 | 8.47e-4 | **9.16e-5** | **4.81e-7** | **4.44e-7** |
| CASE3 | Fc_NMSE | JSLCNN | **4.42e-5** | **4.41e-5** | **4.28e-5** | **4.17e-5** | **4.11e-5** | **4.05e-5** | **4.05e-5** |
| | Tm/Ts_NMSE | JSLCNN | **1.20e-4** | **6.58e-5** | **3.79e-5** | **2.38e-5** | **1.50e-5** | **1.11e-5** | **8.04e-6** |
| CASE4 | Fc_NMSE | JSLCNN | **9.97e-4** | **3.86e-4** | **2.45e-4** | **1.64e-4** | **1.45e-4** | **1.47e-4** | **1.54e-4** |
| | | FrFT | 1.41e-1 | 1.30e-1 | 9.88e-2 | 1.06e-1 | 8.76e-2 | 8.35e-2 | 8.52e-2 |
| | Tm_NMSE | JSLCNN | **4.79e-4** | **1.96e-4** | **1.53e-4** | **1.08e-4** | **8.97e-5** | **6.22e-5** | **3.73e-5** |
| | | FrFT | 9.18e-2 | 9.09e-2 | 9.90e-2 | 8.91e-2 | 8.37e-2 | 8.58e-2 | 8.33e-2 |
| | Ts_NMSE | JSLCNN | **4.79e-4** | **1.96e-4** | **1.53e-4** | **1.08e-4** | **8.97e-5** | **6.22e-5** | **3.73e-5** |
| | | FrFT | \ | \ | \ | \ | \ | \ | \ |

LFM_3. Case 2 is a three-component signal overlapped by SFM_1, SFM_2, and SFM_3. Case 3 is a three-component signal overlapped by BPSK_1, BPSK_2, and BPSK_3. Case 4 is a three-component signal overlapped by FRANK_1, FRANK_2, and FRANK_3. Case 5 is a four-component signal overlapped by LFM_1, SFM_1, BPSK_1, and FRANK_1.

*4) Performance for Monocomponent Signals:* The average NMSEs for monocomponent LFM ($f_c$, $B$, and $T_m$), SFM ($f_c$, $B$, and $T_m$), BPSK ($f_c$ and $T_s$), and Frank ($f_c$ and $T_m$) signals employing the JSRCNN and baselines are shown in Table III. The performance of monocomponent signals utilizing the developed JSRCNN is overall better than baselines. SVD-mse better estimates the parameters of SFM signals under high SNRs, while the JSLCNN reveals more robustness under intense noise. Table IV shows the time cost of various systems. The total computation time, including the generation of STFT and the inference of the JSLCNN, of the proposed method is $4.21\mathrm{e}^{-2}$ s, which is remarkably fast than all baselines.

To sum up, the designed JSLCNN provides a generic parameter estimation method for multiple types of signals and achieves outstanding overall results compared with algorithms specially designed for certain types of signals. Furthermore, the JSLCNN gives a significant real-time performance boost.

*5) Performance for Overlapping Signals:* To further examine the capabilities of the JSLCNN for overlapping signals, we have carried out 100 Monte Carlo experiments from $-2$ to 10 dB at intervals of 2 dB for overlapping LPI radar signals.

Table V shows the estimation performance for three-component signals. Compared to FrFT, the JSRCNN is superior in computing $f_c$ and more robust in estimating $T_m$ and $B$. In addition, FrFT needs to execute three times estimations for signals with various chirp rates, while the JSLCNN still maintains the same calculation time for overlapping signals as that for single-component signal. Hough will fail in estimating parameters of overlapping signals because multicomponent cross-product terms in PWVD restrict its applications.

TABLE VI
NMSE OF ESTIMATION FOR OVERLAPPING SIGNALS CASE 5

| Constituent | Metric | Method | -2 dB | 0 dB | 2 dB | 4 dB | 6 dB | 8 dB | 10 dB |
|---|---|---|---|---|---|---|---|---|---|
| LFM | Fc_NMSE | JSLCNN | 6.67e-4 | 4.08e-4 | 1.41e-4 | 1.42e-4 | 1.45e-4 | 1.54e-4 | 1.91e-4 |
| | B_NMSE | | 2.61e-3 | 2.00e-3 | 2.06e-3 | 1.46e-3 | 1.68e-3 | 1.49e-3 | 1.44e-3 |
| | Tm_NMSE | | 1.22e-3 | 4.62e-4 | 4.92e-4 | 4.04e-4 | 4.38e-4 | 6.17e-4 | 8.05e-4 |
| SFM | Fc_NMSE | JSLCNN | 1.31e-4 | 1.02e-4 | 6.79e-5 | 8.31e-5 | 6.62e-5 | 6.93e-5 | 7.44e-5 |
| | B_NMSE | | 1.91e-3 | 9.92e-4 | 5.18e-4 | 3.42e-4 | 2.36e-4 | 1.98e-4 | 1.42e-4 |
| | Tm_NMSE | | 5.53e-4 | 3.94e-4 | 4.09e-4 | 5.46e-4 | 5.86e-4 | 6.57e-4 | 7.67e-4 |
| FRANK | Fc_NMSE | JSLCNN | 2.51e-3 | 1.23e-3 | 7.55e-4 | 3.38e-4 | 2.78e-4 | 3.24e-4 | 2.58e-4 |
| | Tm/Ts_NMSE | | 4.79e-4 | 2.33e-4 | 1.58e-4 | 5.02e-5 | 3.94e-5 | 3.34e-5 | 2.80e-5 |
| BPSK | Fc_NMSE | JSLCNN | 4.70e-1 | 3.00e-1 | 4.01e-2 | 9.68e-5 | 1.01e-2 | 9.54e-5 | 9.48e-5 |
| | Tm/Ts_NMSE | | 4.74e-1 | 3.05e-1 | 4.29e-2 | 1.14e-3 | 1.07e-2 | 1.16e-4 | 5.92e-5 |

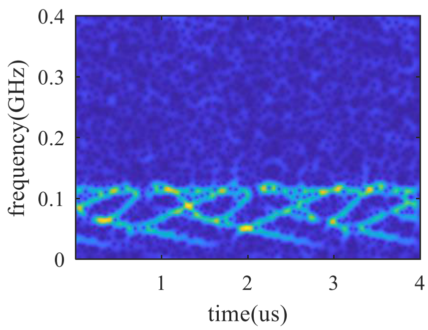


Fig. 11. TFI of the four-component signal at the SNR of 10 dB.

Estimation-and-elimination-based SVD-MLE leads to better recognition under high SNR, while the JSLCNN performs more robustly. Computational efficiency is another advantage for the JSLCNN compared with cyclically executed SVD-MLE. For SVD-IRT, signal energy distribution must be in the center of TFI. However, for overlapping SFM signals, especially in the case of frequency aliasing, it is difficult to use SVD-IRT to calculate $f_c$ of each component. ZAM suffers difficulties in locating $f_c$ for each component and can hardly estimate the parameter of overlapping BPSK signals. The proposed JSLCNN maintains an NMSE less than $4.5e^{-5}$ from $-2$ to 10 dB for the estimation of $f_c$. The JSLCNN can also compute $T_m$ with an NMSE less than $6.58e^{-5}$ when the SNR is over 0 dB, which is beyond the capability of ZAM. The JSLCNN outperforms FrFT for all SNRs when computing parameters of three-component signals overlapped by Frank signals.

Table VI shows the estimation performance for four-component signals in Case 5. The TFI of the overlapping signal at the SNR of 10 dB is shown in Fig. 11. The energy distributions of the four components are aliased together, which dramatically increases the analysis difficulty. Facing these overlapping signals, all baselines fail, while the proposed JSLCNN can still calculate parameters accurately. Parameter estimation results for each component are nearly positively related to SNRs. Estimation performances for SFM, BPSK, and Frank signals are better than that of LFM signals. This is mainly because the energy distribution of the LFM signals suffers from more sabotage by other components.

Generally, the JSLCNN provides a better estimation for overlapping signals while maintaining the same computational time as single-component signals. Desirable estimation performance and outstanding real-time capability further validate the superiority of the designed JSLCNN.

*6) Estimation Performance Against PR:* Given the masking problem between strong and weak signals, estimation experiments on two-component signals with different PRs are carried out to verify the robustness of the JSLCNN. LFM_1, SFM_1, BPSK_1, and FRANK_1 are mixed in pairs to form six sets of two-component signals. To better investigate the influence of PR, the SNR is fixed at 20 dB to suppress the effect of noise. The variation range of PR is set from $-9$ to 9 dB at intervals of 3 dB. The NMSE for signal components that are not detected will be set to 1; 100 Monte Carlo experiments are conducted under each PR. It can be seen from Table VII that as the power difference between components gradually increases, the estimation results will also become worse. Overall, the system can maintain desirable estimation performance when the PR is from $-6$ to 6 dB. Experimental results show that the proposed method can deal with the masking problem to a certain degree.

## C. Ablation Studies

This section illustrates the validity of the FRB and the necessity of DJL. The testing data are the same as in Section IV-A3. In the JSLCNN, the loss weight of FRB is set as [0.83, 0.5 0.33, 0.16, 0.09], which is adjusted every ten epochs. The loss weights of FRB and the analysis network are equal in the JSLCNN without DJL. The recognition performance in the absence of FRB and DJL is shown in Fig. 12. The best performance is degraded when FRB or DJL is excluded. It can be observed that DJL makes a more positive impact on the recognition results, boosting about 1%–2%, while FRB seems slightly weaker. This is because the equal weights of FRB and the analysis network make the network learn two objectives at the same time instead of focusing more on analysis.

Fig. 13 presents the estimation performance for four-component signals Case 5 without FRB and DJL. Results show that FRB brings about a more robust estimation even in the absence of DJL. FRB can remove the destructive influence of noise and provide a purer feature representation, contributing to boundary regression. DJL further enhances the robustness of the estimation by allocating more attention to the analysis subnet. In addition, the recovery subnetwork is activated only

TABLE VII
NMSE OF ESTIMATION AGAINST PR

| PR | -9 dB | -6 dB | -3 dB | 0 dB | 3 dB | 6 dB | 9 dB |
|---|---|---|---|---|---|---|---|
| NMSE | 4.20e-2 | 1.88e-4 | 1.05e-4 | 9.20e-5 | 1.05e-4 | 1.97e-4 | 4.20e-2 |

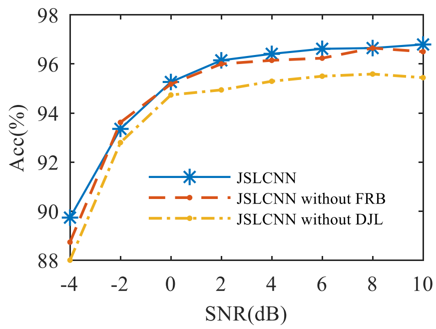


Fig. 12. Recognition performance in the absence of FRB and DJL.

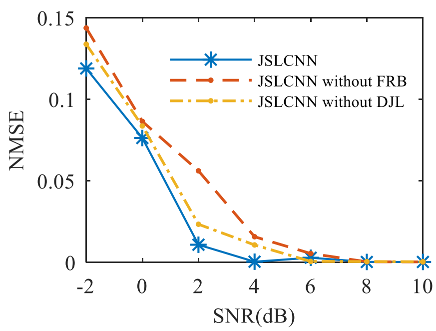


Fig. 13. Estimation performance in the absence of FRB and DJL.

in the training process to assist in the generation of clean features, which does not add any parameters and computation costs in the test phase. Namely, the above performance boost is realized at no cost.

## V. CONCLUSION

Analysis of each constituent in the overlapping signal is a crucial ingredient in modern electronic reconnaissance. This article reframes overlapping LPI signal analysis as a semantic learning task, straight from TFIs to parameter estimation and modulation classification based on shared features. Meanwhile, a feature recovery subnetwork is jointly optimized to enhance the clarity of features at no computational cost.

Extensive experiments compared with other modulation identification and parameter estimation approaches have been carried out to reveal the validity and superiority of the developed JSLCNN. According to the results of comparative experiments, the following conclusions can be drawn. 1) the JSLCNN has desirable extensibility to recognize unseen signal combinations, eliminating dependence on covering all possible signal combinations; 2) the JSLCNN avoids performance degradation under unknown jamming and the robustness to unknown interferences makes the JSLCNN more practical.; and 3) the regression-based JSLCNN realizes a better real-time and generalized performance on parameter estimation.

Nevertheless, further work is needed to improve the proposed JSLCNN from the following aspects.

1) JSLCNN relies on considerable training samples. The improved method driven by the small sample data should be explored.
2) JSLCNN can reject several unknown jamming signals. However, modern electronic reconnaissance also requires the system to identify unknown categories when the corresponding labels are progressively received. Further research on incrementally learning of recognition systems should be investigated.
3) For the practical deployment of the JSLCNN, we will attempt to validate the developed approach on deep learning chips and hardware platforms.

## REFERENCES


[1] C. Aydogdu et al., "Radar interference mitigation for automated driving: Exploring proactive strategies," *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 72–84, Jul. 2020.
[2] D. Ma et al., "Joint radar-communication strategies for autonomous vehicles: Combining two key automotive technologies," *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 85–97, Jul. 2020.
[3] J. Xu et al., "Focus-before-detection radar signal processing: Part I—Challenges and methods," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 32, no. 9, pp. 48–59, Sep. 2017.
[4] G. Hu, L. Xu, P. Zhao, and W. Wang, "Confidence evaluation of BPSK signal analysis based on EVT," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 1515–1537, Apr. 2020.
[5] M. Greco, F. Gini, and A. Farina, "Radar detection and classification of jamming signals belonging to a cone class," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1984–1993, May 2008.
[6] S. Huang, Y. Yao, Z. Wei, Z. Feng, and P. Zhang, "Automatic modulation classification of overlapped sources using multiple cumulants," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6089–6101, Jul. 2017.
[7] H. Zhu, S. Zhang, and H. Zhao, "Single-channel source separation of multi-component radar signal based on EVD and ICA," *Digit. Signal Process.*, vol. 57, pp. 93–105, Oct. 2016.
[8] V. Gouldieff, J. Palicot, and S. Daumont, "Blind modulation classification for cognitive satellite in the spectral coexistence context," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3204–3217, Jun. 2017.
[9] H. Zhu, S. Zhang, and H. Zhao, "Single channel source separation and parameter estimation of multi-component PRBCPM-SFM signal based on generalized period," *Digit. Signal Process.*, vol. 40, pp. 224–237, May 2015.
[10] Q. Qu, S. Wei, S. Liu, J. Liang, and J. Shi, "JRNet: Jamming recognition networks for radar compound suppression jamming signals," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15035–15045, Dec. 2020.
[11] M. D. M. Zhang and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE Access*, vol. 5, pp. 11074–11082, 2017.
[12] K. Chen, L. Zhu, S. Chen, S. Zhang, and H. Zhao, "Deep residual learning in modulation recognition of radar signals using higher-order spectral distribution," *Measurement*, vol. 185, Nov. 2021, Art. no. 109945.
[13] Y. Dong, X. Jiang, H. Zhou, Y. Lin, and Q. Shi, "SR2CNN: Zero-shot learning for signal recognition," *IEEE Trans. Signal Process.*, vol. 69, pp. 2316–2329, 2021.
[14] L. M. Hoang, M. Kim, and S.-H. Kong, "Automatic recognition of general LPI radar waveform using SSD and supplementary classifier," *IEEE Trans. Signal Process.*, vol. 67, no. 13, pp. 3516–3530, Jul. 2019.
[15] K. Chen, J. Zhang, S. Chen, S. Zhang, and H. Zhao, "Automatic modulation classification of radar signals utilizing X-Net," *Digit. Signal Process.*, vol. 123, Apr. 2022, Art. no. 103396.

[16] K. K. Guner et al., "FPGA-based Wigner–Hough transform system for detection and parameter extraction of LPI radar LFMCW signals," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 2003515.
[17] T. O. Gulum, A. Y. Erdogan, T. Yildirim, and L. Durak Ata, "Parameter extraction of FMCW modulated radar signals using Wigner–Hough transform," in *Proc. IEEE 12th Int. Symp. Comput. Intell. Informat. (CINTI)*, Nov. 2011, pp. 465–468.
[18] A. Y. Erdogan, T. O. Gulum, L. Durak-Ata, T. Yildirim, and P. E. Pace, "FMCW signal detection and parameter extraction by cross Wigner–Hough transform," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 1, pp. 334–344, Feb. 2017.
[19] Z. Wang, Y. Wang, and L. Xu, "Parameter estimation of hybrid linear frequency modulation-sinusoidal frequency modulation signal," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1238–1241, Aug. 2017.
[20] L. Li, T. Niu, H. Ji, H. Han, and Y. Liu, "Accurate parameter estimation of chirp class signals under low SNR," in *Proc. IEEE Int. Conf. Signal Image Process. (ICSIP)*, Aug. 2016, pp. 412–416.
[21] X. Wang, G. Zhang, and R. Zhou, "Detection and parameter estimation of polyphase codes radar signal based on fractional autocorrelation," in *Proc. Int. Conf. Electron. Sci. Autom. Control*, 2015, pp. 12–15.
[22] X. Zeng, B. Tang, and Y. Xiong, "Interception algorithm of S-cubed signal model in stealth radar equipment," *Chin. J. Aeronaut.*, vol. 25, no. 3, pp. 416–422, Jun. 2012.
[23] S. Huang, Y. Jiang, X. Qin, Y. Gao, Z. Feng, and P. Zhang, "Automatic modulation classification of overlapped sources using multi-gene genetic programming with structural risk minimization principle," *IEEE Access*, vol. 6, pp. 48827–48839, 2018.
[24] M. Zaerin and B. Seyfe, "Multiuser modulation classification based on cumulants in additive white Gaussian noise channel," *IET Signal Process.*, vol. 6, no. 9, pp. 815–823, Dec. 2012.
[25] P. Jain and A. Kapoor, "Active learning for large multi-class problems," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 762–769.
[26] M.-L. Zhang and Z.-H. Zhou, "A review on multi-label learning algorithms," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 8, pp. 1819–1837, Jun. 2014.
[27] M. Zhu, Y. Li, Z. Pan, and J. Yang, "Automatic modulation recognition of compound signals using a deep multi-label classifier: A case study with radar jamming signals," *Signal Process.*, vol. 169, Apr. 2020, Art. no. 107393.
[28] W. Si, C. Wan, and Z. Deng, "Intra-pulse modulation recognition of dual-component radar signals based on deep convolutional neural network," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3305–3309, Oct. 2021.
[29] Y. Hong-Hai, Y. Xiao-Peng, L. Shao-Kun, L. Ping, and H. Xin-Hong, "Radar emitter multi-label recognition based on residual network," *Defence Technol.*, vol. 18, no. 3, pp. 410–417, Mar. 2022.
[30] Z. Pan, S. Wang, M. Zhu, and Y. Li, "Automatic waveform recognition of overlapping LPI radar signals based on multi-instance multi-label learning," *IEEE Signal Process. Lett.*, vol. 27, pp. 1275–1279, 2020.
[31] L. Stankovic et al., "Inverse radon transform–based micro-Doppler analysis from a reduced set of observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 2, pp. 1155–1169, Apr. 2015.
[32] M. Dakovic and L. Stankovic, "Estimation of sinusoidally modulated signal parameters based on the inverse radon transform," in *Proc. 8th Int. Symp. Image Signal Process. Anal. (ISPA)*, Sep. 2013, pp. 302–307.
[33] L. Guo, Y. Hu, S. Xu, X. Dong, and M. Li, "Separate estimation schemes for time-frequency overlapped micro-Doppler signal in single channel infrared laser detection," *Infr. Phys. Technol.*, vol. 94, pp. 134–141, Nov. 2018.
[34] A. Bouchikhi, A.-O. Boudraa, J.-C. Cexus, and T. Chonavel, "Analysis of multicomponent LFM signals by teager Huang–Hough transform," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 1222–1233, Apr. 2014.
[35] A. Serbes, "On the estimation of LFM signal parameters: Analytical formulation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 2, pp. 848–860, Apr. 2018.
[36] O. Aldimashki and A. Serbes, "Performance of chirp parameter estimation in the fractional Fourier domains and an algorithm for fast chirp-rate estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 5, pp. 3685–3700, Oct. 2020.
[37] Z. Xie, J. Gao, K. Wu, and J. Zhang, "Brief survey on image semantic analysis and understanding," in *Proc. Int. Conf. Soft Comput. Pattern Recognit. (SoCPaR)*, Oct. 2011, pp. 179–183.
[38] D. Freedman, *Statistical Models: Theory and Practice*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[39] D. H. Vu, K. M. Muttaqi, A. P. Agalgaonkar, and A. Bouzerdoum, "Short-term forecasting of electricity spot prices containing random spikes using a time-varying autoregressive model combined with kernel regression," *IEEE Trans. Ind. Informat.*, vol. 15, no. 9, pp. 5378–5388, Sep. 2019.
[40] J.-S. Chou and T.-K. Nguyen, "Forward forecast of stock price using sliding-window Metaheuristic-optimized machine-learning regression," *IEEE Trans. Ind. Informat.*, vol. 14, no. 7, pp. 3132–3142, Jul. 2018.
[41] K. Liu, J. Ye, Y. Yang, L. Shen, and H. Jiang, "A unified model for joint normalization and differential gene expression detection in RNA-seq data," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 16, no. 2, pp. 442–454, Mar. 2019.
[42] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.
[43] J. Yu et al., "UnitBox: An advanced object detection network," in *Proc. 24th ACM Int. Conf. Multimedia*, 2016, pp. 516–520.
[44] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO series in 2021," 2021, *arXiv:2107.08430*.
[45] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 936–944.
[46] R. Caruana, "Multitask learning," *Mach. Learn.*, vol. 28, no. 1, pp. 41–75, Dec. 1997.
[47] S. Huang, T. Le, and D. Jaw, "DSNet: Joint semantic learning for object detection in inclement weather conditions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 8, pp. 2623–2633, Aug. 2021.
[48] H. Zhang, V. Sindagi, and V. M. Patel, "Multi-scale single image dehazing using perceptual pyramid deep network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 1015–101509.
[49] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.
[50] X. Zhu and M. Bain, "B-CNN: Branch convolutional neural network for hierarchical classification," 2017, *arXiv:1709.09890*.
[51] T. Haifeng, L. Wei, W. Tan, F. Jian, and F. Zhiyong, "The analysis on the candidate frequency bands of future mobile communication systems," *China Commun.*, vol. 12, pp. 140–149, Dec. 2015.
[52] P. E. Pace, *Detecting and Classifying Low Probability of Intercept Radar*. Norwood, MA, USA: Artech House, 2009.
[53] E. Sejdić, I. Djurović, and J. Jiang, "Time–frequency feature representation using energy concentration: An overview of recent advances," *Digit. Signal Process.*, vol. 19, no. 1, pp. 153–183, 2009.
[54] B. Boashash, "Note on the use of the Wigner distribution for time–frequency signal analysis," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-36, no. 9, pp. 1518–1521, Sep. 1988.
[55] H.-I. Choi and W. J. Williams, "Improved time–frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, Jun. 1989.
[56] S. Ren et al., "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1–9.
[57] A. Neubeck and L. Van Gool, "Efficient non-maximum suppression," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, 2006, pp. 850–855.

**Kuiyu Chen** was born in Huaian, Jiangsu, China, in 1995. He is currently pursuing the Ph.D. degree in electronic and optical engineering with the Nanjing University of Science and Technology, Nanjing, China.

He is also a Visiting Ph.D. Student with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include blind signal separation, automatic modulation recognition, and jamming technology.

**Lipo Wang** (Senior Member, IEEE) received the bachelor's degree from the National University of Defense Technology, Changsha, China, in 1983, and the Ph.D. degree from Louisiana State University, Baton Rouge, LA, USA, in 1988.

He is currently a Faculty Member with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He has published more than 340 publications. He has coauthored two monographs and coedited 15 books. He holds a U.S. patent in neural networks and a patent in systems. He has more than 8800 Google scholar citations, with H-index of 43. His research interests include artificial intelligence with applications to image/video processing, biomedical engineering, and data mining.

Dr. Wang was a member of the Board of Governors of the International Neural Network Society, the IEEE Computational Intelligence Society (CIS), and the IEEE Biometrics Council. He received the APNNA Excellent Service Award. He was the Founding Chair of the EMBS Singapore Chapter and the CIS Singapore Chapter. He serves/served as the chair/committee member for over 200 international conferences. He was a keynote speaker for 36 international conferences. He served as the CIS Vice President for Technical Activities and the Chair for the Emergent Technologies Technical Committee, as well as the Chair for the Education Committee of the IEEE Engineering in Medicine and Biology Society (EMBS). He is/was an associate editor/an editorial board member of 30 international journals, including four IEEE TRANSACTIONS, and a guest editor for ten journal special issues. He was the President of the Asia–Pacific Neural Network Assembly (APNNA).

**Jingyi Zhang** (Student Member, IEEE) was born in Xuzhou, Jiangsu, China, in 1995. She received the B.S. degree in electronic and information engineering from the Nanjing University of Science and Technology, Nanjing, China, in 2017, where she is currently pursuing the Ph.D. degree in electronic and optical engineering.

Her research interests include communication engineering.

**Si Chen** (Member, IEEE) was born in Xuancheng, Anhui, China, in 1988. He received the B.S. and M.S. degrees in communication and information engineering from the Nanjing University of Science and Technology, Nanjing, China, in 2010 and 2015, respectively.

He is currently an Associate Professor with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology. His research interests include missile-borne synthetic aperture radar (SAR) imaging, target detection, and signal processing.

**Shuning Zhang** was born in Weihai, China, in 1977. She received the Ph.D. degree in signal and information processing from the Nanjing University of Science and Technology, Nanjing, China, in 2006.

She is currently a Professor at the School of Electronic and Optical Engineering, Nanjing University of Science and Technology. Her research interests include nonstationary signal processing, parameters extraction, and multifractional signal analysis.