Parameter Extraction Method of Overlapping Radar Signals Using Modulation Recognition-Guided Semantic Segmentation

Weibo Huo, *Member, IEEE*, Yang Luo, *Graduate Student Member, IEEE*, Hao Wang, Jifang Pei, *Member, IEEE*, Yin Zhang, *Member, IEEE*, Yulin Huang, *Senior Member, IEEE*, and Jianyu Yang, *Member, IEEE*

***Abstract*—Parameter extraction of radar signals is an important but challenging task in electronic warfare. In the modern electromagnetic environment, the radiation sources greatly increase, causing different radar signals to overlap, making the parameter extraction of radar signals difficult. Meanwhile, using radar signal parameter extraction methods that are not suitable for dealing with overlapping signals can lead to serious errors in this case. To address this, we propose a parameter extraction network for overlapping radar signals using modulation recognition-guided semantic segmentation. Specifically, we first design an encoder–decoder to segment overlapping radar signals, which uses channel rearrangement and modulation type filtering to increase the accuracy of segmentation. In this encoder–decoder, channel rearrangement is an optimization of convolution operation, aiming to increase the perceptual field while reducing feature information loss. And modulation type filtering can convert the results of semantic segmentation into masks corresponding to each radar signal, increasing the accuracy of segmentation. After the encoder–decoder, signal segmentation masks are obtained. Then, we compress these segmentation masks in the time and frequency domains, and extract the span of them to achieve accurate extraction of the pulsewidth and bandwidth of each radar signal. The experiments validate the feasibility of the proposed method.**

***Index Terms*—Modulation recognition, overlapping radar signals, parameter extraction, semantic segmentation.**

I. Introduction

PARAMETER extraction of radar signals, as an important part of electronic reconnaissance [1], refers to the analysis and processing of the received radar signals to extract signal parameters such as pulsewidth and bandwidth [2]. Parameter extraction of radar signals is used in various fields, such as signal sorting [3], jamming decision optimization [4], etc. However, the parameter extraction of radar signals in complex electromagnetic environments, especially with overlapping signals, can be challenging [5]. Therefore, the study of overlapping radar signal parameter extraction is of significant practical application.

Traditional parameter extraction methods for radar signals are mainly divided into maximum likelihood estimation [6], Fourier transform [7], time–frequency analysis method [8], etc. These methods are computationally simple to use. The parameter extraction method based on maximum likelihood estimation generally includes: modeling the system as a statistical model, deriving the maximum likelihood estimation of the parameter, and providing compact expressions of the Cramer–Rao bound [9], [10], [11]. However, this stochastic modeling method usually requires linear, stationary, and Gaussian assumptions. The signals in the modern electromagnetic environment are usually nonlinear, nonstationary, and non-Gaussian, when the actual scenario does not meet its assumptions, the accuracy of maximum likelihood estimation greatly decreases. Fourier transform-based method is a non-time–frequency analysis method, which converts time-domain signals into frequency-domain signals for signal analysis [12], [13]. This method utilizes the Fourier transform to extract frequency-domain parameters such as bandwidth [14]. However, the Fourier transform-based method does not consider the time domain information of the signal, which is not conducive to improving signal parameter extraction performance with time–frequency joint information.

The time–frequency characteristics can accurately reflect the frequency variation of the signal with time, which are useful for the parameter extraction of traditional radar signals [8]. The radar parameter extraction method based on time–frequency analysis usually consists of time–frequency diagram generation and parameter calculation [15], [16], [17]. First, the time–frequency diagram generation can be achieved using short-time Fourier transform (STFT) [18], Wigner-Ville distribution (WVD) [19], etc. Zhang et al. [20] observe that STFT calculation is fast and easy to implement, but a tradeoff exists between time resolution and frequency resolution. And improved methods for generating time–frequency diagrams have been studied, such as [21] and [22]. Then, the main idea of parameter calculation is extracting parameters from the time–frequency diagram using numerical methods such as inherit accumulation [23], scalar

Manuscript received 30 June 2023; revised 12 December 2023; accepted 25 January 2024. Date of publication 5 February 2024; date of current version 15 February 2024. This work was supported by the National Natural Science Foundation of China under Grant 61901091 and Grant 61901090. *(Corresponding author: Jifang Pei.)*

The authors are with the Department of Electrical Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China (e-mail: hwbuyi@uestc.edu.cn; 302920402@qq.com; 15320698817@163.com; peijfstudy@126.com; yinzhang@uestc.edu.cn; yulinhuang@uestc.edu.cn; jyyang@uestc.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3361905

transform [20], etc. The aforementioned time–frequency analysis methods are applicable to traditional radar signals. However, in complex electromagnetic environments, radar signals employ various modulation techniques, which are increasingly intricate and varied, causing misjudgment during the demodulation process and making it difficult to accurately distinguish different modulation types [24]. Moreover, the noise interference present in the contemporary electromagnetic environment is more complex, encompassing diverse forms of electromagnetic interference, multipath propagation interference, Doppler effect [25], etc. As a result, traditional time–frequency analysis methods face significant challenges in accurately extracting signal parameters.

With the great development of deep network technology, some studies attempt to improve the performance of parameter extraction of radar signals using deep networks. Parashar et al. [26] propose a method for learning micro-Doppler features using autoencoders, which addresses the issues of high computational complexity and slow response speed. Lv et al. [27] propose a CNN-based jamming recognition and antijamming target detection algorithm that can extract the real targets' range and velocity, which has higher accuracy and robustness compared to traditional methods. Su et al. [28] present a method to extract the central frequency of a linear frequency modulation (LFM) signal by inputting a WVD time–frequency diagram into CNN, which solves the problems of low accuracy and poor robustness in traditional parameter extraction methods. Besides these methods, support vector machine [29], modified CNN [30], [31], domain adaptation [32], and transformer [33] are also employed to solve the parameter extraction problem of radar signals. These parameter extraction methods are effective for individual radar signals. However, when radar signals overlap, parameters such as the bandwidth and amplitude of the signals change, leading to contamination of individual radar signal features. Consequently, using these methods in overlapping situations may lead to serious errors in parameter extraction. Therefore, it is necessary to propose a parameter extraction method for overlapping radar signals.

In the case of parameter extraction of overlapping radar signals, it is necessary to first segment the overlapping radar signals before extracting parameters. Wang et al. [34] propose a squeeze-and-excitation Inception net with adaptive thresholds to segment overlapping radar signals, which has a high recognition rate and improved performance compared to traditional methods. Chen et al. [35] develop a joint semantic learning CNN to do semantic segmentation on low probability of intercept (LPI) radar signals, which has higher separation and recognition accuracy than traditional methods. Besides these methods, attention [36] and reinforcement learning [37] are also employed to segment overlapping radar signals. Although some research has been conducted on the segmentation of overlapping radar signals, the aforementioned algorithms do not incorporate parameter extraction into overlapping radar signal segmentation, so the aforementioned algorithms cannot solve the problem of parameter extraction of overlapping radar signals. Therefore, further research is needed to combine deep networks with parameter extraction of overlapping radar signals, especially in the cases of dense radar signal pulse streams. To handle this issue, we propose a parameter extraction network using modulation recognition-guided semantic segmentation for extracting parameters of overlapping radar signals. Specifically, we use modulation recognition-guided semantic segmentation to segment the overlapping radar signals, and then, use a parameter extraction module to extract the pulsewidths and bandwidths of the overlapping radar signals from the segmentation results.

This article's primary contributions are outlined as follows.

1) We propose a parameter extraction network of overlapping radar signals that combines signal segmentation and parameter extraction, which can effectively separate overlapping radar signals and obtain the pulsewidth and bandwidth of each signal.
2) The channel rearrangement and modulation type filtering are proposed and integrated into the semantic segmentation. The channel rearrangement layer can effectively expand the perception field, and decrease the feature information loss. And modulation type filtering can increase the accuracy of segmentation.

Meanwhile, we conduct comparative experiments, ablation experiments, and experiments on different modulation types to validate the efficacy of the proposed method.

The rest of this article is organized as follows. Section II examines the characteristics of overlapping radar signals. Section III offers a thorough explanation of the proposed parameter extraction network. In Section IV, we evaluate and discuss the performance of our method using a generated dataset. Section V discusses the performance and influence factors of the proposed method, and the limitations are listed and future work is outlined. Finally, Section VI concludes this article.

II. Problem Formulation

This section provides a concise overview of overlapping radar signals and the parameter extraction method after semantic segmentation, aiming to describe the method for extracting parameters of overlapping radar signals.

A. Analysis of Overlapping Radar Signals

Overlapping radar signals refer to the reception of more than one radar signal within the sample period. The schematic diagram of an overlapping signal consisting of two radar pulse signals is shown in Fig. 1.

The length of overlapping time is often not equal to the total time length of the signal. Therefore, the overlap degree is worth studying for overlapping radar signals and can be used for dataset generation. The overlap degree (Od) of the signal is calculated as

$$\mathrm{Od} = \frac{nt_0}{\sum_{i=1}^{n} t_i} \tag{1}$$

where t_0 represents the overlapping duration, n represents the number of the radar signals, and $t_i\{i = 1, 2, \ldots, n\}$ represents the pulsewidth of the ith radar signal.

In addition, the amplitudes of two radar signals are often not consistent, and signals with strong energy may drown out radar

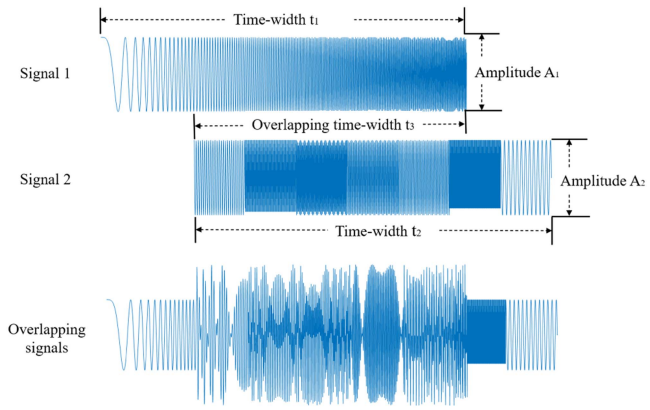

Fig. 1. Schematic diagram of overlapping signals. The received overlapping radar signal is formed by signal 1 and signal 2, with a time width of t_1 and an amplitude of A_1 for signal 1; the time width of signal 2 is t_2 and the amplitude is A_2; the overlapping time width is t_3.

signals with low energy. Therefore, when analyzing overlapping radar signals, it is necessary to limit the amplitudes of the two radar signals to a certain range. The signal power ratio (Pr) is used to characterize the energy intensity ratio of overlapping signals, and its expression is

$$\mathrm{Pr} = 20 \log \left(\frac{A_1}{A_2} \right) \tag{2}$$

where A_1 represents the amplitude of one signal, and A_2 represents the amplitude of the other.

B. Parameter Extraction Problem of Overlapping Radar Signals

This article examines the extraction of bandwidth and pulsewidth from overlapping two radar signals. The pulsewidth and bandwidth of a single radar signal can be obtained from its time–frequency diagram. The time–frequency diagram's horizontal axis projection represents the signal's temporal width, that is, the pulsewidth; while the vertical axis projection represents the signal's frequency width, that is, the bandwidth. By analyzing the time–frequency diagram, we can extract information about a signal's pulsewidth and bandwidth.

If radar signals are overlapping to a certain extent in both the time and frequency domains, it can be difficult to obtain each signal's distribution directly. Consequently, it is necessary to separate overlapping radar pulse signals first, and then, extract the parameters of each signal individually. In this article, the separation process of the overlapping radar signals is implemented by modulation recognition-guided semantic segmentation. As a typical computer vision problem [38], semantic segmentation aims to classify each pixel of input image data. By semantic segmentation of the time–frequency diagrams of received overlapping signals, all pixels can be divided into background and individual signals. Through the semantic segmentation of the time–frequency diagram, different modulated radar signals can be separated.

The segmentation results may contradict the ground truth due to the lack of guidance from physical prior information [39]. To avoid this situation, we incorporated modulation type prior information at the end of the segmentation process, which we refer to as "modulation recognition guided semantic segmentation." The prior knowledge about modulation types provides essential understanding of the expected characteristics of radar signals, which can be utilized to guide the segmentation process. Specifically, different modulation types of radar signals exhibit distinct features in the time–frequency diagram. When two signals with similar local features overlap, it becomes challenging to accurately segment them at the overlap. By filtering the prior information about modulation types, the segmentation algorithm can effectively differentiate and localize the signals based on their anticipated modulation characteristics, thus enhancing the accuracy and precision of the segmentation results.

This article considers six modulated radar signals two-by-two overlapping: single carrier (SC) signal, LFM signal, nonlinear frequency modulation (NLFM) signal, frequency agility (FA) signal, Costas frequency encoded signal, and P4 phase-encoded signal, and the schematic diagram of parameter extraction through semantic segmentation is shown in Fig. 2.

The problem of semantic segmentation of overlapping radar signals has two main aspects [40]. First, the local segmentation proble, refers to the accurate pixel-level segmentation of specific signals within a time–frequency diagram, typically used to describe the texture, edges, corners, and other local features. This involves precise segmentation of the boundaries and details of each radar signal, requiring the model to capture fine features and edge information. Second, the global segmentation problem, which can help the model understand the overall semantic content of the time–frequency diagram, such as identifying background, scenes, and the relationships between different objects. It requires the model to have a holistic understanding of the entire time–frequency diagram and accurately segment different radar signals and background. To address the local segmentation problem, this article uses channel rearrangement instead of traditional pooling operations to retain more signal features. In addition, the jump connections between the encoder and decoder pass high-resolution features directly to the decoder, enhancing local segmentation. To tackle the issue of global segmentation, a global feature extraction layer is incorporated to extract signal features on a larger scale. In addition, modulation recognition results are used in segmentation to optimize the global segmentation effect. Later on, we will delve into these methods with more comprehensive explanations.

III. Methodology

This section covers the discussion of the proposed parameter extraction network, starting with an explanation of the time–frequency analysis method. Then, we present the architecture of the proposed parameter extraction network and provide details of each module in the proposed network.

A. Time–Frequency Analysis

The input for the suggested parameter extraction network is the time–frequency diagram derived from overlapping radar signals. Therefore, this subsection will present an introduction

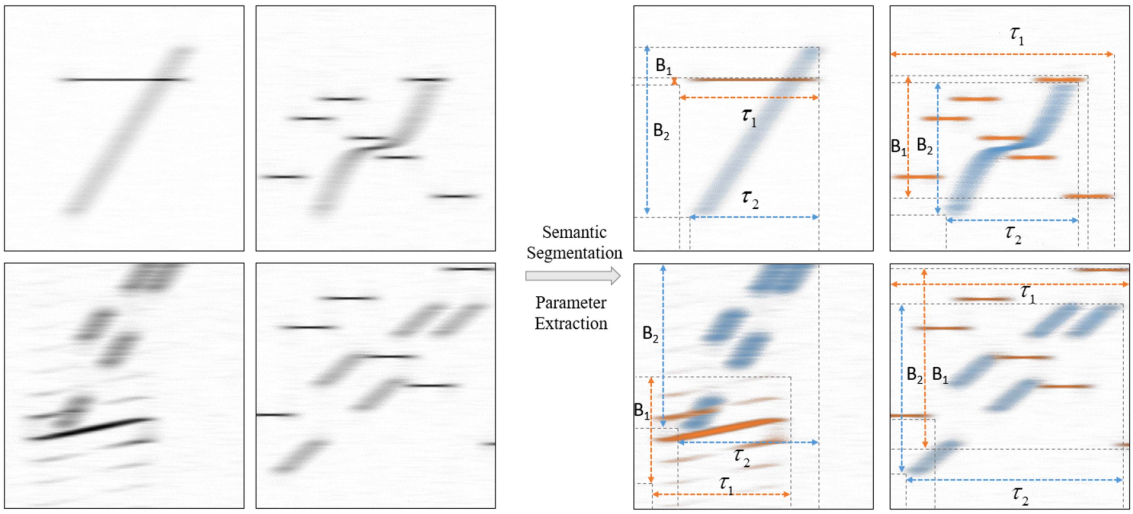

Fig. 2. Schematic diagram for extracting overlapping signal parameters through semantic segmentation. Overlapping radar signals can be separated into signals marked orange and blue through semantic segmentation steps. Then, the bandwidth parameters (B_1 and B_2) and the pulsewidth parameters (τ_1 and τ_2) of each signal can be extracted separately by projecting the signal on the vertical and horizontal axes of the time–frequency diagram.

to the time–frequency analysis method used in this article. For radar signals, common time–frequency transformation methods include STFT, WVD, and synchronous compression short-time Fourier transform (FSST). Besides, compared with STFT, FSST can make time–frequency analysis more precise [41]. Compared with WVD, FSST does not have cross-interference terms, which is extremely important when studying overlapping radar signals. Therefore, this article will use FSST to extract the time–frequency features of radar signals. FSST is developed based on synchronous compressed wavelet transforms (SSWT), and its core idea is to rearrange the results of STFT, which can compress the scattered energy distribution to near the true frequency. The expression for FSST is as

$$\mathrm{FSST}(t,\omega)=\frac{1}{g(0)}\int V_f(\eta,t)\delta\left(\omega-\widehat{\omega}_f(\eta,t)\right)d\eta \quad (3)$$

where $g(0)$ represents the value of the sliding window $g(t)$ at 0, $\delta(\cdot)$ represents the Dirichlet function, $V_f(\eta,t)$ represents the STFT result of the signal, and the definition of $\widehat{\omega}_f(\eta,t)$ is as

$$\widehat{\omega}_f(\eta,t)=\frac{1}{2\pi}\partial_t\arg V_f(\eta,t)=\mathrm{Re}\left(\frac{1}{2i\pi}\frac{\partial_t V_f(\eta,t)}{V_f(\eta,t)}\right) \quad (4)$$

where $\partial_t\cdot$ represents differentiating $\cdot$ over t.

B. Structure of the Proposed Parameter Extraction Network

Fig. 3 shows the structure of the proposed parameter extraction network using modulation recognition-guided semantic segmentation.

1) Time–frequency diagrams of the overlapping radar signals based on FSST are simultaneously input to the modulation recognition module and the encoder.
2) Encoder consists of channel rearrangement layers, lower basic blocks, a global feature extraction layer, and a interpolate layer. The encoder captures high-dimensional semantic features from the radar signal, and then, feeds them into the decoder.
3) Decoder consists of upper basic blocks, interpolate layers, and a semantic segmentation layer.
4) The time–frequency diagrams are also input to the modulation recognition module. The results of the modulation recognition are used to assist the semantic segmentation to improve segmentation accuracy, and exclude classes that do not exist in recognition results.
5) After the semantic segmentation layer, the parameter extraction layer is followed, which is used to extract the bandwidths and pulsewidths of the segmented signals. And the model parameter table is shown in Table I.

The structure of the whole network is discussed in detail later.

In this article, semantic segmentation is utilized for separating overlapping radar signals. Semantic segmentation based on the CNN is generally structured as an encoder–decoder system [42]. The encoder is typically a network designed for feature extraction, with the primary objective of extracting high-dimensional semantic features from the input image using convolutional layers. The decoder is responsible for reprojecting the high-dimensional features obtained by the encoder onto the pixel space of the original image size, enabling the classification of each pixel. Subsequently, parameters can be extracted from the segmentation masks.

C. Encoder Design

The structure of the encoder is shown in Fig. 3. The input time–frequency diagrams undergo multiscale feature extraction initially, which includes four sets of two lower basic blocks. The primary role of the multiscale feature extraction module is to obtain features from different perceptual fields and integrate these features. Following that, there are four identical feature extraction modules, each containing a channel rearrangement layer and two lower basic blocks. The channel rearrangement

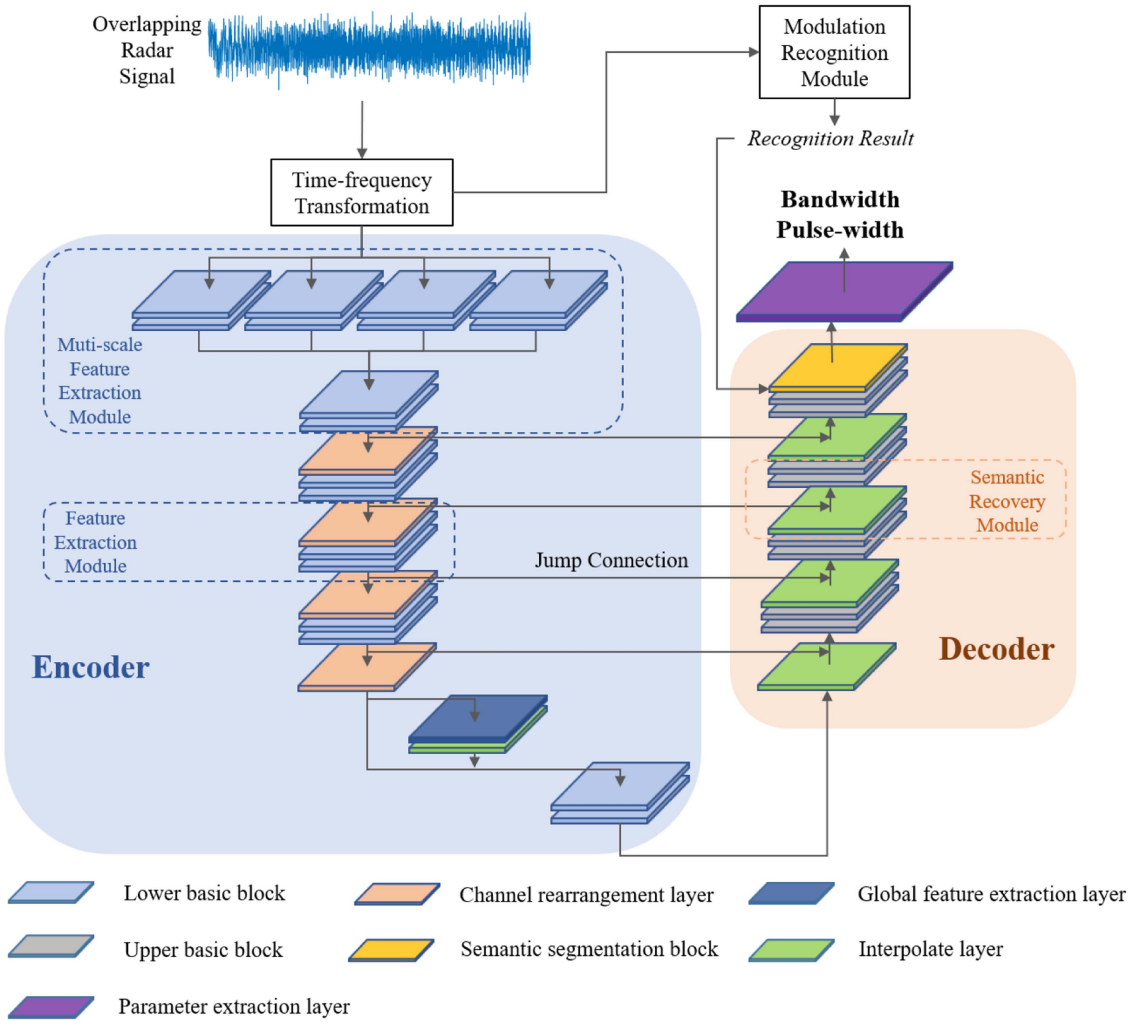

Fig. 3. Structure of the proposed parameter extraction network using modulation recognition-guided semantic segmentation.

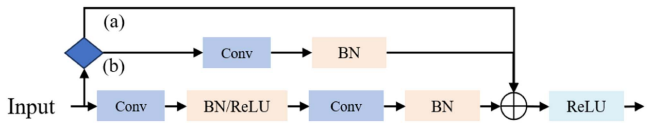

Fig. 4. Structure diagram of the lower basic block. The blue diamond box represents the judgment structure. If the number of input and output channels is consistent, choose (a) to sum; if inconsistent, choose (b).

layer expands the convolutional perceptual field without losing feature local information. The lower basic block uses residual convolution to better propagate gradients. A global feature extraction layer and an interpolation layer are inserted into the last feature extraction module to obtain global features. Each of these structures in the encoder will be further elaborated as follows.

1) Lower Basic Block: The schematic diagram of the lower basic block is shown as Fig. 4. The lower basic block is a fundamental part of the encoder. The complexity of segmentation of overlapping radar signals calls for a deepening in the network layers to enhance its feature extraction capability [43]. In order to better propagate gradients, and alleviate problems such as gradient vanishing and gradient exploding, we used residual convolution in the lower basic block. The lower basic block has two convolution layers, two batch normalization (BN) layers, and two ReLU activation functions.

2) Multiscale Feature Extract Module: Inspired by [44], multiscale feature extraction module is designed as shown in Fig. 3. This module employs parallel residual convolution operations, with each branch utilizing two lower basic blocks to extract features at different scales. By using convolutional kernels at different scales, time–frequency diagram information at different scales can be captured, thereby improving the segmentation ability of overlapping radar signals and enhancing the robustness of the proposed model. When the bandwidth to the pulsewidth ratio (BPR) of the radar signal is small, it appears as a flat curve in the time–frequency plot. Therefore, using a 3×7 convolution kernel is more advantageous for extracting the signal's features. Conversely, when BPR is large, it appears as a slender curve in the time–frequency plot. In this case, using a 7×3 convolution kernel is more effective for extracting the signal's features. The schematic diagram of multiscale feature extraction is shown as Fig. 5.

The parameter settings can be found in Table I, which are named as multiscale 1 to multiscale 5. Channel 1 adopts 3×7

TABLE I
PROPOSED NETWORK PARAMETER TABLE

Name	Type	kernel size/stride/padding	Depth	Input size	Output size	Connected To
Input				(1,512,512)		Multiscale 1,2,3,4
Multiscale 1	Lower Basic Block	3×3/1/1	2	(1,512,512)	(8,512,512)	Multiscale 5
Multiscale 2	Lower Basic Block	5×5/1/1	2	(1,512,512)	(8,512,512)	Multiscale 5
Multiscale 3	Lower Basic Block	3×7/1/(1,3)	2	(1,512,512)	(8,512,512)	Multiscale 5
Multiscale 4	Lower Basic Block	7×3/1/(3,1)	2	(1,512,512)	(8,512,512)	Multiscale 5
Multiscale 5	Lower Basic Block	3×3/1/1	2	(32,512,512)	(32,512,512)	Rearrange 1, De-layer 4
Rearrange 1	Channel Rearrangement Layer		1	(32,512,512)	(128,256,256)	En-layer 1
En-layer 1	Lower Basic Block	3×3/1/1	2	(128,256,256)	(128,256,256)	Rearrange 2, De-layer 3
Rearrange 2	Channel Rearrangement Layer		1	(128,256,256)	(512,128,128)	En-layer 2
En-layer 2	Lower Basic Block	3×3/1/1	2	(512,128,128)	(512,128,128)	Rearrange 3, De-layer 2
Rearrange 3	Channel Rearrangement Layer		1	(512,128,128)	(2048, 64, 64)	En-layer 3
En-layer 3	Lower Basic Block	3×3/1/1	2	(2048,64,64)	(2048,64,64)	Rearrange 4, De-layer 1
Rearrange 4	Channel Rearrangement Layer		1	(2048,64,64)	(8192,32,32)	Global 1, En-layer 4
Global 1	AdaptiveAvgPool2d(1)		1	(8192,32,32)	(8192,1,1)	Interpolate 1
Global 2	Conv2d/ReLU	1×1/1/0	1	(8192,1,1)	(8192,1,1)	Interpolate 1
Interpolate 1	Bilinear Interpolate		1	(8192,1,1)	(8192,32,32)	En-layer 4
En-layer 4	Lower Basic Block	3×3/1/1	2	(16384,32,32)	(16384,32,32)	Interpolate 2
Interpolate 2	Bilinear Interpolate		1	(16384,32,32)	(16384,64,64)	De-layer 1
De-layer 1	Upper Basic Block	3×3/1/1	2	(18432, 64, 64)	(2048, 64, 64)	Interpolate 3
Interpolate 3	Bilinear Interpolate		1	(2048, 64, 64)	(2048, 128, 128)	De-layer 2
De-layer 2	Upper Basic Block	3×3/1/1	2	(2560, 128, 128)	(512, 128, 128)	Interpolate 4
Interpolate 4	Bilinear Interpolate		1	(512, 128, 128)	(512, 256, 256)	De-layer 3
De-layer 3	Upper Basic Block	3×3/1/1	2	(640, 256, 256)	(128, 256, 256)	Interpolate 5
Interpolate 5	Bilinear Interpolate		1	(128, 256, 256)	(128, 512, 512)	De-layer 4
De-layer 4	Upper Basic Block	3×3/1/1	2	(160, 512, 512)	(64, 512, 512)	Segmentation 1
Segmentation 1	Conv2d/SoftMax	1×1/1/0	1	(64, 512, 512)	(7, 512, 512)	Segmentation 2
Segmentation 2	Modulation Type Filtering		1	(7, 512, 512)	(3, 512, 512)	Output

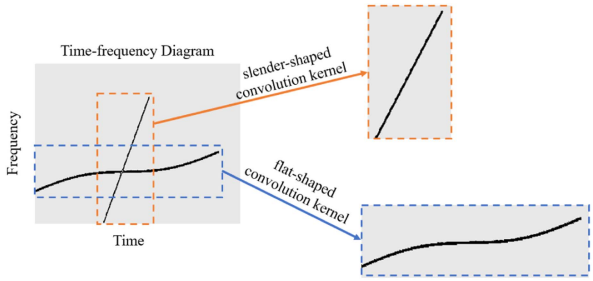

Fig. 5. Schematic diagram of the multiscale feature extraction.

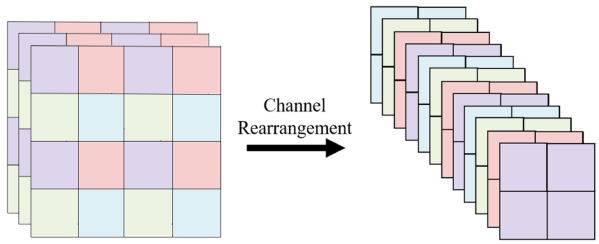

Fig. 6. Schematic diagram of the channel rearrangement layer.

convolution kernels to extract radar signal features with small BPR; and channel 2 adopts 7 × 3 convolution kernels to extract radar signal features with large BPR. Channel 3 adopts a 5 × 5 convolution kernel to extract the features of radar signals with BPR close to 1. Channel 4 uses a 3 × 3 convolution kernel to extract the subtle features. After four channels of feature extraction, the features are concatenated, and then, input to two lower basic block for feature fusion.

3) Channel Rearrangement Layer: To better capture the overall structure of overlapping signals and reduce the occurrence of blurred boundaries and mis-segmentation in the segmentation results, increasing the receptive field is crucial. Traditional semantic segmentation networks often achieve this by employing max-pooling operations. However, for overlapping radar signals, max-pooling operations may loss the critical features such as corners and edges, reducing the accuracy of signal segmentation. To address this issue, we propose the channel rearrangement instead of max-pooling, whose effect is shown in Fig. 6. Channel rearrangement is an optimization of the pooling operation. It preserves the information by separating the feature maps without deleting them. Specifically, for the input feature map, we rearrange it in the height and width dimensions with a stride of 2, resulting in four nonoverlapping subfeature maps. This approach requires that the height and width of the input feature map are both even, which is satisfied in our network. This method achieves the expansion of the perception field while reducing information loss. Assume that the input feature map is

$$X = [x_{p,q}]_{m\times n} \tag{5}$$

where m and n present the height and weight of input feature map. p and q present the height index and weight index. And the formula of channel rearrangement is as follows:

$$\begin{aligned}
\text{Feature1} &= [x_{2i,2j}]_{\frac{m}{2}\times\frac{n}{2}} \\
\text{Feature2} &= [x_{2i,2j+1}]_{\frac{m}{2}\times\frac{n}{2}} \\
\text{Feature3} &= [x_{2i+1,2j}]_{\frac{m}{2}\times\frac{n}{2}} \\
\text{Feature4} &= [x_{2i+1,2j+1}]_{\frac{m}{2}\times\frac{n}{2}}
\end{aligned} \tag{6}$$

where $i = 0, 1, \ldots, \frac{m}{2} - 1$, and $j = 0, 1, \ldots, \frac{n}{2} - 1$. And Feature1, Feature2, Feature3, and Feature4 are four channels of the output.

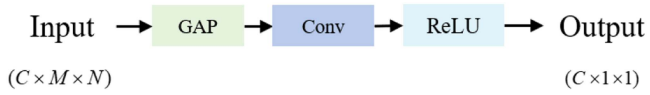

Fig. 7. Structure diagram of the global feature extraction layer. The input size of the global feature extraction layer is $c \times m \times n$, where c denotes the feature channel dimension, and $m \times n$ denotes the size of the feature map. After the GAP processing, the size of the features becomes $c \times 1 \times 1$. At this time, there is only one feature value for each channel, and this feature value is calculated from the global features of each channel.

4) Global Feature Extraction Layer: Inspired by [40], the structure of the global feature extraction layer is shown in Fig. 7. The purpose of the global feature extraction layer is to improve the overall feature connectivity. In the context of semantic segmentation, the features belonging to the same category in long-distance communication pose challenges to the network. This hinders the network's ability to effectively establish connections between distant features during the feature processing stage. Consequently, the network may be limited to performing semantic segmentation based solely on local features. More seriously, in cases where a complete signal exhibits varying frequency characteristics across different time intervals, it may be partitioned into multiple distinct signals. Therefore, the proposed semantic segmentation extracts the connection between distant features through the global feature extraction layer.

Global average pool (GAP) can replace the features of each channel with global information. GAP only focuses on the role of each channel in the entire task while reducing the computational complexity. The calculation formula is

$$\mathrm{GAP} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x(i, j) \tag{7}$$

where H and W represent the height and the width of the input feature map, and $x(i, j)$ represents the value of the input feature at position (i, j). After GAP, the input feature dimension becomes $1 \times 1 \times C$, where *C* represents the number of channels for the input feature.

5) Interpolate Layer: The proposed method uses bilinear interpolation operation to interpolate after global feature extraction, and to do up-sampling in the decoder. The bilinear interpolation is one of the most widely used up-sampling methods in semantic segmentation networks. It calculates the current pixel value based on the distance between the interpolation position and surrounding pixels. Assume that Point *P* is the target point to be interpolated, and its interpolation is calculated from four points: Q_{11}, Q_{12}, Q_{21}, and Q_{22}. The interpolation formula for calculating the target point *P* is as follows:

$$\begin{cases} f(R_1) = \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}) \\ f(R_2) = \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}) \\ f(P) = \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2) \end{cases} \tag{8}$$

where $(x_1, y_1), (x_1, y_2), (x_2, y_1)$, and$(x_2, y_2)$ presents the coordinates of Q_{11}, Q_{12}, Q_{21}, and Q_{22}.

Fig. 8. Structure diagram of the upper basic block.

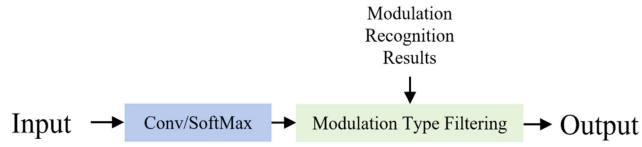

Fig. 9. Structure diagram of the semantic segmentation block.

D. Decoder Design

This section discusses the design of a decoder for the semantic segmentation of overlapping radar signals. The structure diagram of the decoder is presented in Fig. 3. The proposed decoder consists of four semantic recovery modules and a semantic segmentation block. In each semantic recovery module, two upper basic blocks and a interpolate layer are included. In addition to the output of the lower layer, each semantic recovery module takes input from a jump connection at the same level as the encoder. Following processing by three semantic recovery modules, the features are combined with the recognition result generated by the modulation recognition module, and fed into the semantic segmentation block. The ultimate output of the semantic segmentation block is the final segmentation results, which represent the signal masks of various modulation types.

1) Upper Basic Block: The schematic diagram of upper basic block is shown as Fig. 8. Upper basic block is a fundamental part of the decoder. The upper basic block includes a convolution layer, a BN layer, a ReLU activation function, and a dropout layer.

2) Jump Connection: To solve the local segmentation problem mentioned previously, a jump connection is used in this article. The role of the jump connection is to merge the shallow, high-resolution features of the encoder with the deep, high-semantic information features of the decoder. This feature combination is conducive to recovering edge details in the process of overlapping signal segmentation.

3) Semantic Segmentation Block: The structure of the semantic segmentation block is the convolution layer plus the softmax layer, which is shown in Fig. 9. After the softmax layer, modulation type filtering is used, which is discussed later. As one of the most common probability calculation functions in multiclassification tasks, we leverage the softmax layer to compute the category probability for each pixel, and its expression is shown as

$$p_i = \frac{\exp(v_i)}{\sum_{j=1}^{M} \exp(v_j)} \tag{9}$$

where M represents the total number of categories to be classified, which is also equal to the length of the output vector of the target, and $v_i (i = 1, 2, \ldots, M)$ represents the *i*th value in the output vector. After the softmax layer's calculation, the resulting vector indicates the predicted probability of the classification target belonging to each category, each value of the vector is between 0 and 1, and the sum of each value of the vector is 1.

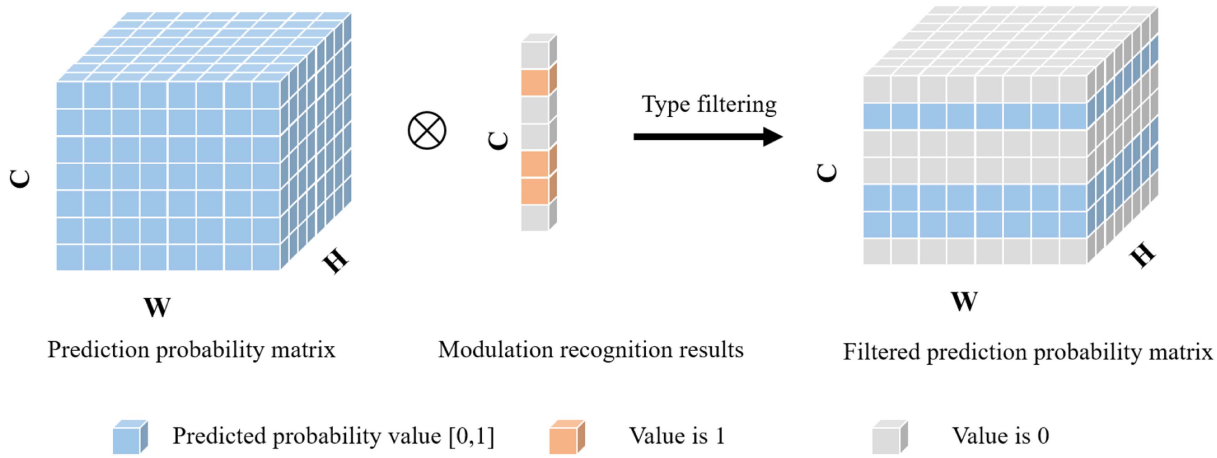

Fig. 10. Schematic diagram of modulation type filtering. The size of the prediction probability matrix after the softmax layer is $W \times H \times C$, where C is the total number of categories for classification, here it is the number of modulation types plus one, the reason for adding one is to include the background class. $W \times H$ is the size of the output features, also equal to the size of the input time–frequency diagram. The modulation recognition result is a vector of length C, where the recognition presence class and background class values are 1, and the rest positions are 0.

4) Modulation Type Filtering: To solve the aforementioned global segmentation problem, modulation type filtering is designed. Modulation type filtering uses the output of modulation recognition module [34], i.e., the recognition results, as prior information to help increase the segmentation accuracy. The modulation recognition module will be discussed later. By filtering the recognition results with the output of the softmax layer, the filtered prediction probability can be obtained.

As shown in Fig. 10, before modulation type filtering, the network produces a tensor of size (7, 512, and 512). This tensor comprises seven channels, with each channel representing a modulation type or background class. Each point in the output tensor represents the probability of belonging to a specific class. The output of the modulation type recognition network is a tensor of length 7, where each value is either 0 or 1, representing the predicted modulation type recognition results by the network. Modulation type filtering means multiplying the two tensors in the channel dimension to obtain the filtered probability matrix. The filtered matrix contains zeros in positions where the modulation type is recognized as nonexistent. Furthermore, classes that are not recognized by the modulation will not be present in the semantic segmentation results.

Following modulation type filtering, semantic segmentation is performed by selecting the class with the highest predicted probability for each pixel as the class to which the pixel belongs (argmax operation).

E. Modulation Recognition Module

The structure of the modulation recognition module is shown in Fig. 11 [34]. The recognition result is output by the modulation recognition module. Within this module, an Inception block is employed to extract features from varying receptive fields, while a squeeze-and-excitation (SE) block is used to reduce noise. An adaptive threshold block concludes two fully connected (FC) layers and the ReLU/Sigmoid layer, which offers an adaptive threshold and eliminates the difficulty of having to choose a threshold in multiclassification tasks.

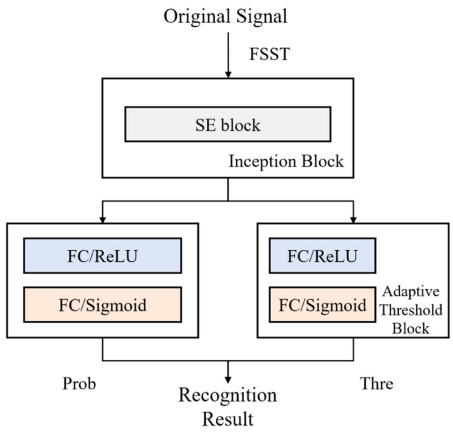

Fig. 11. Structure diagram of the modulation recognition module.

F. Parameter Extraction Layer

After performing modulation recognition guided-semantic segmentation on the overlapping radar signals of two different modulation types, we obtain a probability matrix of size (3, 512, 512), where each pixel represents the probability of belonging to a specific modulation type or the background class. By applying the argmax operation to this probability matrix, we obtain an image of size (512, 512), where each pixel represents the most probable modulation type or the background class. We refer to this image as the mask. In the mask, different modulation types of radar signals can be separated based on the pixel's assigned class. Subsequently, performing temporal compression on the radar signal's mask yields its projection on the frequency axis, while similar frequency domain compression yields their projection on the time axis. Finally, by extracting the span of the compressed mask in the time and frequency domains, we can calculate the pulsewidth and bandwidth of the signal.

The schematic diagram of parameter extraction is shown in Fig. 12. The time–frequency masks of two signals are obtained through semantic segmentation. By separating the masks, the range of the two types of modulation signals on the

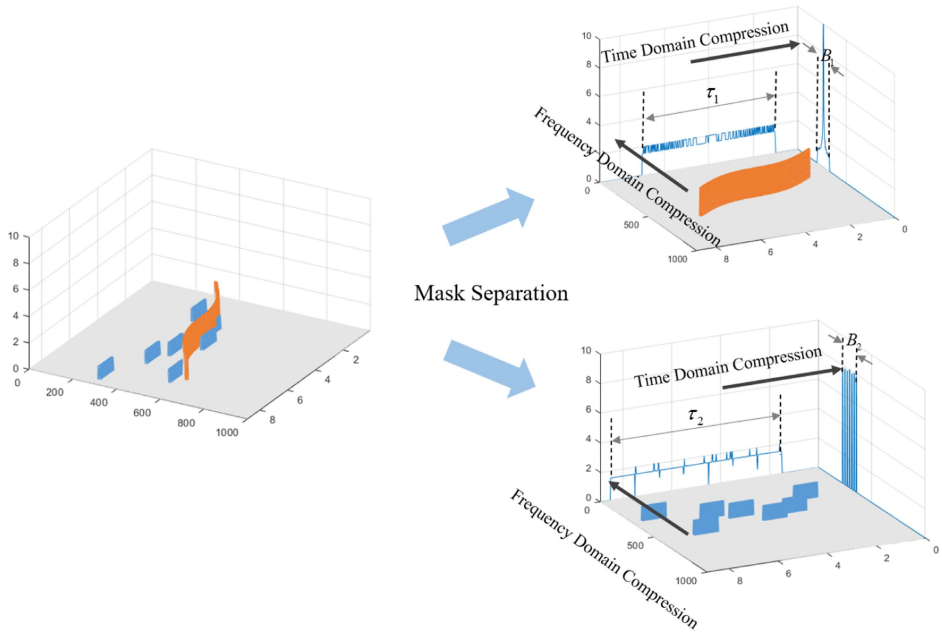

Fig. 12. Schematic diagram of the parameter extraction layer. Typically, the time–frequency representation of the STFT transform is visualized using a heatmap. However, for the purpose of intuitive visualization during the parameter extraction process, we have adopted the use of 3-D images to represent this data. Due to space constraints, we do not list the cases of overlapping signals one by one. The overlapping signals of different modulation types can be seen in heatmap Fig. 2.

time–frequency diagram can be obtained. The bandwidth and pulsewidth can be obtained by compressing the separated signal masks and counting their span range on the time and frequency domain. The calculation of the bandwidth and pulsewidth can be written as

$$B = \frac{n_{Bl} - n_{Br}}{n_B} f_s \tag{10}$$

$$\tau = \frac{n_{tl} - n_{tr}}{n_t} t \tag{11}$$

where B denotes the bandwidth of the signal, n_{Bl} denotes the left edge subscript on the frequency axis after mask compression, n_{Br} denotes the right edge subscript on the frequency axis after mask compression, n_B denotes the total number of samples on the frequency axis, and f_s denotes the sampling rate. Similarly, τ denotes the pulsewidth of the signal, n_{tl} denotes the left edge subscript on the time axis after mask compression, n_{tr} denotes the right edge subscript on the time axis after mask compression, n_t denotes the total number of samples on the time axis, and t denotes the total observation time.

IV. Experiments and Evaluations

In this section, experiments and evaluations of the proposed method will be discussed. First, network training details and the evaluation criterion are introduced. After that, to exhibit the efficacy of the proposed method, four experiments are conducted. The first experiment shows the parameter extraction of common overlapping radar signals. The second experiment displays the capability of the proposed method under varying SNRs, and compares it with conventional semantic segmentation networks and parameter extraction methods. The third is the ablation experiment, which verifies the effectiveness of channel rearrangement and modulation type filtering methods. The fourth experiment demonstrates the differences in the effectiveness of the proposed method on different modulation types.

TABLE II
Experimental Simulation Environment

Platform	Version
RAM	960GB
CPU	Inter(R) Xeon(R) Gold 6128 (3.4GHz)
GPU	NVIDIA Tesla P100-PCIE-16GB
Operating System	Windows Server 2019 Datacenter 64 b
Programming Language	Python 3.7
Deep Learning Framework	Pytorch 1.12.1 and CUDA 11.6
Signal Simulation Platform	MATLAB R2022a

The experimental platform and environment used in this section are shown in Table II.

A. Network Training

1) Dataset: This section generates a dataset of overlapping radar signals. The dataset consists of FSST time–frequency diagrams with six modulated radar signals two-by-two overlapping: SC signal, LFM signal, NLFM signal, FA signal, Costas signal, and P4 signal. The formulaic representation of six signals is presented in Table IV (due to space constraints, the specific meanings represented by each symbol are not elaborated upon). After generating simulated signals, Gaussian white noise is added to meet the desired signal-to-noise ratio (SNR) requirements. The length of the dataset is about 100 000. The pulsewidth, bandwidth, power ratio, overlap degree, and SNR of the signals are used as variable parameters, and their ranges are shown in Table III. After generating the dataset, it is randomly divided into training and testing sets using an 80-20 split.

TABLE III
SIGNAL PARAMETERS

Parameter	Range	step
Pulsewidth	2 μs∼5 μs	1 μs
Bandwidth	10 MHz∼300 MHz	41 MHz
Overlap Degree	30%∼100%	14%
Power Ratio	0 dB∼10 dB	3.3 dB
SNR	-10 dB∼15 dB	5 dB

TABLE IV
FORMULAS OF SIX SIGNALS

Modulation	Formula
SC	$s(t) = \text{rect}\left(\frac{t}{\tau}\right) e^{j(2\pi f_c t+\varphi)}$
LFM	$s(t) = \text{rect}\left(\frac{t}{\tau}\right) e^{\left(j2\pi f_c t+jk\pi t^2\right)}$
NLFM	$s(t) = \text{rect}\left(\frac{t}{\tau}\right) e^{j\left(2\pi f_c t+\pi B\tau\left[\frac{K}{\pi^2}\sin^2\left(\frac{\pi f}{B}\right)-\frac{f^2}{B^2}\right]\right)}$
FA	$s(t) = \text{rect}\left(\frac{t}{\tau}\right) \sum_{i=0}^{N} \text{rect}(t-i\cdot\Delta t) e^{j(2\pi f_i t+k_i\pi t^2)}$
Costas	$s(t) = \text{rect}\left(\frac{t}{\tau}\right) \sum_{i=0}^{N} \text{rect}(t-i\cdot\Delta t) e^{j(2\pi f_i t+\varphi)}$
P4	$u(t) = \frac{1}{T}\sum_{n=1}^{N} u_n \,\text{rect}\left(\frac{t-(m-1)t_b}{t_b}\right)$

2) Loss Function: On the time–frequency diagram of the radar signal, since the pixel number occupied by the background is far greater than those occupied by the signal, the same weight of pixels occupied by the background and those occupied by the signal would cause the loss function to fall into a local optimum. Consequently, the focal loss function is applied to tackle the challenge of imbalanced sample distribution, and its expression is shown as follows:

$$\text{loss} = -\alpha_t \left(1-p_t\right)^{\gamma} \log\left(p_t\right) \tag{12}$$

where α_t represents the weight of the category, p_t represents the prediction probability, and γ is a hyperparameter. In this article, α_t of the background is taken as 0.1; α_t of the signal is taken as 0.9; and γ is taken as 2.

3) Network Configuration: The network parameters are set according to the Table I. The input images are normalized and transformed to the size of (512, 512). The network was trained using the Adam optimization algorithm with an initial learning rate of 0.001. We employed a batch size of 32 and trained the network using the aforementioned focal loss function. In addition, we employed a learning rate scheduler to adjust the learning rate based on the training progress, ensuring stable and efficient convergence. Specifically, we decayed the learning rate by a factor of 0.5 if the validation loss did not improve for five consecutive epochs. If the validation loss does not improve after continuously changing the learning rate five times, stop training.

B. Evaluation Criterion

The effectiveness of the experiments is assessed through two metrics. In this article, mean intersection over union (MIoU) is employed to gauge the efficacy of semantic segmentation for overlapping radar signals; root mean squared error (RMSE) is adopted to assess the efficacy of parameter extraction.

1) MIoU: MIoU is defined as the average of the intersection ratio for each category [45]. Fig. 13 displays the schematic

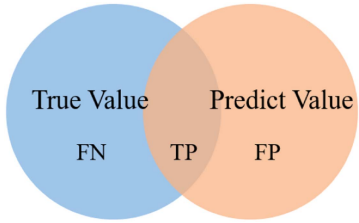

Fig. 13. Schematic diagram of the segmentation results.

TABLE V
OVERLAPPING SIGNAL PARAMETERS

Signal type	Pulsewidth	Bandwidth
Costas	5.5 μs	200 MHz
NLFM	4.8 μs	190 MHz

diagram of the segmentation outcomes. True positive (TP) denotes positive samples that have been accurately predicted by the model; false negative (FN) represents positive samples that have been inaccurately predicted by the model; and false positive (FP) signifies negative samples that have been inaccurately predicted by the model. The expression of MIoU is shown as

$$\begin{aligned}\text{MIoU} &= \frac{1}{k+1}\sum_{i=0}^{k}\frac{\text{TP}}{\text{FN}+\text{FP}+\text{TP}} \\ &= \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{i=0}^{k}p_{ij}+\sum_{i=0}^{k}p_{ji}-p_{ii}}\end{aligned} \tag{13}$$

where i denotes true class, j denotes predict class, p_{ij} denotes predicting i as j, i.e., FN, p_{ji} denotes predicting j as i, i.e., FP, p_{ii} denotes predicting i as i, i.e., TP.

2) RMSE: The definition of RMSE is the distinction between the predicted parameters and the actual parameters [46]. Its expression is shown as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2} \tag{14}$$

where Y_i denotes the true parameter, and $\hat{Y}_i$ denotes the predicted value.

C. Parameter Extraction of Typical Overlapping Radar Signals

This experiment demonstrates the parameter extraction process of the proposed method using overlapping Costas signal and NLFM signals as an example, and the results are shown in Fig. 14. The parameters of the two signals are shown in Table V. Set SNR as 0 dB, the sampling rate as 600 MHz, and the total signal duration as 7 μs.

After the semantic segmentation of the overlapping radar signals, the time–frequency masks for both NLFM and Costas signals are acquired. The masks for these two signals are then separated and compressed in the time and frequency dimensions, respectively. In this experiment, the real pulsewidth of the Costas signal is 5.5 μs and the bandwidth is 200 MHz, and the estimated values are 5.53 μs and 193 MHz, with the relative errors of 0.5% and 3.5%. The real pulsewidth and bandwidth of the NLFM

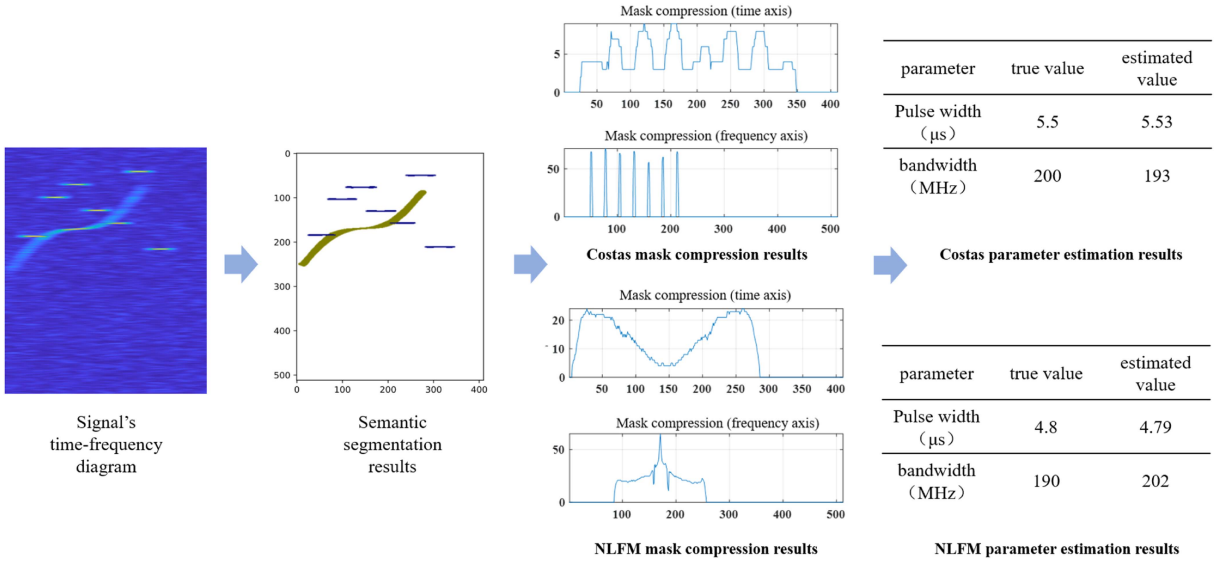

Fig. 14. Parameter extraction results of typical overlapping radar signals using the proposed method.

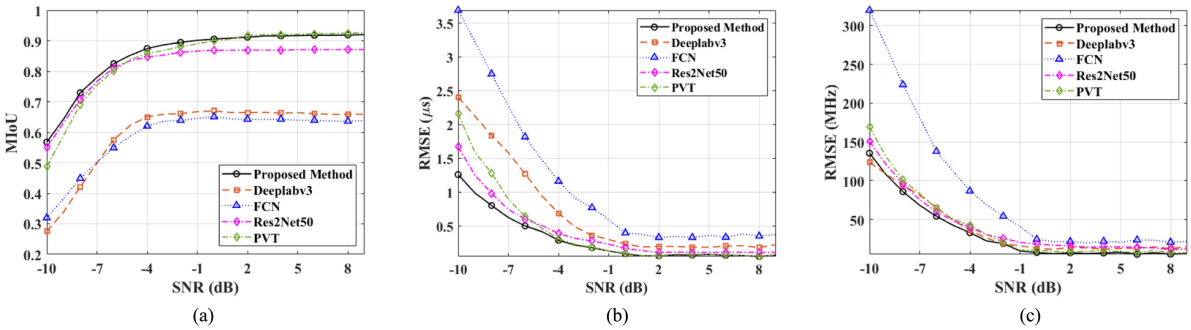

Fig. 15. Comparison experiment. (a) Segmentation MIoU comparison. (b) Pulsewidth RMSE comparison. (c) Bandwidth RMSE comparison.

signal are 4.8 μs and 190 MHz, and the estimated values are 4.79 μs and 202 MHz, with relative errors of 0.2% and 6.3%.

In addition, the signal time–frequency map size is 512×512, the sampling rate is 600 MHz, the total signal duration is 7 μs, the frequency accuracy of each pixel point in the time–frequency map is 1.17 MHz, and the time accuracy is 0.016 μs. In this overlapping case, the Costas signal has an error of 2 and 6 pixels on the time and frequency axes; the NLFM signal has an error of 1 and 10 pixels on the time and frequency axes, respectively. The outcomes of this experiment illustrate that the proposed method can proficiently extract the pulsewidth and bandwidth of overlapping radar signals.

D. Parameter Extraction Performance Comparison at Different SNRs

This experiment compares the signal segmentation and parameter extraction performance of the proposed method with four commonly used segmentation networks: fully connected network (FCN) [47], a milestone network in image segmentation; Deeplab v3 [48], a widely used CNN based network; Res2Net [49], an improved network for ResNet; Pyramid vision transformer (PVT) [50], a transformer-based network. After employing these networks for semantic segmentation, the same operation is carried out to extract parameters as the proposed method. The results of the comparison experiments are shown in Fig. 15.

1) Semantic Segmentation Comparison: The MIoU of the proposed method is higher than 0.9 at 0 dB. Across all SNRs, the proposed method's MIoU is significantly higher than that of FCN and Deeplabv3. At SNRs greater than 0 dB, the MIoU of the proposed method is approximately 3% higher than that of Res2Net and roughly equivalent to that of PVT. At SNRs below 0 dB, the proposed method exhibits the most gradual decrease in MIoU compared to the other four methods. Specifically, at -10 dB, the proposed method achieves an MIoU 1.7% higher than that of Res2Net and approximately 8% higher than that of PVT. The MIoU results illustrate that the segmentation results obtained by the proposed method exhibit superior efficacy compared to those generated by Res2Net, PVT, Deeplabv3, and FCN.

2) Parameter Extraction Comparison: The pulsewidth RMSE of the proposed method is 0.06 μs at 0 dB, and the bandwidth RMSE is about 8 MHz. Across all SNRs, the

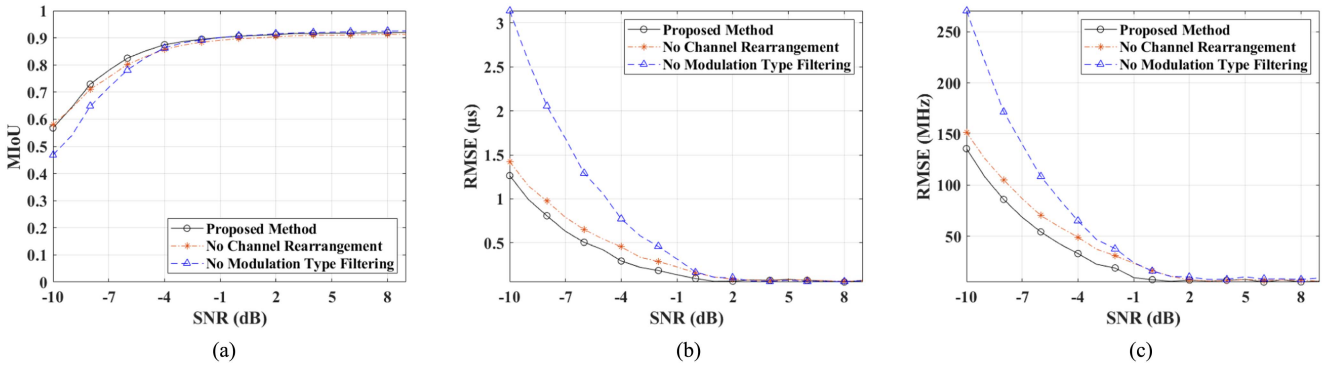

Fig. 16. Ablation experiment. (a) Segmentation MIoU comparison. (b) Pulsewidth RMSE comparison. (c) Bandwidth RMSE comparison.

proposed method exhibits significantly higher pulsewidth RMSE compared to that of FCN and Deeplabv3, as well as significantly higher bandwidth RMSE compared to that of FCN. Above 0 dB, the proposed method's pulsewidth RMSE is similar to that of PVT, and approximately 0.07 μs higher than that of Res2Net; and the proposed method's bandwidth RMSE is similar to PVT, Deeplabv3, and Res2Net. Below 0 dB, the proposed method's pulsewidth RMSE and bandwidth RMSE are significantly higher than those of Res2Net and PVT, except for at -10 dB, where it is slightly lower than that of Deeplabv3. The RMSE outcomes indicate that the proposed network exhibits improved global performance in parameter extraction.

3) Discussion: Compared to the proposed method, the receptive field of FCN and Deeplabv3 is smaller, so the ability to obtain global information is poorer, and the modulation type recognition result is not used as a prior to modify the segmentation result. The rationale behind the effectiveness of the proposed method in contrast to FCN and Deeplabv3 is that channel rearrangement and modulation type filtering improve the accuracy of segmentation.

We observe that at SNR above 0, the performance difference between PVT and the proposed method is relatively small. However, when the SNR falls below 0, the performance of PVT significantly deteriorates, leading to an increasing disparity compared to the proposed method. This is due to the PVT model's ability to better utilize clean input data, resulting in a relatively minor impact from noise, thus minimizing the performance gap with the proposed method. Conversely, as the SNR decreases, the influence of noise becomes more pronounced, causing a notable decline in the performance of PVT. In contrast, the proposed method employs channel rearrangement and modulation type filtering, which are specific operations designed to address the overlapping radar signals under low SNRs, enhancing robustness against noise.

In addition, we observe that when the SNR is less than 0, the performance difference between res2net and the proposed method is relatively small; however, when the SNR is greater than 0, the performance gap of res2net is significant. This is because when the SNR is less than 0, the impact of noise is predominant, and both res2net and the proposed method are similarly affected. However, when the SNR is greater than 0, the significance of the signal becomes apparent, and the proposed method, which employs channel rearrangement and modulation type filtering, is able to better utilize the signal information, leading to the observed performance disparity with res2net.

E. Ablation Experiments

To verify the effects of channel rearrangement and modulation type filtering, ablation experiments are designed in this section.

1) Ablation of Channel Rearrangement: In this ablation experiment, the channel rearrangement method is replaced by the maximum pooling method, and the results are shown in Fig. 16. When ablating the channel rearrangement, the maximum reduction in pulsewidth RMSE is 0.7 μs, and the maximum reduction in bandwidth RMSE is 83 MHz. After adding the channel rearrangement, there is an improvement in both signal segmentation and parameter estimation. This is because channel rearrangement preserves more features, which is conducive to the segmentation of details such as the edges of the signal time–frequency diagram.

2) Ablation of Modulation Type Filtering: In this ablation experiment, the modulation type filtering is removed, and the results are shown in Fig. 16. When ablating the modulation type filtering, the maximum reduction in MIoU was 0.22, and the maximum reduction in RMSE of pulsewidth and bandwidth was 2.96 μs and 248 MHz, respectively. The modulation type filtering method has a relatively small impact on the segmentation but has a significant impact on the parameter estimation. This is because modulation type filtering can pull the wrong part of the signal segmentation back to the correct class. If the filtering operation is ablated, a small part of the signal will be segmented incorrectly. The reduced value of MIoU is related to the area of segmentation error. When extracting parameters, only the distance between the compressed mask edge affects the extraction of bandwidth and pulsewidth. Even if only a small portion is segmented incorrectly, it could lead to significant parameter extraction errors.

F. Parameter Extraction for Different Modulation Types

This experiment demonstrates the parameter extraction performance of the proposed method from different overlapping radar signals. This experiment involves the preparation of six

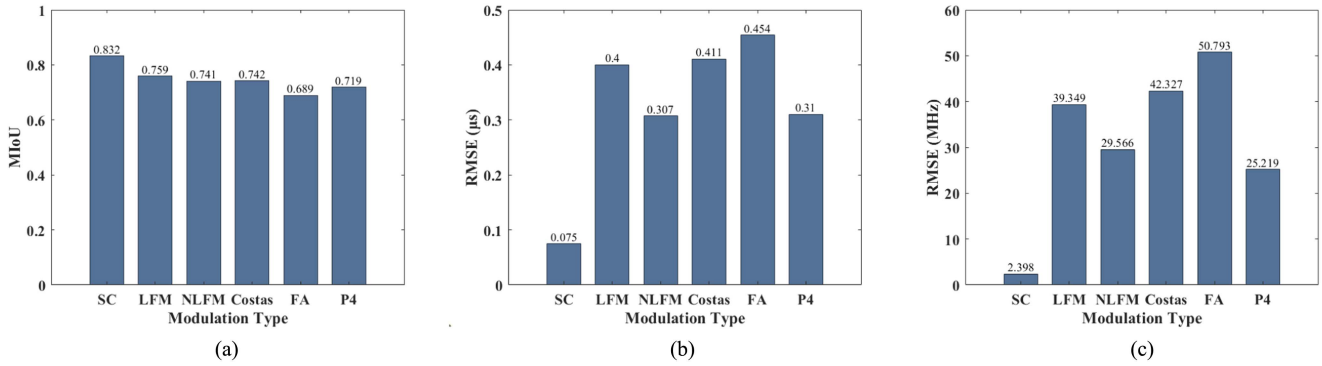

Fig. 17. Effect of the proposed method for different modulation types. (a) Segmentation MIoU. (b) Pulsewidth RMSE. (c) Bandwidth RMSE.

modulation radar signal types. The outcomes of this experiment are illustrated in Fig. 17, where the signal's modulation type is exhibited on the horizontal axis, while MIoU and RMSE are showcased on the vertical axis. Experiments show that signal segmentation and parameter extraction of each type of radar signal are effective. The experimental result shows that the signal segmentation and parameter extraction effect of the SC signal is the best, because the SC signal is reflected in the time–frequency diagram as a straight line, which is the simplest modulation type. The signal segmentation and parameter extraction effect of the FA signal is the worst, which is because the frequency variation of the FA signal with time is discontinuous and the modulation type is more complex.

In the aforementioned four experiments, the performance of the proposed method in signal segmentation and parameter extraction is evaluated under different modulation types and SNR conditions. And the effectiveness of channel rearrangement and modulation type filtering is validated through ablation experiments. The results indicate that the proposed method outperforms traditional four methods in terms of separation and extraction accuracy.

V. Discussion

This section begins by analyzing the performance of the proposed channel rearrangement and modulation type filtering. Subsequently, it delves into the examination of misidentification cases and their respective causes. Finally, the section outlines the limitations of the proposed algorithm and provides an overview of future work.

Channel rearrangement play a crucial role in the proposed method. It is an optimization of the pooling operation in the context of overlapping radar signals. The comparative effects of this operation are illustrated in the Fig. 18. The raw feature map, with a size of (32, 32), can be transformed into a size of (16, 16) through max pooling or channel rearrangement. In the result of max-pooling, the edge, protrusion, and corner information of the raw feature map (as indicated by the colored box) are disrupted, which would result in incomplete semantic segmentation masks, reducing the accuracy of parameter extraction. In comparison, the proposed channel rearrangement offers the advantage of working with four channels, preserving crucial information for the parameter extraction of overlapping radar signal.

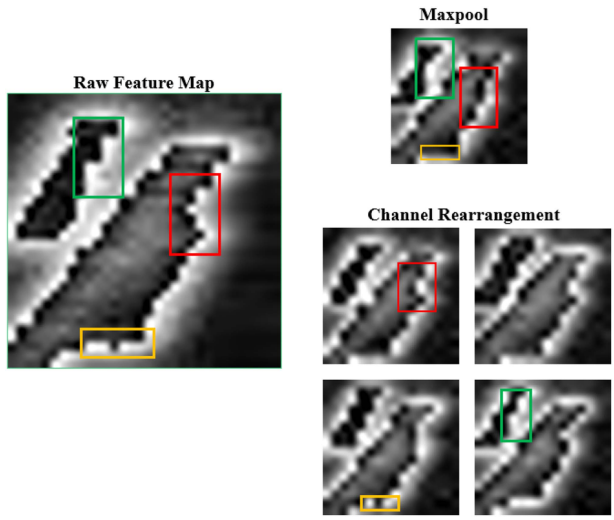

Fig. 18. Comparison effect of channel rearrangement and max-pooling operation.

The effectiveness of modulation type filtering is shown in Fig. 19, where an LFM signal overlaps with a Costas signal. In segmentation without using modulation type filtering, the overlapping signal is incorrectly divided into two LFM signals and one Costa signal. However, when implementing of modulation type filtering, the overlap part of signals is effectively segmented, demonstrating the effectiveness of prior knowledge about modulation types in improving the accuracy of signal segmentation.

It should be noted that in certain specific scenarios, some modulation types may be prone to confusion, leading to a decrease in parameter extraction accuracy. On one hand, when the subbands of a Costas signal are relatively long, they may be identified as several SC signals. This is due to the fact that the time–frequency representation of an SC signal appears as a horizontal line, while that of a Costas signal appears as multiple horizontal lines. On the other hand, in scenarios with relatively low SNRs, there is often a tendency to confuse SC signals with P4 signals. This is because, the time–frequency representations of these two types of signals under lower SNRs appear more similar. Therefore, when dealing with overlapping Costas, SC, and P4 radar signals, it is highly likely that they may be misidentified as the same type of signal, leading to errors.

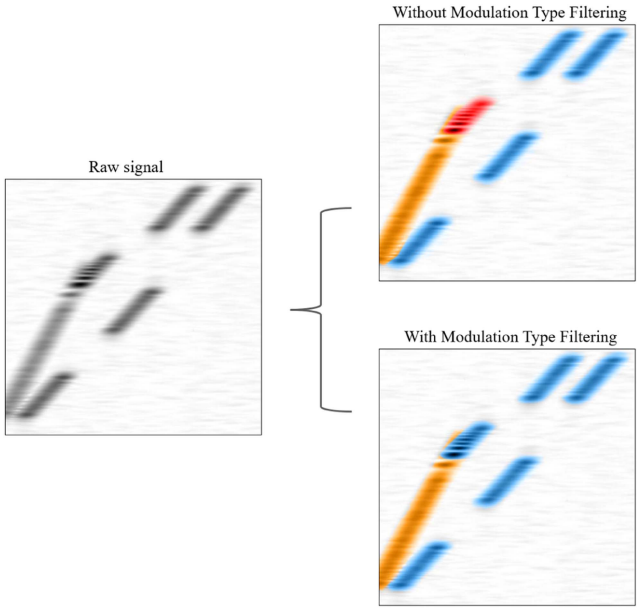

Fig. 19. Ablation effect of modulation type filtering. The yellow, red, and blue parts in the figure represent the segmented LFM signal 1, LFM signal 2, and Costas signal, respectively.

The proposed method still has certain limitations. First, this method focuses primarily on the extraction of the most significant parameters for electronic reconnaissance: bandwidth and pulsewidth, while disregarding other parameters such as pulse amplitude, the LFM slope, and the Costas encoding, etc. Second, the proposed network structure can be further streamlined, which can accelerate the convergence of the network and reduce computational resources. Finally, the method utilizes the encoder–decoder network architecture, which is well-established in the field of radar signal segmentation, instead of employing more advanced backbones developed for semantic segmentation tasks. Therefore, in future work, we intend to address parameter extraction for radar signals with unknown modulation types, explore more advanced networks for complex situations, and explore more effective methods to extraction more valuable parameters.

VI. Conclusion

In this article, we propose a parameter extraction method for overlapping radar signals, which primarily comprises two steps: the first step involves separating the overlapping radar signals by semantic segmentation, and the second step is to compress the separated signal masks to extract signal parameters. As the effect of signal segmentation determines the accuracy of parameter extraction, the channel rearrangement is added to the semantic segmentation network to minimize the feature information loss. In addition, the output of semantic segmentation is multiplied by the modulation recognition result to increase the accuracy of segmentation, which is named modulation type filtering. Comparison experiments, ablation experiments, and experiments on different modulation types verify the effectiveness of the proposed method in extracting parameters from overlapping radar signals. At 0-dB SNR, the segmentation MIoU of the proposed method can reach over 0.9; the pulsewidth extraction RMSE is 0.06 μs, and the bandwidth extraction RMSE is 8 MHz.

References

[1] D. C. Schleher, *Electronic Warfare in the Information Age*. Norwood, MA, USA: Artech House, 1999.
[2] S. Vakin, L. Shustov, and R. Dunwell, *Fundamentals of Electronic Warfare*. Norwood, MA, USA: Artech House, 2001.
[3] S. Q. Wang, G. P. Hu, Q. L. Zhang, C. Y. Gao, and T. Cai, "The background and significance of radar signal sorting research in modern warfare," *Procedia Comput. Sci.*, vol. 154, pp. 519–523, 2019.
[4] C. Zhang, L. Wang, R. Jiang, J. Hu, and S. Xu, "Radar jamming decision-making in cognitive electronic warfare: A review," *IEEE Sensors J.*, vol. 23, no. 11, pp. 11383–11403, Jun. 2023.
[5] K. Kulpa, K. Lukin, W. Miceli, and T. Thayaparan, "Signal processing in noise radar technology," *IET Radar, Sonar Navigation*, vol. 4, no. 2, pp. 229–232, 2008.
[6] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proc. IEEE*, vol. 86, no. 2, pp. 421–441, Feb. 1998.
[7] D. Brandwood, *Fourier Transforms in Radar and Signal Processing*. Norwood, MA, USA: Artech House, 2012.
[8] V. C. Chen and H. Ling, "Joint time-frequency analysis for radar signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 2, pp. 81–93, Mar. 1999.
[9] B. Cantrell, W. Gordon, and G. Trunk, "Maximum likelihood elevation angle estimates of radar targets using subapertures," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES- 17, no. 2, pp. 213–221, Mar. 1981.
[10] C. Magnard, M. Frioud, D. Small, T. Brehm, and E. Meier, "Analysis of a maximum likelihood phase estimation method for airborne multibaseline SAR interferometry," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 3, pp. 1072–1085, Mar. 2016.
[11] B. Tang, J. Tang, Y. Zhang, and Z. Zheng, "Maximum likelihood estimation of DOD and DOA for bistatic MIMO radar," *Signal Process.*, vol. 93, no. 5, pp. 1349–1357, 2013.
[12] J. Zheng, T. Su, G. Liao, H. Liu, Z. Liu, and Q. H. Liu, "ISAR imaging for fluctuating ships based on a fast bilinear parameter estimation algorithm," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 3954–3966, Aug. 2015.
[13] S. Jardak, S. Ahmed, and M.-S. Alouini, "Low complexity moving target parameter estimation for MIMO radar using 2D-FFT," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4745–4755, Sep. 2017.
[14] E. Brookner, *Radar Technology*. Dedham, MA, USA: Artech House, 1977.
[15] J. Yang and Y. Zhang, "An airborne SAR moving target imaging and motion parameters estimation algorithm with azimuth-dechirping and the second-order keystone transform applied," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 3967–3976, Aug. 2015.
[16] Z. Liu, Z. Li, C. Huang, J. Wu, and J. Yang, "Bistatic forward-looking SAR KDCT-FSFT-based refocusing method for ground moving target with unknown curve motion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 8, pp. 4848–4858, Aug. 2020.
[17] J. Zhang, T. Su, J. Zheng, and X. He, "Novel fast coherent detection algorithm for radar maneuvering target with jerk motion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 1792–1803, May 2017.
[18] D. Griffin and J. Lim, "Signal estimation from modified short-time Fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-32, no. 2, pp. 236–243, Apr. 1984.
[19] B. Bouachache and P. Flandrin, "Wigner-Ville analysis of time-varying signals," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, 1982, vol. 7, pp. 1329–1332.
[20] L. Zhang, X.-H. He, Y.-C. Li, M.-D. Xing, and Z. Bao, "Parameters estimation of LFM signals based on STTFD," in *Proc. 9th Int. Conf. Signal Process.*, 2008, pp. 2351–2355.
[21] C. Zeng, D. Li, X. Luo, D. Song, H. Liu, and J. Su, "Ground maneuvering targets imaging for synthetic aperture radar based on second-order keystone transform and high-order motion parameter estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 11, pp. 4486–4501, Nov. 2019.
[22] Y. Huang, Y. Song, Y. Gu, and W. Gao, "A fast parameter estimation method for multicomponent LFM signal based on FRFT," in *Proc. Int. Conf. Virtual Reality Intell. Syst.*, 2020, pp. 393–396.

[23] G. Qiao, D. Dai, and C. Zhang, "Parameter estimation of LFM signal based on improved fractional Fourier transform," in *Proc. 6th Int. Conf. Digit. Signal Process.*, 2022, pp. 223–229.
[24] F. A. Butt and M. Jalil, "An overview of electronic warfare in radar systems," in *Proc. Int. Conf. Technological Adv. Elect. Electron. Comput. Eng.*, 2013, pp. 213–217.
[25] A. Spezio, "Electronic warfare systems," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 3, pp. 633–644, Mar. 2002.
[26] K. N. Parashar, M. C. Oveneke, M. Rykunov, H. Sahli, and A. Bourdoux, "Micro-doppler feature extraction using convolutional auto-encoders for low latency target classification," in *Proc. IEEE Radar Conf.*, 2017, pp. 1739–1744.
[27] Q. Lv, Y. Quan, M. Sha, W. Feng, and M. Xing, "Deep neural network-based interrupted sampling deceptive jamming countermeasure method," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, no. 10, pp. 9073–9085, Oct. 2022.
[28] H. Su, J. Pan, Q. Bao, and Z. Chen, "Parameter estimation of linear frequency modulated signals based on a Wigner–Ville distribution complex-valued convolutional neural network," *J. Appl. Remote Sens.*, vol. 14, no. 3, 2020, Art. no. 0 36512.
[29] S. Wagner, "Combination of convolutional feature extraction and support vector machines for radar ATR," in *Proc. 17th Int. Conf. Inf. Fusion*, 2014, pp. 1–6.
[30] C. Wang, J. Wang, and X. Zhang, "Automatic radar waveform recognition based on time-frequency analysis and convolutional neural network," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2017, pp. 2437–2441.
[31] L. Peng, X. Liao, and M. Chen, "Resampling parameter estimation via dual-filtering based convolutional neural network," *Multimedia Syst.*, vol. 27, pp. 363–370, 2021.
[32] Y. Shi, L. Du, and Y. Guo, "Unsupervised domain adaptation for SAR target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, no. 6, pp. 6372–6385, Jun. 2021.
[33] Z. Zhang, M. Zhu, Y. Li, Y. Li, and S. Wang, "Joint recognition and parameter estimation of cognitive radar work modes with LSTM-transformer," *Digit. Signal Process.*, 2023, Art. no. 104081.
[34] H. Wang, W. Huo, Y. Lu, J. Pei, and Y. Huang, "Modulation recognition of overlapping radar signals under low SNR based on SE-incepatnet," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 2734–2737.
[35] K. Chen, L. Wang, J. Zhang, S. Chen, and S. Zhang, "Semantic learning for analysis of overlapping LPI radar signals," *IEEE Trans. Instrum. Meas.*, vol. 72, no. 2, Feb. 2023, Art. no. 8501615.
[36] Z. Yue, F. Gao, Q. Xiong, J. Wang, A. Hussain, and H. Zhou, "A novel attention fully convolutional network method for synthetic aperture radar image segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 8, pp. 4585–4598, Aug. 2020.
[37] Z. Qu, C. Hou, C. Hou, and W. Wang, "Radar signal intra-pulse modulation recognition based on convolutional neural network and deep q-learning network," *IEEE Access*, vol. 8, pp. 49125–49136, 2020.
[38] H. Yuet al., "Methods and datasets on semantic segmentation: A review," *Neurocomputing*, vol. 304, pp. 82–103, 2018.
[39] A. Daw, A. Karpatne, W. D. Watkins, J. S. Read, and V. Kumar, "Physics-guided neural networks (PGNN): An application in lake temperature modeling," in *Knowledge Guided Machine Learning*. Boca Raton, FL, USA: CRC Press, 2022, pp. 353–372.
[40] W. Liu, A. Rabinovich, and A. C. Berg, "Parsenet: Looking wider to see better," in *Proc. Int. Conf. Learn. Represent.*, 2016, pp. 1–11.
[41] L. Li, H. Cai, H. Han, Q. Jiang, and H. Ji, "Adaptive short-time Fourier transform and synchrosqueezing transform for non-stationary signal separation," *Signal Process.*, vol. 166, 2020, Art. no. 107231.
[42] S. Hao, Y. Zhou, and Y. Guo, "A brief survey on semantic segmentation with deep learning," *Neurocomputing*, vol. 406, pp. 302–321, 2020.
[43] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
[44] C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.
[45] M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *Advances in Visual Computing*. Berlin, Germany: Springer, 2016, pp. 234–244.
[46] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature,"*Geoscientific Model Develop.*, vol. 7, no. 3, pp. 1247–1250, 2014.
[47] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.
[48] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 801–818.
[49] S. Gao, M. Cheng, K. Zhao, X. Zhang, M. Yang, and P. Torr, "Res2Net: A new multi-scale backbone architecture," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 2, pp. 652–662, Feb. 2021.
[50] W. Wang et al., "Pyramid vision transformer: A versatile backbone for dense prediction without convolutions," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 568–578.

Weibo Huo (Member, IEEE) received the B.S. degree in electronic engineering from the School of Science, Shandong Jianzhu University, Shandong, China, in 2010, and the Ph.D. degree in information and communication engineering from the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2019.

He is currently an Associate Research Fellow with the School of Information and Communication Engineering, UESTC. His research interests include statistical signal processing, machine learning, synthetic aperture radar image processing, radar target detection, and sea clutter modeling and simulation.

Yang Luo (Graduate Student Member, IEEE) received the B.S. degree in communication engineering from the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China, in 2022. He is currently working toward the master degree in information and communication engineering with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China.

His research interests include radar countermeasure, cognitive radar jamming, pattern recognition, radar signal processing, reinforcement learning, and machine learning.

Hao Wang received the master's degree in electronic and information engineering from the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China, in 2023.

His research interests include radar operation mode recognition, radar signal processing, and machine learning.

Jifang Pei (Member, IEEE) received the B.S. degree in information engineering from the College of Information Engineering, Xiangtan University, Hunan, China, in 2010, the M.S. degree in electronic engineering from the School of Electronic Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2013, and the Ph.D. degree in electronic engineering from the School of Information and Communication Engineering, UESTC, in 2018.

From 2016 to 2017, he was a joint Ph.D. Student with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. He is currently an Associate Research Fellow with the School of Information and Communication Engineering, UESTC. His research interests include radar signal processing, machine learning, and automatic target recognition.

Yin Zhang (Member, IEEE) received the B.S. and Ph.D. degrees in electronic engineering from the School of Electronic Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2008 and 2016, respectively.

From 2014 to 2015, he was a Visiting Student with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, USA. He is currently a Research Fellow with the UESTC. His research interests include signal processing and radar imaging.

Yulin Huang (Senior Member, IEEE) received the B.S. degree in electronic engineering and Ph.D. degree in information and communication engineering from the School of Electronic Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2002 and 2008, respectively.

He is currently a Full Professor with the UESTC. His research interests mainly include radar signal processing and synthetic aperture radar imaging.

Jianyu Yang (Member, IEEE) received the B.S. degree in electronic engineering from the National University of Defense Technology, Changsha, China, in 1984, and the M.S. and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1987 and 1991, respectively.

From 2001 to 2005, he served as the Dean with the School of Electronic Engineering, UESTC. In 2005, he was a Senior Visiting Scholar with the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. He is currently a Professor with the UESTC. His research interests include radar signal processing and synthetic aperture radar imaging.

Dr. Yang serves as a Senior Editor for the *Chinese Journal of Radio Science* and the *Journal of Systems Engineering and Electronics*. He was selected as the Vice-Chairman of the Radar Society of the Chinese Institute of Electronics (CIE) in 2016 and a Fellow of CIE in 2018.