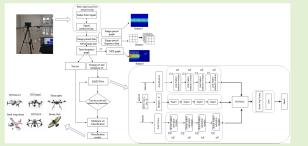
Sensors Council

A Low Slow Small Target Classification Network Model Based on K-Band Radar Dynamic Multifeature Data Fusion

Wang Yuan, Xiaolong Chen[®], *Senior Member, IEEE*, Xiaolin Du[®], *Member, IEEE*, Jian Guan, Jinhao Wang, and Tiange Lan

Abstract—Micro-Doppler (m-D) signals are susceptible to interference from a large number of Doppler signals and ambient noise, and the single use of m-D signatures (MDSs) for the classification of small, slow, and low-speed targets poses certain limitations. In this article, a dynamic multifeature data fusion neural network (DMFFNNet) classification method is proposed. First, K-band frequency-modulated continuous-wave (FMCW) radar is used to collect echo data from five types of rotor drones and bionic bird. After preprocessing the data, 2-D range–period graphic and 2-D time–frequency (TF) spectrograms are obtained. We investi-



gate the construction of new data representations in the range-periodic domain, designing networks to extract dynamic time-varying features of the data. To be able to obtain accurate localized features, a local feature extraction module is proposed to extract local features from the range-period graph, while a global feature extraction module is used to extract global features from the TF spectrograms. To be able to extract dynamic information about the data, a 3-D network is used to capture dynamic change feature in the 3-D range-period data. Finally, a feature fusion module is designed to integrate the extracted features, and to be able to better extract the features of the target, an attention mechanism is added to the fusion network to extract the temporal and spatial features in the spectrogram and fuse them to further improve the overall performance of the model. Experimental results show that compared with single-channel CNN classification methods, incorporating dynamic feature data enables the network to achieve better classification accuracy.

Index Terms—Classification, frequency-modulated continuous-wave (FMCW) radar, low and slow target, micro-Doppler (m-D) features, neural networks.

I. INTRODUCTION

S MALL unmanned aerial vehicles (UAVs) have become increasingly popular due to their compact size, low cost, and ease of control. They have found widespread applications in fields such as aerial photography, environmental monitoring, and mobile communications [1]. While UAVs' technology offers many potential benefits, it also poses significant chal-

Received 6 September 2024; revised 19 October 2024; accepted 4 November 2024. Date of publication 18 November 2024; date of current version 2 January 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62222120 and in part by the Natural Science Foundation of Shandong under Grant ZR2024JQ003. The associate editor coordinating the review of this article and approving it for publication was Prof. Fangqing Wen. (*Corresponding authors: Xiaolong Chen; Xiaolin Du.*)

Wang Yuan and Xiaolin Du are with the School of Computer and Control Engineering, Yantai University, Yantai 264005, China (e-mail: 2896512634@qq.com; duxiaolin168@vip.163.com).

Xiaolong Chen, Jian Guan, and Jinhao Wang are with the Naval Aviation University, Yantai 264001, China (e-mail: cxlcxl1209@163.com; guanjian68@163.com; wjhhhhao@163.com).

Tiange Lan is with the East China Institute of Electronic Engineering, Hefei 230031, China (e-mail: tingerlan@163.com).

Digital Object Identifier 10.1109/JSEN.2024.3496493

lenges to air traffic management and security systems. Some noncooperative UAVs will invade the privacy of individuals or organizations, and their improper use in urban no-fly zones can pose safety threats to citizens [2], [3]. Drones are now considered a new threat to public safety, leading to a sharp increase in the demand for their detection and removal. Due to the similar size and speed of drones and birds, they are easily confused in detection processes [4]. Furthermore, different types of drones may present varying levels of threat, necessitating their differentiation based on specific preventive measures. The use of radar to differentiate between UAVs and birds of prey is currently an effective tool, but there are still significant challenges to more accurately differentiate between low, slow, and small targets [3].

Radar cross sections (RCSs) already provide useful information for target classification; however, the extremely low radar scattering cross-sectional area and slow flight speed of small UAVs increase the difficulty of target detection [5], [6], [7], [8]. When two targets are similar in size or when the target is small, the RCS cannot reliably distinguish between the two targets. Moreover, RCS is influenced by the physical material, size, and shape of the scatterer, which reduces the

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/ accuracy of target classification. Micro-Doppler (m-D) refers to the Doppler frequency shift caused by small components of target other than the main body and was first used by Luo et al. [9] and Hanif et al. [10]. The UAVs' blade and the wings of bird generate Doppler effects in radar echoes. Micro-motion features are closely related to target type and other factors, making m-D an effective feature for classifying UAVs and birds [11], [12], [13]. When analyzing m-D, the frequency in the spectrogram is a crucial factor, as the micro-movements of the target are repeatedly represented in the spectrogram. By performing time-frequency (TF) transformation on the micro-motion signals, the rotation frequency and blade length of the drone's rotor can be estimated, thereby distinguishing the type of target [14]. However, due to the effect of TF resolution, the phenomenon of overlapping UAVs' rotor blades occurs in actual measurements, making it difficult to clearly observe the blinking of each blade, which in turn affects target classification [15]. In addition, feature extraction, a key step, requires a large amount of data, and most feature extraction is manually designed, making it challenging to obtain essential features. To further analyze micro-motion features, the cadence-velocity diagram (CVD) method was introduced. The CVD is obtained by iteratively performing a Fourier transform on each row of the spectrogram [16]. In the literature, CVD has been considered an important feature for classification. Given that MDS is a time-domain Doppler signal, while CVD is a frequency-domain Doppler signal, the CVD method is also effective for analyzing radar echo signals [17].

Radar target classification and recognition tasks have been combined with deep learning to greatly improve target classification and recognition capabilities [18]. Deep learning methods extract high-dimensional features, avoiding the limitations of manual feature extraction [19], [20], [20], [21], [22]. Xu et al. [23] proposed a target-aware recurrent attention network (TARAN) that leverages temporal dependencies between range cells to recognize planar targets, while [24] proposed a reused recursive long short-term memory (RLSTM) network based on Doppler angular trajectories and frequencymodulated continuous-wave (FMCW) to extract temporal and spatial features of echo signal. For 3-D data analysis, 3D-CNN and hybrid complex architectures such as 3D-CNN-LSTM have been proposed, with experiments conducted on radar data collected by ultrawideband (UWB) radar [25]. Rahman and Robertson [26] used the GoogleNet framework to train a model that distinguishes between drones and birds based on their m-D features. Kim et al. [27] combined m-D features and CVD features into a new image and used a CNN structure to improve classification accuracy. Chen et al. [28] and Park et al. [29] used FMCW radar to collect data on drones and birds, merging range spectrogram and m-D features, and used a multichannel DCNN for target classification, which improved the accuracy of drone and bird classification. The existing CNN-based object classification methods do not extract features from the data itself, and there are certain limitations on object classification due to the influence of noise and other factors on micro-motion features.

In this article, a dynamic multifeature fusion network model classification method based on FMCW radar is proposed. The contributions of this article are given as follows. 1657

- We introduced a dynamic data representation in the RP domain: RP sequence tensor data, which allows the extraction of additional time-varying features from RP data. We also constructed a 3-D distance-period dataset, a 2-D distance-period graph dataset, and a 2-D spectrum dataset.
- 2) We developed a deep learning network to extract features from various data types and proposed a novel feature fusion network. This network further extracts features and performs feature fusion through an attention mechanism, enabling the extraction and deep mining of features from the data. The network model dynamic multifeature data fusion neural network (DMFFNNet) is designed for target classification, which improves the classification accuracy compared with a single feature extraction network.
- Extensive experiments were conducted to evaluate the effectiveness of the proposed framework for target recognition. We explored the noise robustness of the network classification using computational data with different signal-to-noise ratio (SNR).

The rest of this article is organized as follows. The signal model of the target is described in Section II. The principles of the proposed DMFFNNet are introduced in Section III. The detailed target classification process and the constructed dataset are presented in Section IV. Comprehensive experimental results and analysis of DMFFNNet are given in Section V. Finally, the conclusions will be drawn in Section VI.

II. SIGNAL MODEL

The FMCW radar echo model is introduced, and the signals of UAVs and flying birds are modeled. An echo model for FMCW radar signals is proposed, as shown in Fig. 1. From this FMCW radar model, the individual signals of this signal can be written as

$$S_k(t_k) = A_k e^{j(2\pi f_c t_k + \pi \phi t_k^2)}$$
(1)

where A_k denotes the amplitude, t_k denotes the fast time referred to the time index within the *m*th chirp, f_c denotes the carrier frequency, and ϕ denotes the slope of the chirp signal. The received echo signal [30] is expressed as

$$S_R(t_k) = \sum_{d=1}^{D} A_d e^{j(2\pi f_c(t_k - \tau_{d,m}) + \pi \phi(t_k - \tau_{d,m})^2)} + n(t_k)$$
(2)

where *D* denotes the number of parts illuminated by radar, A_d denotes the amplitude of signals reflected by different parts, $\tau_{d,m}$ denotes the time delay between the *m*th transmitted signal and the echo signal of the *d*th part, and $n(t_k)$ denotes the noise signal.

A. Signal Model of Rotary Wing UAVs

The model of the radar acquisition UAVs is shown in Fig. 2 Assuming that the velocity does not change during T_c . Then, $S_{\text{Receive}}(t_k)$ and $S_{\text{Transmitting}}(t_k)$ are fed into the mixer simultaneously. The output from the mixer is represented as

$$S_{\text{Multi}}(t_k) = S_{\text{Receive}}(t_k) \cdot \text{conj}(S_{\text{Transmitting}}(t_k))$$
(3)

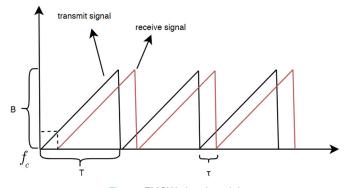


Fig. 1. FMCW signal model.

and the intermediate frequency (IF) signal is obtained by mixing signal through the low-pass filter, which is expressed as

$$S_{\rm IF}(t_k) = \sum_{d=1}^{D} A_k A_d e^{-j(2\pi\phi\tau_{d,m}t_k + 2\pi f_c\tau_{d,m})} + n_{\rm IF}(t_k).$$
(4)

Assuming that the main reflecting components of the UAVs are the body and rotating blades, the IF signal is represented as [31]

$$S_{\rm IF}(t_k, t_s) = A_{\rm body} \, e^{-j \left(2\pi f_R t_k + \frac{4\pi R_0}{\lambda} + 2\pi f_d^{\rm body} t_s\right)} \\ + \sum_{i=1}^{I} \sum_{c=1}^{C} A_{\rm blade}^{i,c} \, e^{-j \left(2\pi f_R t_k + \frac{4\pi R_0}{\lambda} + 2\pi \left(f_{i,c}^{\rm blade} + f_d^{\rm body} t_s\right)\right)} \\ + n_{\rm IF}(t_k, t_s)$$
(5)

where A_{body} denotes the amplitude of UAVs' body, $A_{\text{blade}}^{i,c}$ denotes the amplitude of rotating blades, *i* denotes the number of the rotors, *c* denotes the number of the blades per rotor, t_s denotes the slow-time, f_d^{body} indicates the body Doppler of the UAVs, and $f_{i,c}^{\text{blade}}$ indicates the m-D of rotor. FMCW radar obtains an IF signal by mixing the received and transmitted signals after low-frequency filtering, which is expressed as

$$s_d(f_R, t_s) = A_{\text{body}} e^{-j\left(\frac{4\pi R_0}{\lambda} + 2\pi f_d^{\text{body}} t_s\right)} + \sum_{i=1}^{I} \sum_{c=1}^{C} A_{\text{blade}}^{i,c} e^{-j\left(\frac{4\pi R_0}{\lambda} + 2\pi \left(f_{i,c}^{\text{blade}} + f_d^{\text{body}} t_s\right)\right)} + n_{\text{IF}}(f_R, t_s).$$
(6)

B. Signal Model of Flying Birds

Assuming a bird is flying facing the radar, as shown in Fig. 3, its flapping wings consist of a combination of upward and downward movements. When birds approach the radar, m-D is mainly generated by up and down beat of the wings, and the echo signal of the target is expressed as [32]

$$S_d(f_R, t_s) = A_{\text{body}} e^{-j\left(\frac{4\pi R_0}{\lambda} + 2\pi f_d^{\text{body}} t_s\right)} + 2A_{\text{wing}} e^{-j\left(\frac{4\pi R_0}{\lambda} + 2\pi \left(f_s^{\text{wing}} + f_d^{\text{body}} t_s\right)\right)} + n_{\text{IF}}(f_R, t_s)$$
(7)

where A_{wing} denotes the amplitude of wing, and f_s^{wing} denotes the instantaneous m-D frequency at flapping wing, which is

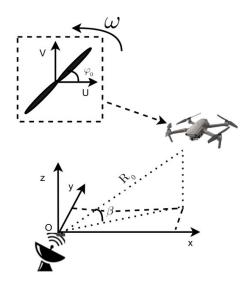


Fig. 2. FMCW radar acquisition of UAVs' echo models.

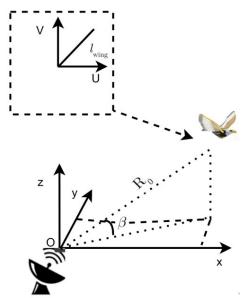


Fig. 3. FMCW radar acquisition bionic bird echo models.

expressed as

$$f_s^{\text{wing}} = -\frac{4l_{\text{wing}}\pi}{\lambda} A_s f_{\text{flap}} \sin(2\pi f_{\text{flap}} t_s)$$
(8)

where A_s , l_{wing} , and f_{flap} denote the range of up and down flapping motions, the length of the single wingspan, and the flapping frequency, respectively. The instantaneous frequency and amplitude of the m-D signal show periodic variations. Doppler is affected by the rotational or beat frequency of each rotor and the initial phase.

III. PRINCIPLE OF DMFFNET AND TARGET CLASSIFICATION

A. Principle of Target Classification Based on DMFFNet

This article proposes a classification method for low, slow, and small targets. First, the data are preprocessed and the dataset is constructed. Then, a network branch is constructed to extract features based on various features of the target echo data. The structure of the target classification network model is shown in Fig. 4, where C1 represents the local feature extraction module, and C2 represents the global feature extraction module. This network model includes feature extraction, feature fusion, and target classification. In the feature extraction stage, the ResNet3D network module, local feature extraction module, and global feature extraction module are used to obtain features from different data. Finally, these extracted features are fused through a feature fusion module to achieve target classification.

B. Global Feature Extraction

The MDI spectrogram contains m-D information from rotor drones, which is an important feature for target classification tasks. The network can fully extract the m-D features of the target while suppressing irrelevant features such as noise and stationary clutter. In addition, due to the differences in rotor speed and blade length among different types of drones, their m-D features also vary. To ensure that the network focuses on more comprehensive and discriminative m-D features, this article introduces an attention module to extract global features from an image in the TF spectrogram. In the global feature extraction branch, a windows multihead self-attention (W-MSA) module is introduced. W-MSA is a module in the Swin Transformer [33] that has a computational complexity that grows linearly with the size of the feature map and uses the prior knowledge of image locality. However, there is a lack of information interaction between windows, so the next module introduces shifted window multihead self-attention (SW-MSA). By shifting the windows toward the bottom-right corner, SW-MSA enables pixel interaction between different windows, allowing better capture of interwindow relationships. This maps the features to a smaller size, which, compared with the multihead self-attention module in transformers, effectively reduces computational costs. The specific module structure is shown in Fig. 5

$$\Omega(\text{MSA}) = 4\hat{h}\hat{w}\hat{C}^2 + 2(\hat{h}\hat{w})^2\hat{C}$$
(9)

$$\Omega(\text{W-MSA}) = 4h\hat{w}C + 2M^2h\hat{w}C \qquad (10)$$

where h denotes the height of the map, \hat{w} denotes the width of the map, \hat{C} represents the depth of the feature map, and M denotes the size of each window. At each stage, the feature map passes through a LayerNorm layer and then into a W-MSA followed by a linear layer with a GELU function. This process is illustrated in the following equation:

$$G_i = f(\text{SW-MSA}(\text{LN}(f(\text{W-MSA}(\text{LN}(G_{i-1})))) + G_{i-1})) + f(\text{W-MSA}(\text{LN}(G_{i-1}))) + G_{i-1} \quad (11)$$

where G_i represents the output of global features, f denotes a convolution operation with a kernel size of 1×1 , and LN represents the LayerNorm operation. Finally, the global features extracted from the TF spectrum are input into the fusion module for feature fusion.

C. Local Feature Extraction

Some local features in feature images are also important. The radar-collected echo data of a rotorcraft UAVs can yield range profile data of the target after pulse compression. Due to differences in the size of target and rotor length, the range profiles of different targets also exhibit some variations. Since these differences in the range profile only exist within a certain range of a specific range unit, we aim to extract these variations from the range profile by ensuring that the network focuses only on local features [34]. By leveraging layer normalization (LN) and GELU activation functions from the transformer, good performance can be achieved across different scenarios. The specific local feature extraction module is shown in Fig. 6. The local features extracted from the range profile are input into the feature fusion module

$$L_i = f_{2 \times 2}(\mathrm{LN}(L_{i-1})) \tag{12}$$

$$\hat{L}_i = f(f(\text{LN}(f_{7\times7}(L_i)))) + L_i$$
 (13)

where L_i is the extracted local feature, and $f_{2\times 2}$ and $f_{7\times 7}$ represent a convolution operation with a kernel size of 2×2 and 7×7 , respectively.

D. Dynamic Feature Extraction

This article introduces the ResNet3D model, as shown in Fig. 7, which effectively captures dynamic changes and spatial structures in radar data by performing convolution operations in both the temporal and spatial dimensions, thereby improving the processing efficiency of radar data. ResNet3D [35] is designed for distance periodic sequence tensor data, aiming to extract dynamic features from the data to improve the classification accuracy. We train ResNet3D to extract spatiotemporal variation features from tensor data. ResNet3D has a specific structure similar to ResNet, but its convolution is extended to 3-D convolution to process tensor data. Specifically, the input data dimension is (batch size, channels, depth, height, width). Through a series of 3-D convolutional layers, 3-D pooling layers, and residual connections, feature representations for classification or regression are ultimately obtained. The input of the network is a matrix that divides continuous range periodic data into 16 frames, allowing the network learn continuous relational features between frames for better extraction of temporal and spatial features, which is crucial for accurately identifying complex data. As shown in Fig. 8, the 3-D convolution kernel slides on the tensor data of the range-periodic sequence, which can capture the long-term dependencies and related features between different frame numbers in the sequence. The dataset is convolved to obtain a 3-D feature map

$$F_i = \operatorname{MaxPool}(f_{3 \times 7 \times 7}(x_i)) \tag{14}$$

$$Stagei = \{Block = \{Conv\}\}$$
(15)

where Maxpool() represents the maximum pooling layer, $f_{3\times7\times7}$ is the convolution operation with a convolution kernel size of $3 \times 7 \times 7$, Stagei denotes fourfold convolution operations, and *i* denotes the number of stage.

E. Multichannel Feature Fusion

The feature fusion module can adaptively fuse different levels of local and global features based on input features.

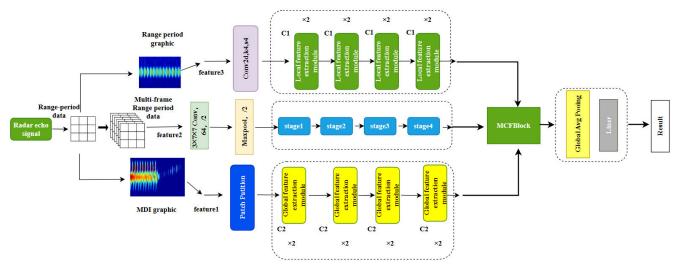


Fig. 4. Multichannel dynamic feature classification model.

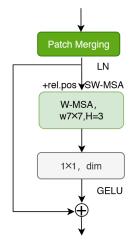


Fig. 5. Global feature extraction module.

This module combines the extracted local features from the distance–periodic graph with the global features from the TF spectrum, as shown in Fig. 9. Since the attention mechanism in global features captures global spatiotemporal information, to some extent, the input global features are captured by the channel attention module (CAM), which uses the interdependence between channel mappings to further extract dependent features. A portion of the features are input to the spatial attention mechanism to enhance local details and suppress irrelevant regions. Finally, feature fusion is performed by splicing the features extracted for each attention. The attention is expressed as

$$CA(x) = \sigma(MLP(AvgPool(x)) + MLP(MaxPool(x)))$$
(16)

$$SA(x) = \sigma(f_{7\times7}(Concat[AvgPool(x), MaxPool(x)])) \quad (17)$$

where σ is the Sigmoid function. The feature fusion operation is denoted as

$$\bar{G}_i = \operatorname{CA}(G_i) \otimes G_i \tag{18}$$

$$L_i = \mathrm{SA}(L_i) \otimes L_i \tag{19}$$

$$T_i = f(\operatorname{Concat}(T_i, G_i, L_i))$$
(20)

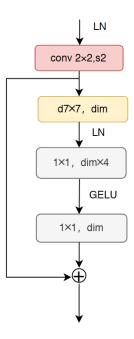


Fig. 6. Local feature extraction module.

$$F_i = \text{LN}(\text{Concat}(\bar{G}_i, \bar{L}_i, \hat{T}_i)) \tag{21}$$

$$\bar{F}_i = f^{3\times3}(F_i) + F_i \tag{22}$$

$$\hat{F}_i = f(f(\bar{F}_i)) + T_i \tag{23}$$

where \otimes denotes element-by-element multiplication, \bar{G}_i is generated by the combination of channel attention, and \bar{L}_i is generated by the combination of spatial attention. \hat{F}_i represents data output after feature fusion.

IV. TARGET CLASSIFICATION METHOD BASED ON DMFFNET

The radar signal is preprocessed to remove the interference in the signal, and the range-period information and TF information in the echo signal are extracted, as shown in Fig. 10. Then the dataset is constructed, and the dataset is input to the designed multifeature extraction network to train and optimize the network parameters, and finally the optimal

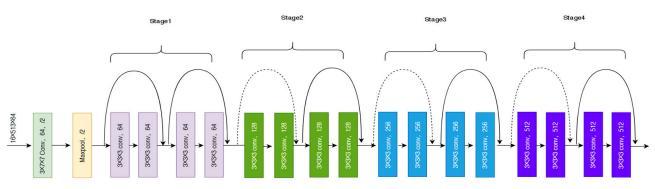


Fig. 7. ResNet3D network framework.

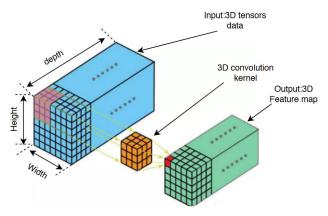


Fig. 8. ResNet3D Convolutional Operation.

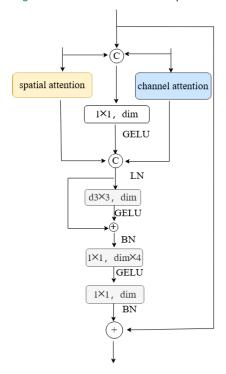


Fig. 9. Multifeature fusion module.

weight obtained by network training is used to complete the task of target classification.

A. Introduction of Radar

The FMCW radar by transmitting continuous wave transmits signal, through the radio frequency module to accept the

TABLE I CONFIGURATION PARAMETERS OF FMCW RADAR

Symbol	Quantity	Configuration	
1	Start Frequency	23.7GHz	
2	modulation period	0.3ms	
3	Bandwidth	100MHz	
4	Sampling Frequency	500MHz	
5	Number of frames	512	
6	Number of Receivers	1	
7	Number of Transmitter	1	

TABLE II SAMPLE SIZE OF DATASET

parameter	Training set	Validation set	Test set
fixed wing UAV	1078	119	47
DJI M350	2384	264	100
hexacopter	1867	207	56
DJI Mavic2	2417	268	100
DJI Inspire2	2220	246	47
Bionic bird	601	66	91

echo signal, and signal mixing, filtering to get the IF signal, the control module to receive the host computer to send parameter commands, to produce the timing control signals required for the transmission and reception, the acquisition module to use multichannel synchronous acquisition of ADC to achieve the collection of echo signals, and through the USB high-speed interface chip, the acquired raw echo signals are transmitted to the host computer for subsequent processing. Finally, through the software module to complete the drive of the host computer system, to ensure that the host computer and the acquisition of normal communication between the card, and have control and display functions, radar processing signal flow as shown in Fig. 11.

This article collects data from various types of rotary wing drones and simulated birds using a K-band FMCW radar, as show in Fig. 12. The working frequency of the radar is 23.7 GHz, and the radar emission waveform is a triangular wave. The specific radar parameters are shown in Table I.

B. Data Preprocessing

In this study, we collected the echo signals of targets using K-band radar and carried out systematic preprocessing steps to improve the accuracy of target detection.

Step 1: First, eliminate the negative frequency signal and dc component in the radar echo data. Then, perform pulse

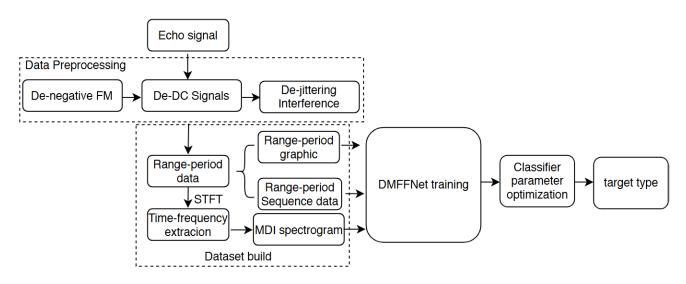


Fig. 10. Classification flowchart of the proposed DMFFNet.

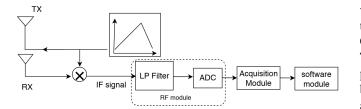


Fig. 11. Radar signal processing flowchart.



Fig. 12. K-band radar.

compression on the original data to obtain a range periodogram containing clutter. Finally, perform moving target indicator (MTI) processing on the data to filter out clutter and obtain the range periodogram of the target.

Step 2: Divide the obtained distance–period data into several data according to the time dimension, store these divided data in a 3-D array, and save them as a mat file to form a 3-D distance–period sequence data.

Step 3: Accurately extract the echo data of the target through the distance-period graph. To further extract the micro-motion features of the target, the short time Fourier transform (STFT) [37] is used to obtain the TF spectrum of the target. Assuming that the target signal obtained through the above steps is S(t), its STFT is expressed as

$$\text{STFT}(t,\omega) = \int_{-\infty}^{+\infty} S(t)g(u-t)\exp(-j\omega t)\,\mathrm{d}\,u \qquad (24)$$

where $g(\cdot)$ represents the sliding window, *t* represents the time dimension, and ω represents the frequency dimension. Construct a dataset of TF spectra of the obtained target. The data matrix is obtained by performing STFT on the preprocessed echo data, and the CFD spectrum [14] is obtained by performing FFT along the time dimension on the TF analysis results

$$CFD = F_t \{STFT\}$$
(25)

where F_t represents performing FFT on the TF graph along the time dimension. The CFD spectrum is equivalent to performing Fourier transform on a series of periodic pulse trains in a TF diagram. The Doppler information dimension in the TF diagram remains unchanged, while the time information is transformed into Cadence frequency information, which represents the repetition frequency, shape, and size information of each Doppler frequency component. Finally, the MDS spectrogram and CFD spectrogram were horizontally spliced to form the MDI spectrogram [27].

The MDI can be expressed as

$$y_{\rm MDI} = [x_{\rm STFT}, x_{\rm CFD}] \tag{26}$$

where y_{MDI} denotes the spliced MDI spectrogram matrix, x_{STFT} denotes the m-D spectrogram matrix, and x_{CFD} denotes the CFD spectrogram matrix.

C. Dataset Construction

In this article, the target echo data collected have been published in Journal of Radars. The dataset [38] can be found at https://radars.ac.cn/web/data/getData?dataType=LLS-LFMCWR. The types of targets collected are shown in Fig. 13. Five different types of rotary wing UAV data are collected simultaneously using K-band and L-band FMCW radar. The target is at the same altitude with the radar in the middle and long distance, and the UAV is in hover state, collecting the echo data of the target through the radar. The collected echo data format is converted into mat format and the files are named according to a rule containing the target distance and



Fig. 13. Types of drones and bionic bird. (a) Fixed wing drone. (b) DJI M350. (c) Hexacopter. (d) DJI Mavic2. (e) DJI Inspire2. (f) Bionic bird.

radar parameters. For different types of rotor UAVs, there are some differences in their range–period graph, among which the range–period echo characteristics are more obvious due to the fact that the fixed wing is a single rotor and the rotational speed is smaller than other types of UAVs, so there are differences in the echo energy with the multirotor UAVs. For quadrotor UAVs, the amplitude of the distance periodogram is different due to the different rotor blade size and length and rotational speed, so there are some differences in the range–period graph.

During the training process, the distance-period map is obtained by preprocessing the radar echo signal, as shown in Fig. 14, and used as a feature to construct a 2-D range-period dataset. After extracting the target signal and performing STFT, TF spectrograms of various types of UAVs are obtained, as shown in Fig. 15. As analyzed in Section II, the bulk Doppler of the target is more prominent than the m-D frequency. There are differences in the MDSs of targets with different micro-motion patterns, and even the MDSs of the same target can vary. We can use these differences to achieve classification and recognition of different targets. To visualize the time-varying characteristics of miniature UAVs, the TF spectrograms of various UAVs are analyzed while they are in a hovering state. The bulk Doppler frequencies of different UAVs are concentrated near zero velocity. By analyzing the MDSs of different types of UAVs, it is observed that the m-D of single-rotor UAVs is more distinct compared with UAVs with different rotor configurations, making it easier to distinguish them from other types of UAVs. Due to the high number of rotors and rapid blade rotation in quadrotor and hexacopter UAVs, the MDSs of the rotors overlap, making it more challenging to differentiate between multirotor UAV types.

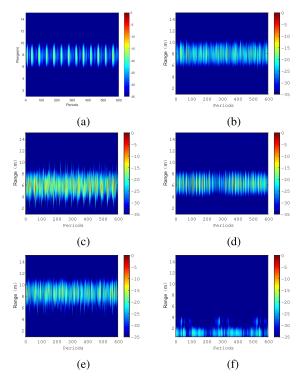


Fig. 14. Range-period graph after MTI. (a) Range-period graph of fixed wing drone. (b) Range-period graph of DJI M350. (c) Range-period graph of hexacopter. (d) Range-period graph of DJI Mavic2. (e) Range-period graph of DJI Inspire2. (f) Range-period graph of flying birds.

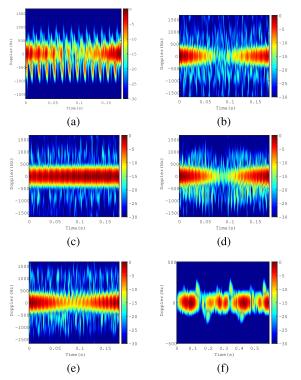


Fig. 15. TF graph of UAVs. (a) TF graph of fixed wing drone. (b) TF graph of DJI M350. (c) TF graph of hexacopter. (d) TF graph of DJI Mavic2. (e) TF graph of DJI Inspire2. (f) TF graph of flying birds.

By calculating the time-domain Fourier transform, the Cadence frequency diagram (CFD) spectrogram is obtained, as shown in Fig. 16. The prominent portions represent the

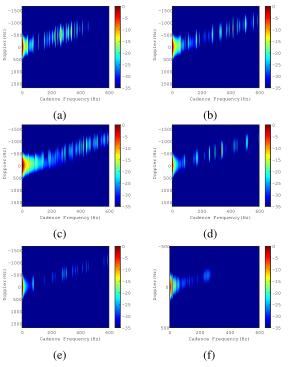


Fig. 16. CFD graph of UAV. (a) CFD graph of fixed wing drone. (b) CFD graph of DJI M350. (c) CFD graph of hexacopter. (d) CFD graph of DJI Mavic2. (e) CFD graph of DJI Inspire2. (f) CFD graph of bionic bird.

dominant components. It can be observed that the dominant components of the CFD for different types of UAVs do not overlap, which aids in distinguishing between different types of UAVs. Since hexacopter UAVs are the largest in size, their dominant components are more pronounced, while the "Mavic" UAV, being the smallest in size, has less prominent dominant components compared with other types of UAVs. The CFD provides information about the shape and frequency of curves in the TF characteristics caused by the moving parts of the target. This method can be used to measure the Doppler frequency repetition rate. Therefore, the TF graph and CFD graph are concatenated together to form an MDI spectrogram to obtain more features and form the dataset. To obtain dynamic features of the data, multiple frames of distance-periodic sequence data were constructed, and dynamic changes in the time dimension were captured through a 3-D network to improve the accuracy of target classification. After preprocessing the collected raw echoes, the IF signal is subjected to fast Fourier transform along the time index to obtain range-period data, which is a covariance matrix representing the distance sampling unit, the number of periods, and the distance offset caused by the motion of target. To extract the dynamic change characteristics of the target [39], we obtain L frames of the range-period sequence within a total imaging time, and then superimpose multiple frames of distance-period sequence data along the slow-time axis to form the range-period sequence tensor data, as shown in Fig. 17.

We set the size of the MDI spectrum dataset to be the same as that of the distance-period dataset. The specific dataset is shown in Table II. We use the training dataset in the table to

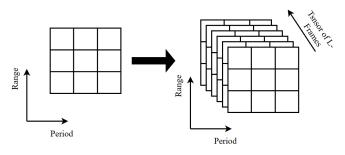


Fig. 17. Flowchart of range-period sequence tensor data.

train the proposed network model and then use the test dataset to test the performance of the trained model.

D. Model Training

This article uses widely used adaptive matrix estimation (AdamW) [40] as the optimizer to adaptively normalize the learning, with hyperparameters of $\beta 1 = 0.9$ and $\beta 2 = 0.999$. The initial learning rate is 3, batch size is set to 32, and epoch is set to 60. This model is built using PyTorch version 1.12.0 and Python version 3.9. The experiment was implemented on a hardware platform: Intel¹ Core² i9-12900KF CPU, 16-GB RAM, NVDIA GeForce RTX 3090 Ti GPU combined with NVIDIA CUDA 12.0. During training, cross-entropy loss is used to calculate the divergence between the real label and the predicted label

Loss =
$$-\frac{1}{N} \sum_{n=1}^{N} \sum_{i=1}^{k} y_i \log(p_i)$$
 (27)

where *N* represents the number of samples, *k* represents the number of categories, y_i represents the true label, and p_i represents the predicted label. This article uses classification accuracy (ACC), precision, and average accuracy as classification metrics. These indicators are all calculated based on confusion matrices. The confusion matrix is represented as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$
(28)

where TP stands for true positive, TN stands for true negative, FP stands for false positive, and FN stands for false negative.

V. EXPERIMENTAL TEST AND RESULTS ANALYSIS

In this section, a large number of experiments were conducted to validate and compare classification networks based on dynamic multifeature fusion. Initial network: The number of training set samples of range–period graph, m-D spectrum, and dynamic data is the same, all set to 10567. The dynamic multifeature data fusion classification model proposed in this article achieved low loss values and high accuracy on measured drone echo data, as shown in Fig. 18. After 60 target classification tasks on the training and validation sets, the training and testing loss values were around 0.0145 and 0.0124,

¹Registered trademark

²Trademarked.

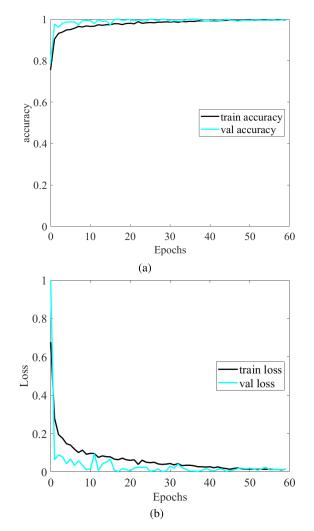


Fig. 18. Accuracy and loss curves for LSS-target classification. (a) Classification accuracy curve. (b) Classification loss curve.

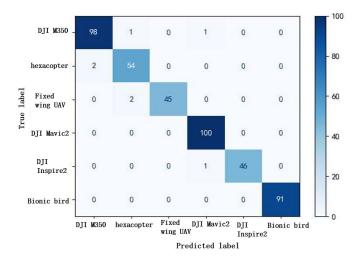


Fig. 19. Classification confusion matrix of the network proposed in this article.

respectively. The average accuracy of training and testing was expressed as 98.9% and 99.2%, respectively. To quantify the proposed framework, the confusion matrix for network classification in this article is shown in Fig. 19.

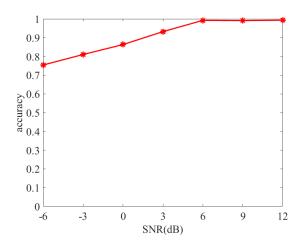


Fig. 20. Accuracy rate of the proposed network under different SNRs.

Due to the use of FMCW radar, the detection range for low altitude observation drones or birds is limited. To demonstrate the universality of the proposed method and further generalize the obtained conclusions, we evaluated the classification ability under different SNRs. The SNR is defined as the average power ratio between the target distance unit and the background unit in the distance-periodic data. Using a dynamic multifeature data fusion model, low and slow small target classification was performed under different SNRs. The classification probability SNR line graph is shown in Fig. 20. When the SNR is not less than 2 dB, the classification probability remains stable at over 90%. To further illustrate the effectiveness of the proposed method, we compared the proposed network with the traditional target classification network. As shown in Fig. 21, we can observe the training loss trend of different networks and the proposed network and the training accuracy of the network. In the initial training stage, the loss value of DMFFNNet is 0.68, while the loss values of other networks are less than 1.6, between 1.2 and 1.6. The loss values of the network have significantly decreased and remained stable. It can be inferred that after 60 epochs of training, the network has almost reached a relatively ideal state.

То compare the performance of using dynamic distance-periodic data with other traditional classification networks, we compared the performance of several commonly used classification networks. The epoch is set to 60, and the best model was selected based on its performance on the test dataset. The results are shown in Table III. The accuracy of the proposed network classification can be seen from the training accuracy of different networks. After 10 epochs, the accuracy of the network reached over 98.6%. Among them, DMFFNNet-1 represents the network without dynamic feature extraction branches. After adding dynamic feature data, the accuracy of DMFFNNet-2 increased by an average of 2.33%. Compared with other traditional classification models, the dynamic feature network model in this article has a higher classification accuracy of 99.56%, which is higher than other classification models. The network, due to its ability to extract more features, has faster convergence and higher accuracy. In terms of parameters, the number of

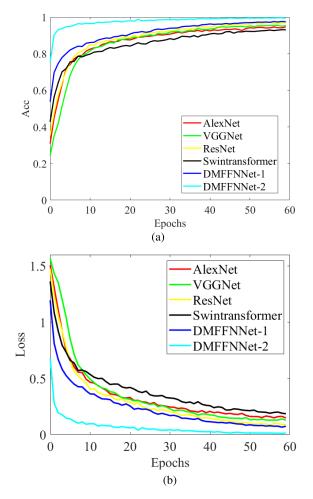


Fig. 21. Accuracy curve and loss curve of each network and the proposed network classification. (a) Classification accuracy curve. (b) Classification loss curve.

TABLE III COMPARISON OF PERFORMANCE OF DIFFERENT NETWORKS

Model	Acc%	Prec%	Parameters(M)	FLOPs(G)
AlexNet	94.20	94.27	14.59	0.51
ResNet50	96.44	96.71	23.52	4.13
VGG19	95.54	95.34	139.59	19.63
Swin-trainsformer	93.40	93.12	86.68	23.55
DMFFNNet-1	97.57	97.23	121.61	18.14
DMFFNNet-2	99.60	99.56	168.10	26.05

network parameters of DMFFNNet is 168.10 M, and the value of FLOPs is the largest, which indicates the higher complexity of this network.

VI. CONCLUSION

This article uses deep learning techniques to conduct research on the classification of low, slow, and small targets, proposing a dynamic feature fusion neural network classification model. Using real measured target signal data, the performance of different feature extraction models and classifiers is tested, improving the accuracy of target classification by extracting dynamic data features from TF spectrograms and range profiles. Data collection was conducted outdoors for five types of rotorcraft UAVs and a bionic bird target, including the DJI Mavic2, DJI Inspire 2, DJI M350, a hexacopter,

and a single-rotor fixed-wing UAV. Based on K-band FMCW radar detection data, range profile datasets, MDI spectrograms datasets, and range profile tensor datasets were constructed. The multichannel network can extract global features from TF spectrograms, local features, and long-term dynamic features from range profile data. These extracted features are then fused using the proposed dynamic multifeature fusion network to classify different types of low, slow, and small targets, thereby improving the classification accuracy. The choice of training set has a significant impact on the performance of the deep learning network. Compared with single-channel CNNs and traditional image processing methods, the multichannel dynamic feature fusion network demonstrates better feature extraction and classification accuracy for target samples. Validation with real measured data from different targets shows that compared with traditional classification methods such as the Swin Transformer, the proposed multichannel dynamic feature fusion network increases the average correct classification accuracy for the six types of targets to 99.56%.

REFERENCES

- M. Pan, C. Chen, X. Yin, and Z. Huang, "UAV-aided emergency environmental monitoring in infrastructure-less areas: LoRa mesh networking approach," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2918–2932, Feb. 2022.
- [2] Z. Nan and H. Zhang, "Experimental research on radar micro-Doppler of flying bird and rotor UAV," *Chin. J. Radio Sci.*, vol. 36, no. 5, pp. 704–714, 2021.
- [3] X. Chen, W. Chen, Y. Rao, Y. Huang, J. Guan, and Y. Dong, "Progress and prospects of radar target detection and recognition technology for flying birds and unmanned aerial vehicles," *J. Radars*, vol. 9, no. 5, pp. 803–827, 2020.
- [4] H. C. Kumawat, M. Chakraborty, A. A. Bazil Raj, and S. V. Dhavale, "DIAT-µSAT: Small aerial targets' micro-Doppler signatures and their classification using CNN," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [5] C. Song, L. Zhou, Y. Wu, and C. Ding, "An estimation method of micro-movement parameters of UAV based on the concentration of timefrequency," *J. Electron. Inf. Technol.*, vol. 42, no. 8, pp. 2029–2036, 2020.
- [6] S. Björklund, "Target detection and classification of small drones by boosting on radar micro-Doppler," in *Proc. 15th Eur. Radar Conf.* (*EuRAD*), Sep. 2018, pp. 182–185.
- [7] F. Gao, A. Liu, K. Liu, E. Yang, and A. Hussain, "A novel visual attention method for target detection from SAR images," *Chin. J. Aeronaut.*, vol. 32, no. 8, pp. 1946–1958, Aug. 2019.
- [8] R. Amiri and A. Shahzadi, "Micro-Doppler based target classification in ground surveillance radar systems," *Digit. Signal Process.*, vol. 101, Jun. 2020, Art. no. 102702.
- [9] Y. Luo, Y. Chen, Y. Zhu, W. Li, and Q. Zhang, "Doppler effect and micro-Doppler effect of vortex-electromagnetic-wavebased radar," *IET Radar, Sonar Navigat.*, vol. 14, no. 1, pp. 2–9, Jan. 2020.
- [10] A. Hanif, M. Muaz, A. Hasan, and M. Adeel, "Micro-Doppler based target recognition with radars: A review," *IEEE Sensors J.*, vol. 22, no. 4, pp. 2948–2961, Feb. 2022.
- [11] L. Zhu, S. Zhang, H. Zhao, S. Chen, D. Wei, and X. Lu, "Classification of UAV-to-ground vehicles based on micro-Doppler signatures using singular value decomposition and deep convolutional neural networks," *IEEE Access*, vol. 7, pp. 22133–22143, 2019.
- [12] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," *IEEE Signal Process. Mag.*, vol. 36, no. 4, pp. 16–28, Jul. 2019.
- [13] C. Bennett, M. Jahangir, F. Fioranelli, B. I. Ahmad, and J. L. Kernec, "Use of symmetrical peak extraction in drone micro-Doppler classification for staring radar," in *Proc. IEEE Radar Conf.*, Florence, Italy, Sep. 2020, pp. 1–6.

1667

- [14] J. Sun and X. Wang, "Micro-motion feature extraction of micro-rotor UAV based on RSP-CFD method," J. Signal Process., vol. 37, no. 3, pp. 399–408, 2021.
- [15] Y. M. Kassim et al., "Small object bird detection in infrared drone videos using mask R-CNN deep learning," *Electron. Imag.*, vol. 32, no. 8, pp. 85-1–85-8, Jan. 2020.
- [16] X. Zhang, X. H. Wang, X. P. Xu, and Y. C. Zhao, "Demand learning and cooperative deployment of UAV networks," *Chin. J. Electron.*, vol. 31, no. 3, pp. 408–415, May 2022.
- [17] S. Björklund, T. Johansson, and H. Petersson, "Evaluation of a micro-Doppler classification method on mm-wave data," in *Proc. IEEE Radar Conf.*, Atlanta, GA, USA, May 2012, pp. 0934–0939.
- [18] M. Z. Ozturk, C. Wu, B. Wang, and K. R. Liu, "GaitCube: Deep data cube learning for human recognition with millimeterwave radio," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 546–557, Jan. 2022.
- [19] Y. Zhang, H. Yuan, H. Li, J. Chen, and M. Niu, "Meta-learner-based stacking network on space target recognition for ISAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 12132–12148, 2021.
- [20] A. Schumann, L. Sommer, J. Klatte, T. Schuchert, and J. Beyerer, "Deep cross-domain flying object classification for robust UAV detection," in *Proc. 14th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Lecce, Italy, Mar. 2017, pp. 1–6.
- [21] A. J. Garcia, A. Aouto, J.-M. Lee, and D.-S. Kim, "CNN-32DC: An improved radar-based drone recognition system based on convolutional neural network," *ICT Exp.*, vol. 8, no. 4, pp. 606–610, 2022.
- [22] W. S. Chen, J. Liu, and J. Li, "Classification of UAV and bird target in low-altitude airspace with surveillance radar data," *Aeronaut. J.*, vol. 123, no. 1260, pp. 191–211, Feb. 2019.
- [23] B. Xu, B. Chen, J. Wan, H. Liu, and L. Jin, "Target-aware recurrent attentional network for radar HRRP target recognition," *Signal Process.*, vol. 155, pp. 268–280, Feb. 2019.
- [24] Z. Yang and X. Zheng, "Hand gesture recognition based on trajectories features and computation-efficient reused LSTM network," *IEEE Sensors J.*, vol. 21, no. 15, pp. 16945–16960, Aug. 2021.
- [25] N. Hendy, H. M. Fayek, and A. Al-Hourani, "Deep learning approaches for air-writing using single UWB radar," *IEEE Sensors J.*, vol. 22, no. 12, pp. 11989–12001, Jun. 2022.
- [26] S. Rahman and D. A. Robertson, "Classification of drones and birds using convolutional neural networks applied to radar micro-Doppler spectrogram images," *IET Radar, Sonar Navigat.*, vol. 14, no. 5, pp. 653–661, May 2020.
- [27] B. K. Kim, H.-S. Kang, and S.-O. Park, "Drone classification using convolutional neural networks with merged Doppler images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 38–42, Jan. 2017.
- [28] X. Chen, H. Zhang, J. Song, J. Guan, J. Li, and Z. He, "Micro-motion classification of flying bird and rotor drones via data augmentation and modified multi-scale CNN," *Remote Sens.*, vol. 14, no. 5, p. 1107, Feb. 2022.
- [29] D. Park, S. Lee, S. Park, and N. Kwak, "Radar-spectrogram-based UAV classification using convolutional neural networks," *Sensors*, vol. 21, no. 1, p. 210, Dec. 2020.
- [30] C. Zhao, G. Luo, Y. Wang, C. Chen, and Z. Wu, "UAV recognition based on micro-Doppler dynamic attribute-guided augmentation algorithm," *Remote Sens.*, vol. 13, no. 6, p. 1205, Mar. 2021.
- [31] Z. Wu, H. Liu, C. Ma, and Z. Liu, "Aerial target recognition with enhanced micro-Doppler dynamic features based on frequencymodulated continuous-wave radar," *IEEE Sensors J.*, vol. 23, no. 19, pp. 23119–23132, Oct. 2023.
- [32] Y. Sun, S. Abeywickrama, L. Jayasinghe, C. Yuen, J. Chen, and M. Zhang, "Micro-Doppler signature-based detection, classification, and localization of small UAV with long short-term memory neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6285–6300, Aug. 2021.
- [33] Z. Liu et al., "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 10012–10022.
- [34] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* (CVPR), Jun. 2022, pp. 11966–11976.

- [35] T. Jin et al., "UWB-HA4D-1.0: Ultra-wideband radar human activity 4D imaging dataset," J. Radars, vol. 11, no. 1, pp. 27–39, 2022.
- [36] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2018, pp. 7132–7141.
- [37] G. Yu, M. Yu, and C. Xu, "Synchroextracting transform," *IEEE Trans. Ind. Electron.*, vol. 64, no. 10, pp. 8042–8054, Oct. 2017.
- [38] X. Chen et al., "Multiband FMCW radar LSS-target detection dataset (LSS-FMCWR-1.0) and high-resolution micromotion feature extraction method," J. Radars, vol. 13, no. 3, pp. 539–553, 2024.
- [39] L. Han, C. Feng, X. Hu, S. He, and X. Xu, "Ballistic target recognition based on multiple data representations and deep-learning algorithms," *Chin. J. Aeronaut.*, vol. 37, no. 6, pp. 167–181, Jun. 2024.
- [40] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," 2017, arXiv:1711.05101.



Wang Yuan was born in Jining, Shandong, China, in 1997. He received the B.S. degree in computer science and technology from Taishan University, Taian, China, in 2019. He is currently pursuing the M.S. degree in computer and control engineering with Yantai Yantai University, Yantai, China.

His research interests include signal processing and deep learning.



Xiaolong Chen (Senior Member, IEEE) was born in Yantai, Shandong, China, in 1985. He received the bachelor's and master's degrees in signal and information processing and the Ph.D. degree in radar signal processing from Naval Aviation University (NAU), Yantai, in 2008, 2010, and 2014, respectively.

From 2018 to 2022, he was an Associate Professor with the Marine Target Detection Research Group, NAU, where he is currently a Professor. He has published more than 120 aca-

demic articles and three books, and holds 44 national invention patents. He has given more than 30 presentations on radar signal processing, especially marine targets. His research interests include radar signal processing, especially for marine target detection, moving target detection, micro-Doppler, and clutter suppression and also deep learning for radar target detection.

Prof. Chen was a recipient of the Excellent Young Scholar from the National Natural Science Foundation of China. He has been with the Editorial Board of Journal of Radars since 2019 and has been serving as an Associate Editor for IEEE ACCESS since 2018. He is also a reviewer of IEEE TRANSACTIONS ON SIGNAL PROCESSING (TSP), IEEE SIGNAL PROCESSING LETTERS (SPL), IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), IEEE GEOSCIENCE AND REMOTE SENSING (TGRS), IEEE GEOSCIENCE AND REMOTE SENSING (GRSL), IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (J-STARS), *IET Radar, Sonar and Navigation* (RSN), *IET Signal Processing* (DSP), and many international conferences.



Xiaolin Du (Member, IEEE) received the B.S. and M.S. degrees from Taiyuan University of Technology, Taiyuan, China, in 2008 and 2011, respectively, and the Ph.D. degree from Xidian University, Xi'an, China, in 2015.

From 2019 to 2020, he was a Post-Doctoral Researcher with the University of Electronic Science and Technology of China, Chengdu, China. From 2019 to 2020, he was a Visiting Scholar with the Department of Electrical Engineering and Information Technology, University

of Naples "Federico II," Naples, Italy. Since 2016, he has been a faculty with the School of Computer and Control Engineering, Yantai University, Yantai, China, where he is currently an Associate Professor. His research interests include radar signal processing, covariance matrix estimation, and waveform design.



Jinhao Wang was born in Shangqiu, Henan, China, in 1999. He received the B.E. degree in underwater acoustic engineering from Naval Aviation University, Yantai, China, in 2023, where he is currently pursuing the M.S. degree in electronic information.

His research interests include radar "low, slow, and small" target signal processing.



Jian Guan received the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2000.

He is currently a Professor with NAU, Yantai, China. He has authored numerous articles in his areas of expertise and holds 21 national invention patents. His research interests include radar target detection and tracking, image processing, and information fusion.

Prof. Guan is also a Senior Member of CIE and a Committee Member of the Radio Positioning

Technology Branch in CIE. He has won the prize of the National Excellent Doctoral Dissertation, "Realistic Outstanding Youth Practical Engineering Award" of CAST, and was selected for the National Talents Engineering of Ministry of Personnel of China. He has served in the technical committee for many international conferences on radar.



Tiange Lan received the Ph.D. degree in optics from Hefei Institutes of Physical Science, CAS, Hefei, China, in 2008.

He is currently a high-grade Engineer with East China Institute of Electronic Engineering, Hefei. Since 2010, he has been primarily engaged in the demonstration, design, and testing of digital array radar, "low, slow, small" target detection radar, and early warning detection system.