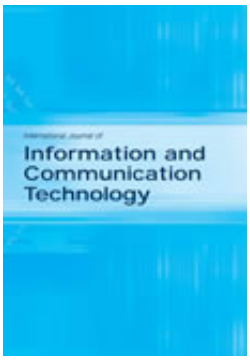
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research




**International Journal of Information and Communication Technology**

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

# An end-to-end radar emitter denoising and recognition method using batch norm removal

Zixin Jiang, Xin Zhao, Bing He, Zhenzhen Wang, Weijie Kang, Jie Zhang

**Article History:**

| | |
|---|---|
| Received: | 22 February 2024 |
| Last revised: | 12 April 2024 |
| Accepted: | 30 April 2024 |
| Published online: | 20 January 2025 |

# An end-to-end radar emitter denoising and recognition method using batch norm removal

Zixin Jiang, Xin Zhao, Bing He*,
Zhenzhen Wang, Weijie Kang and Jie Zhang

Department of Electronic Engineering,
PLA Rocket Force University of Engineering,
Xi'an, China
Email: jiangzixin0214@163.com
Email: zhaoxin20062111@163.com
Email: hebing202301@163.com
Email: wangzhenzhen202404@163.com
Email: milankang@foxmail.com
Email: zhangjie@opt.ac.cn
*Corresponding author

**Abstract:** Radar emitter recognition is a critical part of electronic countermeasures and determines the implementation of subsequent jamming measures. With the rapid development of deep learning (DL), the radar emitter recognition method based on DL is also widely developed. However, methods based on time-frequency analysis to obtain image features suffered from information loss and computational complexity; end-to-end methods based on raw radar signals had low accuracy at low signal-to-noise ratio (SNR). Therefore, we investigated the frequency distribution of signal and noise, analysed the working principle of batch norm, and proposed to suppress the high-frequency noise by removing the batch norm in the network. Furthermore, we constructed a straightforward end-to-end denoising and recognition network as well as utilised the latest classification network training process to improve the accuracy of radar emitter recognition at low SNR. Experiments validated that the proposed method achieved SOTA result on the well-known DeepSig RadioML 2018.01A.

**Keywords:** radar emitter recognition; signal denoising; deep learning; end-to-end; time series.

**Reference** to this paper should be made as follows: Jiang, Z., Zhao, X., He, B., Wang, Z., Kang, W. and Zhang, J. (2025) 'An end-to-end radar emitter denoising and recognition method using batch norm removal', *Int. J. Information and Communication Technology*, Vol. 26, No. 1, pp.23–37.

**Biographical notes:** Zixin Jiang received his BS in Electronic Information Engineering from the Air Force Engineering University (AFEU), China, in 2022 and currently pursuing his PhD in Electronic Information Engineering at the Rocket Force University of Engineering (RFUE). His research interests include signal processing, deep learing and distributed estimation.

Xin Zhao received his MS and PhD in Control Science and Engineering from the Rocket Force University of Engineering (RFUE), China, in 2008 and 2012, respectively, where he is currently an Associate Professor. His research interests include signal process, flight guidance and control theory.

Bing He received his MS and PhD in Armament Science and Technology from the Rocket Force University of Engineering (RFUE), China, in 2008 and 2012, respectively, where he is currently an Associate Professor. His research interests include flight guidance and trajectory planning.

Zhenzhen Wang received his BS in Electronic Information Engineering from the Air Force Engineering University (AFEU), China, in 2022 and currently pursuing his PhD in Electronic Information Engineering at the Rocket Force University of Engineering (RFUE). His research interests include control theory, meta learing and distributed estimation.

Weijie Kang received his BS in Computer Science and Technology from the Xi'an Jiaotong University, China, in 2017 and received his PhD in Weapon Science and Technology from the Air Force Engineering University (AFEU), China, in 2022. Currently, he is a Lecturer in the Rocket Force University of Engineering (RFUE), Chain. His research interests include electronic jamming and equipment reliability.

Jie Zhang received his BS and MS from the Xi'an University of Architecture and Technology in 2016 and he received his PhD in Aeronautical and Astronautical Sciences and Technology from the Rocket Force University of Engineering (RFUE), China, in 2024. His main research interests are X-ray tomography reconstruction and solid mechanics analysis.

## 1 Introduction

Radar systems have a wide range of applications in both military and civilian fields, such as air traffic control, weather monitoring and battlefield surveillance. For both military and civil applications, the ability to accurately identify different radar emitters is critical and important for mission success. In the civilian field, such as air traffic control, accurate recognition of radar emitter is critical to avoiding airborne collisions and improving aviation safety. In the military field, radar emitter recognition is one of the key technologies in electronic reconnaissance and electronic countermeasure systems, and the accuracy of this technology directly affects how well electronic reconnaissance and electronic countermeasures work (Liu et al., 2005). Accurate recognition of radar emitter helps to grasp the radar's working status, the level of power and ribbing and other information and provide intelligence support for battlefield situational awareness, enemy recognition, threat alerts, and battle plan development (Liu, 2020). It is crucial for defense, surveillance and offense, and is an important factor affecting the victory or defeat of electronic countermeasures under the complex battlefield environment.

The traditional radar emitter recognition method mainly relies on the pulse description word, including carrier frequency, pulse amplitude, pulse repetition interval, pulse width and other information, while recognising the radar emitter through expert experience judgment (Yang et al., 2023). However, with the rapid development of

phased array technology and the wide application of multifunctional radar, the density and complexity of radar signals have increased dramatically, and the radar parameters and waveform styles have become more complex and diversified. In this case, the shortcomings of the traditional radar emitter recognition methods have come to the fore. Coupled with the increasing complexity of the electromagnetic environment, the number of electromagnetic signals has increased exponentially, and most radars also adopt low interception probability, making the waveform characteristics become complex and diverse, which also brings difficulties to the traditional emitter recognition methods. In the current environment, the traditional methods face the problems of diverse emitter, similar signals and large noise, which leads to their poor adaptability, weak generalisation ability, and poor universality, and makes it difficult to meet the needs of emitter recognition in complex electromagnetic environments (Sun et al., 2020).

Deep learning (DL), with its excellent feature extraction capability, has made remarkable achievements in the fields of image segmentation (Zhu et al., 2019), object detection (Wu et al., 2022), speech recognition (Shahamiri et al., 2021), and hyperspectral image classification (He et al., 2018). This has led to a rapid development in the research of DL-based radar emitter recognition methods. Time-frequency analysis (López-Risueño et al., 2005) is a common emitter recognition method, which converts the radar sequence signal into a 2D image through time-frequency analysis. Then the image is preprocessed as an input to a deep neural network (DNN), which in turn determines the radar signal category. For example, Huynh-The et al. classified radar signals by using the Choi-Williams distribution based on converting the original radar signals into images and then constructing an LPI-Net to extract image features (Huynh-The et al., 2021). Zhang et al. firstly converted the signal into two kinds of time-frequency images by smooth pseudo wignerville distribution (SPWVD) and Born-Jordan distribution and the image features are extracted by using convolutional neural network (CNN) (Zhang et al., 2019). Subsequently, the joint features are formed by image features and hand-crafted features, and the joint features are fused by applying the multimodality fusion model, which leads to the recognition of radar signals. Xiao et al. converted the original radar signal into a 2D image by short time Fourier transform (STFT), then denoised the image using k-mean algorithm, and subsequently extracted the image features using CNN to identify the radar signal (Xiao and Yan, 2021). Yu et al. convert the radar sequence signal into a frequency domain image by STFT, and then use a deep normalised convolutional neural network (DNCNN) to reduce the noise of the image, secondly use the denoised image to train the established classification model which finally recognises the radar signals using the trained model (Yu et al., 2022). However, converting raw radar signals to 2D images not only leads to loss of information but also increases computational complexity.

Besides the time-frequency analysis method, it is also possible to identify the radar emitters directly from the original radar sequence signals. Zhang et al. proposed a high-order convolutional attention network (HoCAN) based on the higher-order attention mechanism, which introduces a nonlinearly transformed covariance matrix and a higher-order convolutional layer to improve the discriminative power of the signals and the classification accuracy (Zhang et al., 2023). Wang et al. trained two convolutional neural networks using different datasets, where one of the CNNs was trained on samples consisting of two I/Q signals to identify simple modulations and sort out difficult modulations, and the other CNN was trained on samples consisting of signal constellation diagrams to identify difficult modulations (Wang et al., 2019). Huynh-The

et al. proposed a new CNN which is designed with several specific convolutional blocks to learn the correlation of spatio-temporal signals simultaneously via different asymmetric convolutional kernels for automatic modulation recognition (Huynh-The et al., 2020). Although the above methods achieve end-to-end recognition, none of them considered the effect of noise. However, in actualised environments, the SNR of the received signal is low since the signal is usually affected by environmental noise, interference, atmospheric attenuation and other factors during transmission, which makes denoising a key step in improving recognition accuracy.

In the radar emitter recognition task, denoising is a key link because it directly affects the subsequent recognition accuracy. Effective denoising not only improves the quality of the signal, but also enhances the ability to resist interference and noise, thus improving the overall performance of the model. Therefore, finding a method that can effectively reduce the signal noise is an inevitable way to improve the accuracy of emitter recognition. Du et al. proposed a denoising classification network DNCNet for radar emitter recognition consisting of denoising and classification sub-networks. And they used a two-phase training approach, where the first phase trains the denoising sub-network, and the second phase strengthens the mapping between the denoising results and the perceptual representations (Du et al., 2022). Wang et al. used a relatively long signal can implicitly perform noise reduction in a statistical sense, including two steps of constructing sampling points and smoothing filtering (Wang and Gan, 2022). Although the above method achieved reasonable denoising results and improved the classification accuracy, its performance at low SNR is still unsatisfactory. Effective denoising methods should not only suppress the noise, but also ensure that no important information is lost. In this paper, we proposed a simple and direct denoising and recognition method for the problems of DL-based radar emitter recognition method, constructed an end-to-end network, used a new network training method, and improved the recognition accuracy and robustness of the model. The main contributions are summarised as follows:

- Our experiment showed that the network improvement which is not specific to the dataset is not the main reason for the improvement of the classification accuracy. We analysed the signal and the noise in the frequency domain, investigated the frequency distributions of both, and found that suppressing the noise in the signal is what is necessary to improve the accuracy.
- We analysed the function and principle of BN, elaborated its performance at low and high frequencies when the CNN performs the classification task. Therefore, we proposed a method to suppress the high-frequency noise by removing the BN layer, and constructed an simple and straightforward end-to-end denoising and classification network.
- We trained the network using the latest exponential moving average (EMA) method, verified that the proposed method achieved SOTA results on the RadioML 2018.01A. We also validated the applicability of the method using the VGG16 network, and compared the proposed method with other classification methods.

## 2 Problem formulation

### 2.1 Radar signal

In a practical environment, the radar signal received by a receiver can be represented as:

$$x(t) = s(t) + n(t), \tag{1}$$

where, $s(t)$ denotes the radar signal, $n(t)$ denotes the additive noise signal, which is mostly Gaussian white noise in modelling. In order to facilitate computer processing, we discretise the received signal as (Zhao et al., 2021):

$$x(l) = s(l) + n(l),\ 1 \leq l \leq L, \tag{2}$$

where $l \in N$ denotes the sampling point and $L$ denotes the length of the signal. The radar signal $s(l)$ can be expressed as:

$$s(l) = Ae^{j\phi}, \tag{3}$$

$$\phi = 2\pi f_0 l + \varphi(l) + \varphi_0, \tag{4}$$

where $A$ is the amplitude of the radar signal, usually a constant, $f_0$ is the carrier frequency of the radar signal, $\varphi_0$ is the initial phase, $\varphi(l)$ indicates the intra-pulse modulation information of the radar signal. Differences between different emitter are mainly reflected through $\varphi(l)$. The intra-pulse modulation of emitter can be analogue modulation and digital modulation. Analogue modulation is mainly single-frequency modulation (SFM), linear frequency modulation (LFM), nonlinear frequency modulation (NLFM), etc. Digital modulation is mainly phase shift keying (PSK), frequency shift keying (FSK), amplitude shift keying (ASK) and so on.

### 2.2 Noise signal

In radar receivers, noise signals are generally divided into Gaussian noise signals and non-Gaussian noise signals. Gaussian noise signal is the noise signal whose probability density function (PDF) obeys Gaussian distribution, which can be divided into Gaussian white noise and Gaussian colour noise. Non-Gaussian noise signals, also called impulse noise, have a PDF that does not obey a Gaussian distribution, and are usually represented by an alpha-stable distribution in radar and communications.

Thermal noise in most electronic systems such as radar and communication systems is typically additive white Gaussian noise (AWGN). AWGN is a random process consisting of a series of uncorrelated random variables with a constant value of power spectral density (PSD), and is the simplest type of noise (Préaux and Boudraa, 2020). Assuming that Gaussian white noise obeys a Gaussian distribution that satisfies $X \sim N(\mu, \sigma)$, its PDF can be expressed as:

$$P_1(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \tag{5}$$

In addition to the thermal noise inherent in electronic systems, the noise contained in data measured in engineering practice is often coloured noise. Unlike Gaussian white noise, the noise amplitude at any moment in Gaussian coloured noise is not independent and is correlated with the noise amplitudes at other moments, and its PSD is not constant (Wen et al., 2017). Under discrete conditions, assuming that the autocorrelation function of Gaussian coloured noise is $R(l)$, its PSD can be expressed as the Fourier transform of the autocorrelation function according to the Wiener-Khinchin theorem:

$$P_2(\omega) = \sum_{p\in\mathbb{Z}} R(l)\exp(-j\omega l). \tag{6}$$

Except for a few exceptions, there is no closed analytical formula for the probability density of alpha-stable distribution, so it is generally represented by the characteristic function. Alpha-stable distribution is determined by four parameters $(\alpha, \beta, \gamma, a)$, $\alpha$ denoted by the characteristic index, which describes the degree of shock of alpha-stable distribution. $\beta$ is the symmetry parameter, which describes the degree of distortion of alpha-stable distribution. $\gamma$ is the scale parameter, which is equivalent to the variance in Gaussian noise, $a$ is the position parameter, and its characteristic function can be expressed as (Liu et al., 2024):

$$\varphi(t) = \exp\{jat - \gamma \mid t \mid^{\alpha} [1 + j\beta\,\mathrm{sgn}(u)w(t,\alpha)]\}, \tag{7}$$

$$w(t,\alpha) = \begin{cases} \frac{2}{\pi}\log \mid t \mid, & \alpha = 1 \\ \tan\frac{\alpha\pi}{2}, & \alpha \neq 1 \end{cases} \tag{8}$$

$$\mathrm{sgn}(u) = \begin{cases} 1, & u > 0 \\ 0, & u = 0 \\ -1, & u < 0 \end{cases} \tag{9}$$

### 2.3 Signal to noise ratio

As the name suggests, the SNR is the ratio of signal power to noise power, usually expressed in dB (Cosman et al., 1994). Its calculation formula is:

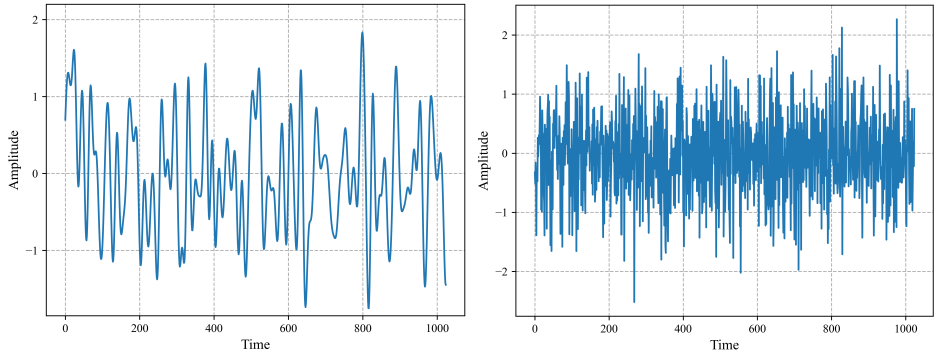
$$SNR = 10lg\frac{P_s}{P_n}, \tag{10}$$

where $P_s$ is the power of the signal and $P_n$ is the noise power. The smaller the SNR, the stronger the noise power in the received signal and the weaker the signal power.

As shown in Figure 1(a) is the signal with SNR equal to 30 dB in the time domain image, and Figure 1(b) is the signal with SNR equal to –20 dB in the time domain image. It can be clearly seen that the signal is annihilated by the noise in the –20 dB condition very seriously, and it is very difficult to distinguish between the two, which is the reason why several end-to-end methods described in Section 1 are ineffective in the case of low SNR. We perform a fast Fourier transform (FFT) on the signal in Figure 1 and convert it to the frequency domain for analysis, as shown in Figure 2. It can be seen that the signals with SNR equal to 30 dB, shown by the orange line in the

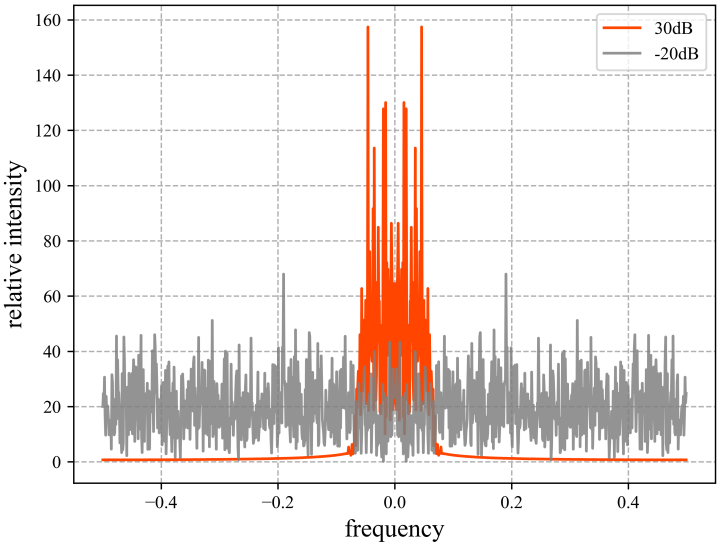
figure, are concentrated in the low frequency range, while the signals with SNR equal to –20 dB condition, shown by the blue line, are distributed in all frequency ranges. This is because the signal in the –20 dB is mainly noise, which is present over the entire frequency band and is more prominent at high frequencies. Therefore, we can use the difference in the frequency distribution of the noise and the signal to suppress the noise.

**Figure 1** The signal with SNR equal to 30 dB and –20 dB in the time domain, (a) the signal with SNR = 30 dB (b) the signal with SNR = –20 dB (see online version for colours)




(a) (b)

**Figure 2** The signal with SNR = 30 dB and SNR = –20 dB in the frequency domain (see online version for colours)



## 3 Proposed method

### 3.1 Batch norm

Batch norm (BN) is a widely used technique to train DNNs faster and more consistently (Santurkar et al., 2018). Ioffe and Szegedy proposed BN in 2015, where they included normalisation as part of the model architecture and performed normalisation in each training mini-batch (Ioffe and Szegedy, 2015). BN performs a certain regularisation function and successfully prevents gradient explosion and vanishing. Currently, BN has been derived from layer normalisation, instance normalisation, group normalisation, switchable normalisation, and various other normalisation processes. According to Ioffe and Szegedy (2015), the purpose of BN is to make the inputs of each layer of the network satisfy the distribution law of zero mean and unit variance.

Therefore, the BN layers is in between the two layers of the network and its role is to take the output from the upper layer, normalise it and pass it on to the next layer. Assuming that the $i^{\text{th}}$ layer has $n$ channels and the input is an n-dimensional data $X_i = (x^{(1)}...x^{(n)})$, then the BN connected to that layer will normalise each dimension before feeding it into the $i^{\text{th}}$ layer. First normalises the evey $x$ to get $\hat{x}$ with zero mean and unit variance:

$$\hat{x}^{(m)} = \frac{x^{(m)} - E[x^{(m)}]}{\sqrt{\text{Var}[x^{(m)}]}} \; m \leq n, m \in \mathbb{R}, \tag{11}$$

where $E[x^{(m)}]$ and $\text{Var}[x^{(m)}]$ are computed on the training set. In order not to reduce the representational power of the network, each $\hat{x}$ is scaled and shifted:

$$y^{(m)} = \gamma^{(m)}\hat{x}^{(m)} + \beta^{(m)}, \tag{12}$$

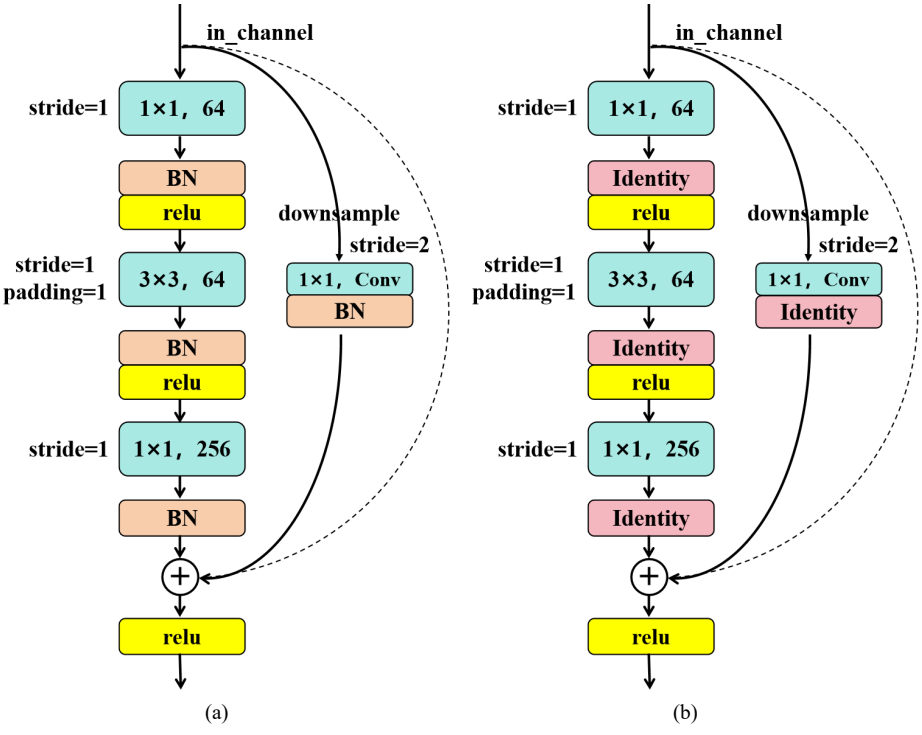
where $\gamma^{(m)}$ and $\beta^{(m)}$ are set parameters, both of which are updated along with the original model parameters as the gradient is backpropagated. If $\gamma^{(m)}$ and $\beta^{(m)}$ eventually converge to $\sqrt{\text{Var}[x^{(m)}]}$ and $E[x^{(m)}]$, respectively, then the BN will recover the original data $x^{(m)}$. Finally, all $y^{(m)}$ will be fed into the $i^{\text{th}}$ layer by composing a new distribution $Y_i = (y^{(1)}...y^{(n)})$.

Basically almost CNN contain BN because of its excellent performance, e.g., ResNet, VGG, etc. and BN is especially widely used in classification problems. However, not all classification problems are suitable for BN. Wang et al. found that BN improves the utilisation of both low-frequency and high-frequency information. And the high-frequency part is generally smaller in amplitude, and BN helps CNN to utilise the high-frequency part well. And if only utilising the low-frequency part, the BN does not always improve the model capability (Wang et al., 2020). As described in Subsection 2.3, they are different frequency distributions of the noise and the signal, with the frequency of the signal tending to be in the lower part of the frequency spectrum and the noise being more prominent at higher frequencies. Thus, the model should make more use of the low-frequency information and discard the high-frequency information appropriately when recognising the signal. Based on this, we remove the BN layers in the model to achieve the purpose of denoising, making the model better extract signal features and improving the recognition ability of the model. In Du et al. (2022), the BN layers was also removed, but it was only empirically found that BN was not very helpful for the classification task, and no theoretical justification was stated.

### *3.2 Network architecture*

We used ResNet50 (He et al., 2016) as the basic architecture of the model with some adjustments to enable it to handle the signal data from the experiments. Since the input is an I/Q signal with a $2 \times N$ format, we adjusted both the convolutional layers and pooling layers to 1D, defined as ResNet50_1D. Specifically, according to the previous section, we adapted the bottleneck layer in ResNet50_1D by replacing the BN layers with the identity layers, and constructed an end-to-end denoised and classification network, which reduces the influence of noise and improves the classification accuracy, as shown in Figure 3.

**Figure 3** Primitive and improved bottleneck architecture, (a) the primitive bottleneck in ResNet50_1D (b) the improved bottleneck in ResNet50_1D (see online version for colours)



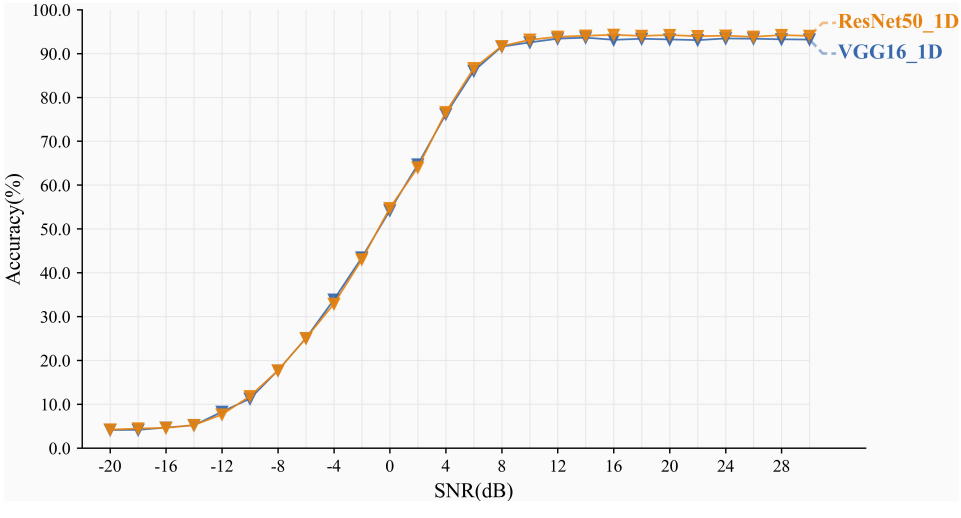

We used the weight averaging (Izmailov et al., 2018) technique and the averaging method used was polyak averaging, which can also be referred to as EMA method, to improve the classification accuracy of the model. EMA is an averaging method that gives higher weights to recent data, which reduces the training time by reducing the number of weight updates required, and utilises the parameters of sliding averaging to improve model robustness and enhance the performance of the model on the training set.

## 4 Experiments

### 4.1 Datasets

We used the public dataset DeepSig RadioML 2018.01A in our experiments, which is a famous and widely used dataset in the modulation recognition. This dataset was generated by O'Shea et al. via GNU radio equipment, which takes into account synthetic analog channel effects and airborne propagation losses, and well reproduces signals in real environments (O'Shea et al., 2018). The dataset contains 24 modulations including OOK, 4ASK, 8ASK, BPSK, QPSK, 8PSK, 16PSK, 32PSK, 16APSK, 32APSK, 64APSK, 128APSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM, AM-SSB-WC,AM-SSB-SC, AM- DSB-WC, AM-DSB-SC, FM, GMSK, and OQPSK. DSB-WC, AM-DSB-SC, FM, GMSK, and OQPSK, covering not only simple modulation, but also high-order digital modulation. Each modulation contains 26 different SNR ranging from –20 dB to +30 dB in 2 dB steps, and each SNR for each modulation contains 4,096 pieces of data, each of which is IQ data, with a sample length of 1,024, and the dataset contains a total of 2,555,904 pieces of data. The signals in Figures 1 and 2 are the 4ASK signals in this dataset. In the experiments, we first classified the DeepSig RadioML 2018.01A dataset into 24 classes according to the modulation mode, and then divided each modulated signal into 26 classes according to the SNR. In accordance with the ratio of 7:1:2, the processed dataset is divided into training set, validation set, and test set, and in order to avoid overfitting, data enhancement was performed on the test set, including both flipping and masking.

**Figure 4** Accuracy of 24-modulation for ResNet50_1D and VGG16_1D (see online version for colours)

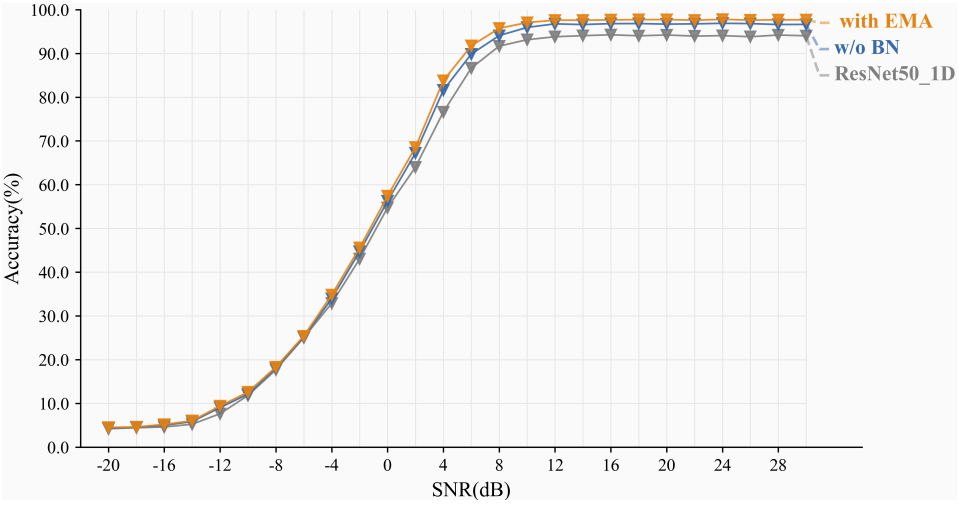



### 4.2 Results

We conducted the experimental on the same NVIDIA GeForce GTX 4090 GPU to avoid errors in randomly dividing the dataset due to platform differences. The training

set, validation set, and test set all include all 24 modulations and 26 SNR in the DeepSig RadioML 2018.01A dataset which can improve the robustness of the network. Firstly, we trained the network on the divided training set; secondly, we adjusted the hyperparameters through the performance of the network on the validation set to improve the network fitting ability; and finally, we tested the trained network on the test set. In order to verify that the model capability is not the main factor to improve the accuracy, we adapted the convolutional and pooling layers in the VGG16 network to 1-dimensional to get VGG16_1D, and conducted contrast experiments with ResNet50_1D and VGG16_1D networks. The training epochs are 50, the batch size is 256 and the learning rate is 0.0001. A cosine simulated annealing method was used to adjust the initial learning rate to 0.0001. To compensate for the regularisation effect provided by the BN, we used the drop path method with the probability set to 0.4. The training set, validation set, test set, and experimental steps used in the two sets of experiments are the same, and only the network structure is different, and the results are shown in Figure 4.

**Figure 5** Accuracy of 24 modulations for three methods (see online version for colours)




In the meantime, we conducted contrast experiments to illustrate the noise suppression effect that can be achieved by removing the BN layers. The first set of experiments uses ResNet50_1D, and the second set of experiments uses the network structure described in Subsection 3.2. To verify that the EMA method can improve the robustness of the network, we setup a third set of experiments to train the network with the EMA method based on the second set of experiments. The steps of the three sets of experiments are the same, and the training parameters were consistent with the first experiment. We performed three tests for each model, and the recognition accuracies in Figure 5 are the average of the three experiments. It is obvious in Figure 5, the performance of the network is somewhat improved by removing the BN layers; the classification accuracy is higher for networks trained using the EMA method. Compared to ResNet50-1D, our proposed method has improved accuracy on all SNR, and the accuracy exceeds 90% at SNR greater than 6 dB. The accuracy of different methods at low SNR is shown in Table 1.

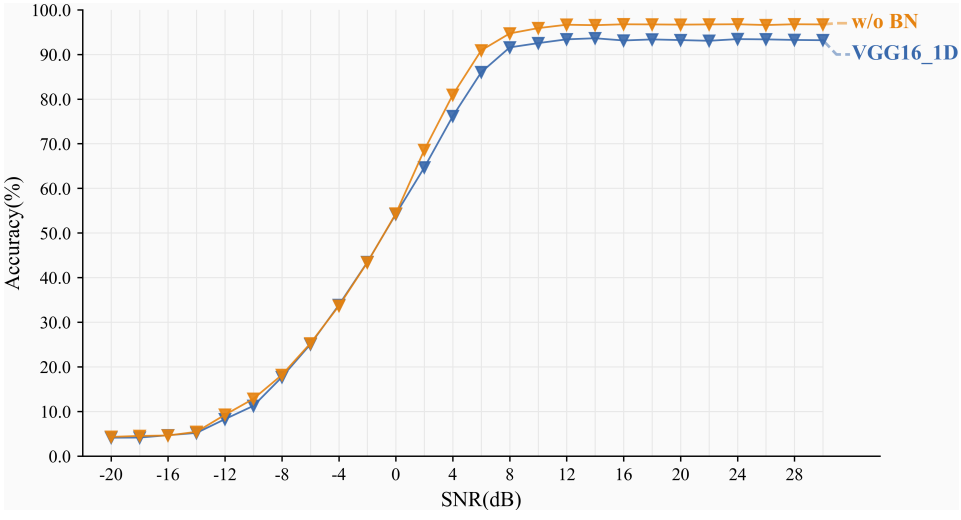
**Table 1** Comparison of accuracy of different methods at low SNR

| *SNR (dB)* | *DNCNet (Du et al., 2022)* | *ResNet50_1D* | *Proposed* |
|---|---|---|---|
| –6 | \ | 25.05% | 25.41% |
| –4 | \ | 32.89% | 34.80% |
| –2 | 40.84% | 43.03% | 45.57% |
| 0 | 56.49% | 54.75% | 57.42% |
| 2 | \ | 64.02% | 68.53% |
| 4 | 83.83% | 76.62% | 83.80% |
| 6 | \ | 85.65% | 91.75% |
| Average | 60.39% | 59.94% | *62.97%* |

**Table 2** Comparison of computational effort between the two sets of experiments

| *Model* | *Total parameters (million)* | *Time for each epoch (NVIDIA GeForce GTX 4090 GPU, minutes)* |
|---|---|---|
| ResNet50_1D | 34.749 | 12.8 |
| Propose | 34.696 | 11.4 |

**Figure 6** Accuracy of 24-modulation for VGG16_1D and w/o BN (see online version for colours)




Through Table 1, it is obvious that the accuracy of the model with the BN layers removed is significantly improved at low SNR, and the average accuracy is improved by 3.03% over ResNet50_1D and 2.56% over the method in Du et al. (2022). As shown in Table 2, removing the BN layers reduces the amount of parameters of the network, shortens the computation time, and reduces the computational burden. Therefore, removing the BN layers can suppress high-frequency noise, improve the model classification accuracy, and reduce the amount of computation. In order to verify that this approach is not a special case on the ResNet structure, we also conducted the

same experiment using the VGG network. We replaced the BN layers in VGG16_1D with Identity layer. Two groups of experiments are also performed, the first experiment adopts the VGG16_1D structure, and the second experiment adopts the above structure. The training set, validation set, test set and experimental steps used are consistent with the ResNet50_1D experiments, and the experimental results are shown in Figure 6.

In Figure 6, the classification effect of the model after removing the BN layers is better than that of VGG16_1D, the accuracy is higher than that of VGG16_1D in the whole SNR, and the classification accuracy is also higher than 90% when the SNR is greater than 6 dB, and the average classification accuracy reaches 62.69%. Therefore, removing the BN layers in the network structure can indeed suppress high-frequency noise, improve the model performance and classification accuracy, and reduce the amount of computation.

## 5 Discussion

The experimental results show that removing the BN layers improves the accuracy and robustness of the radar emitter recognition model. First, BN helps the model to better utilise the high-frequency information, while the noise in the radar signal is mostly in the high-frequency part. Therefore, removing BN can suppress the network's focus on high-frequency information, which achieves the denoising effect. Secondly, removing BN forces the model to learn a more stable internal representation rather than relying on the statistical properties of each batch of data, making the model more robust. Although removing BN may make training more difficult, the model is able to achieve better performance through drop path regularisation, hyperparameter tuning, and EMA training strategies. Finally, removing BN also reduces the number of parameters in the model, which reduces the memory footprint and computational costs.

Through Figures 5 and 6, it can be found that although removing the BN layers improves the overall classification accuracy of the model, the classification accuracy of the model is not high when the SNR is lower than –10 dB. This is because when the SNR is very low, the power of the signal is also very low relative to the noise power. Even if the high-frequency noise is removed, the low-frequency part of the noise is still dominant and the signal is basically annihilated by the noise, so it is difficult to separate it from the noise. Therefore, in the case of a very low SNR, not only is it necessary to remove the high-frequency noise, but also other methods are needed to distinguish the low-frequency part of the signal from the noise in order to improve the classification accuracy.

## 6 Conclusions

In this paper, we proposed to remove the BN layers in the model to suppress the noise, constructed an end-to-end radar emitter denoising and recognition model, and improved the accuracy of the model recognition, especially to enhance the accuracy at low SNR situations. The experimental result showed that our proposed method achieved an average accuracy of 62.97% on the DeepSig RadioML 2018.01A dataset with SOTA results, and the classification accuracy of the model was improved with a lower SNR, while reducing the computation and training time of the model. Through

experiments using VGG16, it was verified that removing the BN layers in the network structure can effectively suppress the noise and improve the model performance. In addition, our study demonstrates the potential of DL in dealing with complex signal processing problems, especially when traditional signal processing methods encounter bottlenecks. The strategy of removing BN and adjusting other hyperparameters shows that the design of DL models can be flexibly adapted to fit the needs of a specific task, which provides new perspectives for future model design. And we will investigate the denoising algorithm in the case of very low SNR in our future work to further improve the classification accuracy.

## References

Cosman, P.C., Gray, R.M. and Olshen, R.A. (1994) 'Evaluating quality of compressed medical images: SNR, subjective rating, and diagnostic accuracy', *Proceedings of the IEEE*, Vol. 82, No. 6, pp.919–932.

Du, M.Y., Zhong, P., Cai, X.H. and Bi, D.P. (2022) 'DNCNet: deep radar signal denoising and recognition', *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 58, No. 4, pp.3549–3562.

He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J. (2016) 'Deep residual learning for image recognition', *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.770–778.

He, N.J., Paoletti, M.E., Haut, J.M., Fang, L.Y., Li, S.T., Plaza, A. and Plaza, J. (2018) 'Feature extraction with multiscale covariance maps for hyperspectral image classification', *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 57, No. 2, pp.755–769.

Huynh-The, T., Hua, C.H., Pham, Q.V. and Kim, D.S. (2020) 'MCNet: an efficient CNN architecture for robust automatic modulation classification', *IEEE Communications Letters*, Vol. 24, No. 4, pp.811–815.

Huynh-The, T., Doan, V.S., Hua, C.H., Pham, Q.V., Nguyen, T.V. and Kim, D.S. (2021) 'Accurate LPI radar waveform recognition with CWD-TFA for deep convolutional network', *IEEE Wireless Communications Letters*, Vol. 10, No. 8, pp.1638–1642.

Ioffe, S. and Szegedy, C. (2015) 'Batch normalization: accelerating deep network training by reducing internal covariate shift', *ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning*, Vol. 37, pp.448–456.

Izmailov, P., Podoprikhin, D., Garipov, T., Vetrov, D. and Wilson, A.G. (2018) 'Averaging weights leads to wider optima and better generalization', *34th Conference on Uncertainty in Artificial Intelligence 2018, UAI 2018*, Vol. 2, pp.876–885.

López-Risueño, G., Grajal, J. and Sanz-Osorio, A. (2005) 'Digital channelized receiver based on time-frequency analysis for signal interception', *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 41, No. 3, pp.879–898.

Liu, J., Lee, J.P., Li, L., Luo, Z.Q. and Wong, K.M. (2005) 'Online clustering algorithms for radar emitter classification', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 8, pp.1185–1196.

Liu, T., Wang, M.J., Feng, H., Luan, S.Y. and Zhang, J.C. (2024) 'Joint estimation of DOA and range for near-field sources in the presence of far-field sources and alpha-stable noise', *IEEE Signal Processing Letters*, Vol. 31, pp406–410.

Liu, Z.M. (2020) 'Recognition of multifunction radars via hierarchically mining and exploiting pulse group patterns', *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 56, No. 6, pp.4659–4672.

O'Shea, T.J., Roy, T. and Clancy, T.C. (2018) 'Over-the-air deep learning based radio signal classification', *IEEE Journal of Selected Topics in Signal Processing*, Vol. 12, No. 1, pp.168–179.

Préaux, Y. and Boudraa, A.O. (2020) 'Statistical behavior of teager-kaiser energy operator in presence of white gaussian noise', *IEEE Signal Processing Letters*, Vol. 27, pp.635–639.

Santurkar, S., Tsipras, D., Ilyas, A. and Madry, A. (2018) 'How does batch normalization help optimization?', *NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pp.2488–2498.

Shahamiri, S.R. (2021) 'Speech vision: an end-to-end deep learning-based dysarthric automatic speech recognition system', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 29, pp.852–861.

Sun, L.T., Huang, Z.T., Wang, X, Wang, F.H. and Li, B.G. (2020) 'Overview of radio frequency fingerprint extraction in specific emitter identification', *Journal of Radars*, Vol. 9, No. 6, pp.1014–1031.

Wang, W.D. and Gan, L. (2022) 'Radio frequency fingerprinting improved by statistical noise reduction', *IEEE Transactions on Cognitive Communications and Networking*, Vol. 8, No. 3, pp.1444–1452.

Wang, Y., Liu, M., Yang, J. and Gui, G. (2019) 'Data-driven deep learning for automatic modulation recognition in cognitive radios', *IEEE Transactions on Vehicular Technology*, Vol. 68, No. 4, pp.4074–4077.

Wang, H.H., Wu, X.D., Huang, Z.Y. and Xing, E.P. (2020) 'High-frequency component helps explain the generalization of convolutional neural networks', *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.8681–8691.

Wen, F.Q., Xiong, X.D., Su, J. and Zhang, Z.J. (2017) 'Angle estimation for bistatic MIMO radar in the presence of spatial colored noise', *Signal Processing*, Vol. 134, pp.261–267.

Wu, Z.H., Liu, C., Wen, J.L., Xu, Y., Yang, J. and Li, X.L. (2022) 'Selecting high-quality proposals for weakly supervised object detection with bottom-up aggregated attention and phase-aware loss', *IEEE Transactions on Image Processing*, Vol. 32, pp.682–693.

Xiao, Z.L. and Yan, Z.Y. (2021) 'Radar emitter identification based on novel time-frequency spectrum and convolutional neural network', *IEEE Communications Letters*, Vol. 25, No. 8, pp.2634–2638.

Yang, S.Y., Peng, T.Q., Liu, H.L., Yang, C., Feng, Z.X. and Wang, M. (2023) 'Radar emitter identification with multi-view adaptive fusion network (MAFN)', *Remote Sensing*, Vol. 15, No. 7, p.1762.

Yu, H.H., Yan, X.P., Liu, S.K., Li, P. and Hao, X.H. (2022) 'Radar emitter multi-label recognition based on residual network', *Defence Technology*, Vol. 18, No. 3, pp.410–417.

Zhang, D., Lu, Y.Y., Li, Y.D., Ding, W.R. and Zhang, B.C. (2023) 'High-order convolutional attention networks for automatic modulation classification in communication', *IEEE Transactions on Wireless Communications*, Vol. 22, No. 7, pp.4600–4610.

Zhang, Z.F., Wang, C., Gan, C.Q., Sun, S.H. and Wang, M.J. (2019) 'Automatic modulation classification using convolutional neural network with features fusion of SPWVD and BJD', *IEEE Transactions on Signal and Information Processing over Networks*, Vol. 10, No. 8, pp.469–478.

Zhao, S.Q., Wang, W.H., Zeng, D.G., Chen, X.N., Zhang, Z.Y., Xu, F.Y., Mao, X.Y. and Liu, X.G. (2021) 'A novel aggregated multipath extreme gradient boosting approach for radar emitter classification', *IEEE Transactions on Industrial Electronics*, Vol. 69, No. 1, pp.703–712.

Zhu, Q.K., Du, B. and Yan, P.K. (2019) 'Boundary-weighted domain adaptive neural network for prostate MR image segmentation', *IEEE Transactions on Medical Imaging*, Vol. 39, No. 3, pp.753–763.