Article
Low SNR Multi-Emitter Signal Sorting and Recognition Method Based on Low-Order Cyclic Statistics CWD Time-Frequency Images and the YOLOv5 Deep Learning Model
Dingkun Huang, Xiaopeng Yan *, Xinhong Hao, Jian Dai and Xinwei Wang

Technology on Electromechanical Dynamic Control Laboratory, School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China
* Correspondence: yanxiaopeng@bit.edu.cn

Abstract: It is difficult for traditional signal-recognition methods to effectively classify and identify multiple emitter signals in a low SNR environment. This paper proposes a multi-emitter signal-feature-sorting and recognition method based on low-order cyclic statistics CWD time-frequency images and the YOLOv5 deep network model, which can quickly dissociate, label, and sort the multi-emitter signal features in the time-frequency domain under a low SNR environment. First, the denoised signal is extracted based on the low-order cyclic statistics of the typical modulation types of radiation source signals. Second, the time-frequency graph of multisource signals was obtained through CWD time-frequency analysis. The cyclic frequency was controlled to balance the noise suppression effect and operation time to achieve noise suppression of multisource signals at a low SNR. Finally, the YOLOv5 deep network model is used as a classifier to sort and identify the received signals from multiple radiation sources. The method proposed in this paper has high real-time performance. It can identify the radiation source signals of different modulation types with high accuracy under the condition of a low SNR.

Keywords: cyclic stationary analysis; CWD time-frequency analysis; noise suppression; YOLOv5 model; radiation source signal sorting and identification

1. Introduction

A radiation source signal has the characteristics of variable parameters, various modulation forms, and comprehensive space coverage. In order to expand the function and anti-interference ability of a radiation source signal, the modulation mode is also developing. The modulation mode has gradually changed from a continuous wave system to multi-waveform frequency modulation, pseudo-random-code phase modulation, frequency-agile, pulse Doppler, and other complex modulation modes. In the actual scene, the measured signal is often mixed with other signals and electromagnetic noise, which will reduce the recognition and analysis ability of the receiver. This phenomenon is more evident in an electromagnetic environment with low SNR. Because the signals of various radiation sources are widely present, the communication equipment of manufactured equipment is likely to be interfered with by them in extreme environments. Therefore, successfully identifying the modulation type of the mixed radiation source signal is essential to ensure communication-equipment stability. After identifying the modulation type of the radiation source signal, the equipment can estimate its detailed parameters and avoid interference. The accurate identification of modulation types of mixed emitter signals is an important signal-processing-system index and one of the core technical problems of blind signal adaptive processing.

The difficulties of emitter signal sorting and identification under a low SNR mainly include the following three points:
(1) The noise-suppression effect of a signal and the contradiction of real-time processing
In an electromagnetic environment with low SNR, the receiver must accumulate signal energy for a long time to achieve a good noise-suppression effect. This method increases the computation amount of the signal-analysis algorithm and reduces the real-time performance. The large working bandwidth disperses the FM signal’s energy, further reducing the identification ability, and the effect of the ordinary noise-suppression method is poor.

(2) There are many factors affecting the recognition accuracy of artificial neural networks. Under the same preprocessing effect, the classifier can only obtain the recognition result of a single category and cannot obtain the location information. This is not conducive to the subsequent parameter-estimation work. It is necessary to use large-scale artificial neural network architecture to obtain high recognition accuracy and a low missed detection rate, which extends the training time.

(3) It is difficult to accurately identify the mixed signal combination mode in the coarse classification work.

The sub-signal features in the mixed radiation source are easily confused, which will mislead the classifier and produce incorrect classification results. An inefficient modulation feature-recognition method will significantly reduce the recognition accuracy of the classifier in the working scene.

To solve the problem of radiation source-signal sorting and identification under low SNR, [1] calculates the second-order cyclic cumulant and fuses two detectors to increase the types of signal recognition and reduce the computational burden. However, this method has a low detection probability for signals below $-10$ dB. The authors of [2] calculate the first-order cyclic mean of a VHF signal, identify the modulation classification mode, and obtain a high recognition rate of the blind signal modulation mode, but they identify fewer signal types. In [3], second-order cyclic autocorrelation function and a convolutional neural network are used to compress the signal dimension and reduce the computation amount, and the recognition accuracy is high. The authors of [4] used a multilevel LPI-NET classifier combined with CWD time-frequency analysis. The scheme can recognize 13 radar waveforms, and the recognition accuracy of the 0 dB signal-to-noise ratio is more than 98%. The authors of [5] study generalized Hankel matrix factorization and propose a new cyclic mean kernel function, which can improve the accuracy of parameter estimation in a low SNR electromagnetic environment. The method proposed in [6] accumulates signal energy and suppresses noise, and uses a deep confidence network to automatically sort and identify radiation source signals. The accuracy of this method is 90% at $-5$ dB. The method in [7] integrates multidimensional information of CWD, STFT, and cyclic spectrum, adopts three kinds of neural networks and flying fish swarm intelligence algorithms to identify the modulation types of signals, and the recognition rate reaches 84.7% at $-4$ dB. The authors of [8] studied the autocorrelation function of the signal and used CNN-DNN to identify the two-dimensional cross-section features, with a high recognition accuracy at $-5$ dB. The method proposed in [9] uses the DCN-BILSTM bidirectional long short-term memory network to recognize the amplitude and phase information of signals, and the recognition rate is higher when the SNR is 4 dB. The authors of [10] propose RRSARNet, a new artificial neural network based on meta-transfer learning, which can recognize STFT images and adaptively recognize signals. In [11], to solve the problem of LPI radar waveform recognition at low SNR, a triple convolutional neural network was designed to recognize PWVD time-frequency images, and the recognition accuracy reached 94% at $-10$ dB. In [12], the CMWT time-frequency graph was input into the ResNet (SEP-ResNet) network, which could classify the enhanced T-F images and complete the high-precision intelligent recognition of radar-modulated signals in a low SNR environment. When the SNR is $-10$ dB, the recognition accuracy is 93.44%. The authors of [13] combined time-frequency analysis with a deep neural network and analyzed the distribution law of electromagnetic noise in the time-frequency diagram, designed a filter to suppress the noise, and obtained a good classification effect. The method in [14] combines CWD and CNN neural networks, uses the binarization operation of preprocessed time-frequency images to
suppress noise and highlight signal features, and the signal-recognition accuracy exceeds 90% at −4 dB, and the recognition accuracy of some signals exceeds 98%. The authors of [15] adjust the hyperparameters and filters of different signals in the CNN network and control the calculation amount through sample averaging, with the detection accuracy exceeding 80% at −15 dB. The method proposed in [16] combines Transformer and YOLOv5 for UAV target detection and obtains a good AP value. The method proposed in [17] adopts deep convolutional network and K-means clustering, and a YOLOv5 classifier to achieve high real-time small-target-object detection. In [18], CAA-Yolo was used to identify ship targets on the sea surface, highlighting the recognition of small targets and increasing the mAP value by 3.4%. The authors of [19] combined trunk learning and a Yolo-Tiny network, transplanted the model to GPU edge devices for use, and obtained a high recognition rate. On the basis of Yolo-X, ref. [20] combines the adaptive activation function Meta-ACon with the SPP module of Backbone to improve the feature extraction ability. The authors of [21] introduce a real-time multi-class target recognition system based on millimeter wave (mmWave) imaging radar for advanced driver assistance systems. Objects in various scenes can be detected with little change in the training dataset, thus reducing the training cost. In [22], radar information is processed by mapping transform neural networks to obtain mask mapping, so that radar information and visual information are on the same scale. This paper proposes a multi-data source deep learning Object Detection network (MS-Yolo) based on millimeter wave radar and vision fusion. The authors of [23−27] used different deep learning models and preprocessing algorithms to classify and recognize radar signals, ships, and intracellular molecules. In [28−30], generative adversarial network and adaptive noise suppression are used to sort and identify modulation features for radar signals, and good results are obtained.

As shown in Figure 1, the technical route of spectrum analysis, time-frequency analysis, and other preprocessing methods combined with an artificial neural network classifier has become the primary solution to the problem of emitter-signal sorting and identification under low SNR. However, the real-time performance of the neural network in identifying long sequence signals is poor. Through the above research, this paper finds that the main solutions have high signal classification accuracy from 0 dB to −10 dB, while sorting and identification accuracy below −10 dB is significantly reduced. Spectral analysis, time-frequency analysis, and other methods have a poor effect on signal processing, which brings difficulties to sorting artificial neural networks. In general, real-time performance is negatively correlated with the effect of preprocessing.

![Figure 1. Roadmap of mainstream radiation source signal-sorting and identification technology under a low SNR.](image-url)

To solve the difficult problem of radiation source-signal sorting and identification under a low SNR, this paper performs the following research work and innovation, as illustrated in Figure 2.
Figure 2. The technical process of this paper.

- Low-order cyclic statistics and CWD time-frequency analysis are used as the preprocessing methods for received signals. Using the granularity of the cycle frequency to adjust the resolution of the cycle mean, combined with the IFFT algorithm, can suppress the noise, reduce the amount of computation, and improve the real-time performance of the preprocessing algorithm.
- To mark and identify the modulation features of multiple signals on the time-frequency graph, we adopted the YOLOv5 deep neural network framework as the classifier. After training, the single recognition time is short, and multiple sub-signal features can be recognized simultaneously. The output result of the classifier is the basic premise of parameter estimation.
- To reduce the missed detection rate and identify the combination pattern of mixed signals, the labeled and unlabeled datasets are fused to train the classifier. Using different kinds and quantities of signals in the test can show the recognition defects of the classifier at different levels and support the subsequent optimization work.

2. Materials and Methods

2.1. Multisource Signal Preprocessing Method Based on Cyclic Mean and CWD Time-Frequency Analysis

The preprocessing method proposed in this paper takes the cyclic mean calculation as the denoising method, carries out domain transformation through the IFFT, and finally generates a time-frequency image, including the multi-emitter signal, through CWD time-frequency analysis. The length and domain of the sequence will change during each step. The input to the preprocessing system is a long one-dimensional signal sequence discrete signal with noise in the time domain. The output is a two-dimensional CWD time-frequency image after noise suppression in the short sequence.

Since the specific order statistical characteristic parameters of cyclically stationary signals change periodically with time and the characteristics of different classes of modulated signals are distinguished, low-order statistics require less computation and have good noise-suppression performance. The cyclic mean is applied as the first step of the preprocessing algorithm. Since most received signals have cyclic stationary characteristics, the cyclic mean values of AM and FM modulation types are theoretically derived in this paper.
2.1.1. Cyclic Mean Analysis of AM Signals

As a common modulation signal type, AM signals are widely used in the field of digital communication. The mathematical model is:

\[ u_{AM}(t) = U_{cm}(1 + m_a \cos(f_m t)) \cos(2\pi f_c t) + n(t) \]  

(1)

In the above equation, \( U_{cm} \) is the carrier signal amplitude, \( m_a = K(U_{Qm}/U_{cm}) \) is the modulation rate, \( K \) is a constant, \( U_{Qm} \) is the modulated signal amplitude, \( f_m \) is the modulated signal frequency, \( f_c \) is the carrier frequency, and \( n(t) \) is the Gaussian white noise. The cyclic mean of the AM signal when the cyclic frequency is \( \alpha \) is:

\[ M_{uA}^\alpha(t) = < U_{cm}(1 + m_a \cos(f_m t)) \cos(2\pi f_c t)e^{-j2\pi \alpha t} > + < n(t)e^{-j2\pi \alpha t} > \]  

(2)

Since the mean of the Gaussian white noise is 0, the value of \( < n(t)e^{-j2\pi \alpha t} > \) is 0, from which we can prove that the first-order cyclic mean has theoretically good denoising properties. Expanding (1) and (2) to obtain the sum of the two parts, part A is:

\[ M_{uA}^\alpha(t) = < U_{cm}\cos(2\pi f_c t)[\cos(2\pi \alpha t) - j \sin(2\pi \alpha t)] > \]  

(3)

Expanded, it is:

\[ M_{uA}^\alpha(t) = < \frac{1}{2}U_{cm}\{\cos(2\pi (f_c + \alpha) t)\} + \{\cos(2\pi (f_c - \alpha) t)\} > \]  

(4)

Part B is:

\[ M_{uB}^\alpha(t) = < U_{cm}m_a \cos(2\pi f_c t) \cos(f_m t)\cos(2\pi \alpha t) - j \sin(2\pi \alpha t) > \]  

(5)

Expanded, it is:

\[ M_{uB}^\alpha(t) = < \frac{1}{2}m_a U_{cm}\cos(2\pi f_c t + f_m t) + \cos(2\pi f_c t - f_m t)\cos(2\pi \alpha t) - j \sin(2\pi \alpha t) > \]  

(6)

\[ R = 0.5 \ast m_a U_{cm}, \]  

and F:

\[ M_{uBc}^\alpha = < R\{\frac{1}{2}\{\cos(2\pi t(f_c + \frac{1}{2\pi}f_m + \alpha)) + \cos(2\pi t(f_c + \frac{1}{2\pi}f_m - \alpha))\} > \]  

\[ M_{uBD}^\alpha = < -j R\{\frac{1}{2}\{\sin(2\pi t(f_c + \frac{1}{2\pi}f_m + \alpha)) - \sin(2\pi t(f_c + \frac{1}{2\pi}f_m - \alpha))\} > \]  

(7)

\[ M_{uBE}^\alpha = < R\{\frac{1}{2}\{\cos(2\pi t(f_c - \frac{1}{2\pi}f_m + \alpha)) + \cos(2\pi t(f_c - \frac{1}{2\pi}f_m - \alpha))\} > \]  

\[ M_{uBF}^\alpha = < -j R\{\frac{1}{2}\{\sin(2\pi t(f_c - \frac{1}{2\pi}f_m + \alpha)) - \sin(2\pi t(f_c - \frac{1}{2\pi}f_m - \alpha))\} > \]

From Equations (3)–(7), the cyclic mean of the AM signal is divided into real and imaginary parts, and a parameter with cyclic frequency \( \alpha \) is added to the Fourier frequency. Because the frequency terms of the two parts are the same, this section only uses the real part mean as an example. The real part of the cyclic mean is composed of three spectral lines with different frequencies, where the cyclic frequency \( \alpha \) represents the frequency under the current transformation. The image of the first-order cyclic mean is the result of time averaging after the Fourier transform of the signal is shifted \( \alpha \) to the left as a whole.

Figure 3 shows the cyclic mean of the AM signal under different SNRs. In the range from 0 to −10 dB, the cyclic mean of the signal has a substantial noise-suppression property, and at −20 dB, the noise-suppression performance is significantly reduced.
From Equations (3)–(7), the cyclic mean of the AM signal is divided into real and imaginary parts:

\[ M_{ufm}(t) = \langle u_{fm}(t)e^{-\Omega t} \rangle = \langle U_m(\cos(2\pi f_c t + mf \sin \Omega t) + n(t) \rangle \]

Expanding (11) and (12):

\[ M_{ufmA}(t) = \langle U_m(\cos(2\pi f_c t + mf \sin \Omega t)) + \cos(\Omega t) \rangle \]

\[ M_{ufmB}(t) = \langle -\frac{i}{2} U_m(\sin(2\pi f_c t + mf \sin \Omega t)) \rangle \]

According to Equations (14) and (15), the cyclic mean of the FM signal can also be divided into two parts: the real part and the imaginary part. Similar to the result after the Fourier transform of the FM signal, a component of cyclic frequency is added to the spectral line.

As shown in Figure 4, when the cyclic frequency \( \alpha \) is close to the harmonic frequency of the signal, the cyclic mean will have a peak value, and when it leaves the harmonic frequency, the peak value will rapidly decrease to zero. As the calculation method of
the cyclic mean is the generalized FFT operation, due to its superior denoising performance and high operation speed, the calculation of the cyclic mean is the main means of denoising and preprocessing the received signals from multiple sources. Because the traditional method of lengthy signal spectrum analysis and time-frequency analysis directly transforming domain operations tends to increase the amount of calculation and consume too much time, choosing an appropriate cyclic frequency is equal to the adjusted cyclic average resolution. With its sequence control points, the average cyclic images can carry long-time domain-sequence information with short sequences and reduce the time for subsequent steps.

![FM cyclic mean diagram -10dB](image1)
![FM cyclic mean diagram -20dB](image2)

**Figure 4.** The cyclic mean of the FM signal under different SNRs.

### 2.2. CWD Time-Frequency Analysis Image Generation

The time-domain sequence is converted into the cyclic domain by calculating the cyclic mean of the multi-emitter signal, and the cyclic frequency controls the sequence length. Since the input to the CWD algorithm is a time-domain sequence, after the cyclic mean suppresses the noise, the IFFT method is used to transform the short sequence in the cyclic domain into a time-domain sequence of the same length. Because the value of the cyclic frequency is small, the frequency offset caused by the IFFT can be ignored. The above steps can maintain the noise-suppression effect, complete the sequence length conversion, and reduce the processing time of the CWD algorithm.

The kernel function determines the type of time-frequency analysis. This paper uses the unified expression of Cohen class time-frequency distribution to derive the theory of CWD time-frequency analysis. Choosing the proper kernel function can reduce the cross-term caused by the operation, if:

\[
\varphi(\tau, \xi) = e^{-\xi^2 \tau^2 / \sigma}
\]  

(16)

Among these, \(\sigma (\sigma > 0)\) is known as the scaling factor, which substitutes (1–11) into the unified expression of the Cohen class time-frequency distribution. If the continuous signal is \(x(t)\), then the CWD is expressed as:

\[
CWD_x(t, \omega) = \int_{-\infty}^{\infty} e^{-j\omega t} \left[ \int_{-\infty}^{\infty} \frac{\sigma}{4\pi T^2} G(\mu, \tau) z(\mu + \frac{\tau}{2}) z^* (\mu - \frac{\tau}{2}) d\mu \right] d\tau
\]  

(17)

where

\[
G(\mu, \tau) = e^{-\sigma (\mu - t)^2 / 4\tau^2}
\]  

(18)

At time \(t\), the angular frequency is \(\omega\) and \(\sigma\) is the scaling factor. The size of the scaling factor is proportional to the amplitude of the crossover term. Therefore, it is crucial for
the choice of the CWD scaling factor; a small scale factor will lead to poor time-frequency aggregation of the signal, and a large scaling factor will lead to a serious cross-term.

In practice, the CWD calculation needs to be discretized by:

$$CWD_s(s, \omega) = 2 \sum_{\tau=\infty}^{\infty} W_N(\tau)e^{-j2\omega \tau} \sum_{\mu=\infty}^{\infty} W_M(\mu) \sqrt{\frac{\sigma}{4\pi \tau^2}} e^{-\epsilon(\mu-\tau)^2/(4\tau^2)} z(\mu + \tau + s) * (\mu - \tau + s)$$  \hspace{1cm} (19)

In Equations (1)–(19), the window functions $W_N(\tau)$ and $W_M(\tau)$ are weighted. The $\tau$ values of these two window functions are between $-n/2 \leq \tau \leq N/2$ and $-m/2 \leq \tau \leq M/2$, respectively.

CWD, as a widespread use of the time-frequency analysis method, has an excellent cross-term inhibiting effect, but it is rather time-consuming for the two-dimensional sequence of the process. To improve the speed of pretreatment, one-dimensional time series are directly processed. For a one-dimensional sequence on noise-suppression generated CWD time-frequency images, the time-frequency transform after image processing minimizes time consumption, and there is no need to operate the two-dimensional image matrix frequently. When CWD time-frequency images are generated, they can be directly fed into the classifier, which greatly improves the processing speed.

2.3. Comparison of the Results of the Preprocessing Algorithm

Multi-emitter signal superposition can be divided into two cases. The first case is that the time-frequency domain of each sub-signal is far apart, and a method of searching for the peak value in the frequency domain can effectively distinguish the modulation type of the sub-signal. The second case is that each sub-signal overlaps in the time-frequency domain. This study focuses on the second case.

Figure 5 shows the cyclic mean of AM and FM signals mixed at $-20$ dB. After noise suppression using the cyclic mean, the characteristics of each sub-signal can be displayed. The cyclic mean value of the received mixed signals is calculated, and the number of sampling points is added to the time-domain sequence to achieve a better noise-suppression effect. The IFFT operation is then performed on the signal to obtain the time-domain signal under short sequence points. At this time, the algorithm has completed the transformation of long and short time series.

![Figure 5. The cyclic mean of the AM and FM multiple signals.](image-url)

Figures 6 and 7 show the CWD time-frequency images after noise suppression at $-20$ dB and $-15$ dB, respectively. The preprocessing method proposed in this paper has a good noise-suppression effect on different combinations of sub-signals. However, the CWD transformation diagram at $-20$ dB still retains certain noise points, while the signal time-frequency diagram at $-15$ dB has fewer noise points. It can effectively recover the emitter signal from the low SNR to facilitate the subsequent classifier processing.
Sampling points is added to the time-domain sequence to achieve a better noise-suppression effect. The IFFT operation is then performed on the signal to obtain the time-domain signal under short sequence points. At this time, the algorithm has completed the transformation of long and short time series.

Figure 5. The cyclic mean of the AM and FM multiple signals.

Figures 6 and 7 show the CWD time-frequency images after noise suppression at −20 dB and −15 dB, respectively. The preprocessing method proposed in this paper has a good noise-suppression effect on different combinations of sub-signals. However, the CWD transformation diagram at −20 dB still retains certain noise points, while the signal time-frequency diagram at −15 dB has fewer noise points. It can effectively recover the emitter signal from the low SNR to facilitate the subsequent classifier processing.

(a) unsuppressed noise (−20 dB), (b) suppressed noise.

Figure 6. (a) unsuppressed noise (−20 dB), (b) suppressed noise.

The cyclic mean is set to:

\[ \alpha = \frac{2\pi k}{K} \]  

The value of the cyclic frequency can be flexibly adjusted according to different needs. K is an arbitrary integer, the length of the cyclic sequence is proportional to the value of k, and the resolution is inversely proportional.

Figure 8 shows a diagram of the sequence length and domain transformation of the preprocessing method. Time series of the same length can be transformed into a one-dimensional sequence of cyclic domains of different lengths by calculating the mean value of different cyclic frequencies. The cyclic sequences are converted into time-domain sequences of the same length by the IFFT algorithm. At this time, the signal noise of multiple radiation sources received by the system has been suppressed. Since the number of size points of the CWD time-frequency analysis graph is the square of the length of the IFFT sequence, a short IFFT sequence length can significantly reduce the computation time of the CWD time-frequency image.

After preprocessing, the time-frequency images of the multi-radiation source signals are obtained, which can be directly fed into the classifier for feature identification. The pseudocode of the preprocessing algorithm is presented in Appendix A.
The value of the cyclic frequency can be flexibly adjusted according to different needs. \( k \) is an arbitrary integer, the length of the cyclic sequence is proportional to the value of \( k \), and the resolution is inversely proportional.

Figure 8 shows a diagram of the sequence length and domain transformation of the preprocessing method. Time series of the same length can be transformed into a one-dimensional sequence of cyclic domains of different lengths by calculating the mean value of different cyclic frequencies. The cyclic sequences are converted into time-domain sequences of the same length by the IFFT algorithm. At this time, the signal noise of multiple radiation sources received by the system has been suppressed. Since the number of size points of the CWD time-frequency analysis graph is the square of the length of the IFFT sequence, a short IFFT sequence length can significantly reduce the computation time of the CWD time-frequency image.

Figure 8. Preprocessing method steps and sequence length comparison diagram.

2.4. Computational Complexity of Preprocessing Algorithm

This paper’s preprocessing computational complexity analysis consists of three parts: cyclic mean, IFFT calculation, and CWD time-frequency analysis. The butterfly calculation method is adopted for the cyclic mean and IFFT, and four multiplications and two additions are defined as one operation. For the cyclic mean of \( N \) points, there are \( N + \frac{N^2}{2} \) operations. Similarly, for an IFFT of \( k \) points, there are \( k + \frac{k^2}{2} \) operations. The calculation amount of CWD time-frequency analysis is:

\[
O(CWD) = \sigma \frac{k^2}{4}
\]  

(21)

In Formula (21), \( \sigma \) is a constant, \( t \) is the number of time points, and the computational complexity of the CWD algorithm is mainly related to \( t \). If \( \sigma = 0.5 \), the CWD calculation of converting the cyclic mean series of 2000 points into an IFFT time series of 1000 points requires 657,750 calculations in the preprocessing process. If a 50 MHz embedded processor is used, the above steps take about 13 ms.

The computing speed and complexity in the natural environment depend on the hardware platform, the specific implementation of the algorithm, and many other conditions, so this paper calculates the running time on a microprocessor based on this theoretical estimation.

3. Multi-Emitter Signal Classification Model Based on a YOLO Deep Network

As a target-recognition algorithm in the field of deep learning, You Only Look Once (YOLO) has a high accuracy rate and short training time for multicategory object recognition. It is widely used in fault recognition, dangerous goods recognition, intelligent driving, and other fields that are characterized by clear returned images, obvious target features, and distinguishable features. Benefiting from the rapid identification of the YOLO algorithm for good multiobjective support, applying a small sample dataset to real-time signal sorting under low-SNR recognition technology has a great advantage, but often the subduction of the electromagnetic signal in the noise and the characteristic of time-frequency domain
overlapping confusion cause the YOLO deep network to have low sorting result accuracy. However, after the multi-emitter signal preprocessing steps proposed in this paper, the problem of obtaining high-quality images is solved, and a high-resolution signal time-frequency map containing signal-feature information suitable for the YOLO algorithm is obtained, which is convenient for sorting and recognition operations.

The YOLO algorithm is based on a convolutional neural network architecture, which can simultaneously mark the features and locations of multiple targets. According to the architecture size of the network, it can be divided into five sizes: N, S, M, L, and X. Each size is composed of four parts: input end, backbone, neck and head (prediction). In general, the YOLO algorithm consists of the following steps:

1. An image is segmented into an \( S \times S \) image matrix grid.
2. If the center point of the predicted target falls within a grid, then the grid is bound to the current target.
3. Within each grid, \( K \) borders containing the center point of the target are predicted.
4. The target confidence of the border within each grid and the probability of each border area on multiple categories are calculated.
5. The target window with low confidence is removed according to the threshold set.
6. Non-maximum rejection (NMS) is used to remove redundant windows.

In the input-processing stage, YOLOv5 uses mosaic data enhancement, adaptive anchoring, image scaling, and other methods to preprocess images and splice multiple images in different ways to increase the number of features and improve the recognition accuracy of datasets with small sample sizes.

The backbone uses focus downsampling, a CSP structure, and an SPP pooling layer to extract image features. The neck adopts an FPN+PAN feature pyramid structure as the feature-aggregation layer of the entire network architecture, which can obtain the multiscale information of images and better solve the multiscale recognition problem of image features after the input to the head. The head uses three loss functions to calculate the feature classification, feature location, and confidence loss, and improves the accuracy of network prediction through NMS. The output of the algorithm includes a vector of prediction box category, confidence, and coordinate position.

The CBL structure of YOLOv5 uses a convolution (Conv) layer and a batch normalization (BN) layer and is passed into a SiLU activation function layer. The SiLU activation function is defined as follows:

\[
\text{SiLU}(x) = x \cdot \sigma(x)
\]

\[
\sigma(z) = \frac{1}{1 + \exp(-z)}
\]

The focus module divides the input images into four layers, splices them into 12-dimensional feature images, extracts \( 3 \times 3 \) feature information, and generates 32-dimensional feature images. Focus downsampling does not lose image information, significantly reduces the amount of computation, and improves the overall training speed.

The bottleneck consists of residual modules. It controls the number of channels using two Conv modules, extracts feature information, and controls residual connections using shortcuts.

The Conv module in SPP reduces the input number of the feature map, subsamples the maximum pooling of three different convolution kernels, and splices the results with the input feature map according to the channel. YOLOv5 uses convolution kernels with consistent step sizes to realize SPP, obtaining the same feature-map size but with different regional sensitivities, and finally realizing feature fusion.

Through the above structure, three steps of image scaling, convolutional neural network and non-maximum value suppression of object detection are realized.
The loss function of YOLOv5 includes three parts: localization loss, classification loss, and confidence loss, where the total loss function is the sum of the three-part loss function. The GIOU function for the confidence loss is calculated as:

\[
GIOU = IOU - \frac{C - (A \cup B)}{C}
\]

where \( A \) is the prediction box of the model, \( B \) is the real box, and \( C \) is the minimum box containing \( A \) and \( B \). For localization loss and classification loss, the YOLOv5 model adopts the cross-entropy loss function:

\[
C = -\frac{1}{n}\sum_x [y \ln a + (1 - y) \ln(1 - a)]
\]

where \( x \) is the sample, \( y \) is the label, \( a \) is the predicted value, and \( n \) is the total number of samples. The pseudocode for the Yolo algorithm has been presented in Appendix A.

4. Research on Multi-Emitter Signal Sorting and Recognition Based on the Preprocessing Method and the YOLOv5 Model

4.1. Experimental Environment and Dataset Setup

The experimental platform environment built in this study is shown in Table 1. An NVIDIA Tesla V100 graphics card and CUDA11.1.1 hardware acceleration greatly improve the training speed.

| Category                  | Configuration                      |
|---------------------------|------------------------------------|
| CPU                       | Intel(R) Core(TM)i7-10700K @3.80 GHz |
| GPU                       | NVIDIA Tesla V100 32 GB             |
| Operating System          | Linux                              |
| Deep Learning Framework   | Pytorch 1.8.1                       |
| Programming Language      | Python 3.8                         |
| Dependent Package         | CUDA 11.1                          |

In this study, seven modulation categories and mixed categories of radiation-source signals are used. The single modulation category includes triangle wave linear frequency modulation (FM_TRI), sawtooth wave frequency modulation (FM_SAW), sinusoidal frequency modulation (SFM), AM, Fre_Agility signal (Fre_Agility), pulse signal, and biphase coding signal. Since this study studies the modulation-type recognition of radiation-source signals with similar frequencies, the recognition rate of single signal features will be greatly affected when the frequencies are similar or partially aliased. To improve the recognition rate of mixed signals with highly similar modulation features, as shown in Figure 9, frequency overlapping features of 15 classes of mixed signals with similar frequencies, such as FM_TRI, FM_SAW, and SFM, are added in the process of dataset annotation, for a total of 22 classes.
Figure 9. Mixed feature markers.

In Table 2, the parameters of the signal of multiple radiation sources are listed. The carrier frequency is set to 800 MHz–1.5 GHz, and the modulation center frequency is 100 kHz. To simulate randomness, signals are randomly generated within 10 kHz of the center frequency of each radiation source signal. The characteristic size of each signal in the time-frequency image is uncertain; for example, the triangle wave linear frequency modulation signal and sinusoidal frequency modulation signal have obvious characteristics in the case of a large bandwidth and are close to the pulse signal, binomial coding signal, and variable speed signal in the case of small bandwidth and modulation slope. The input to YOLOv5 uses Mosaic to optimize the effect of small target-feature detection. In this study, 3500 mixed-signal CWD image datasets are used for classification training tasks. In the process of data enhancement, the datasets can be further expanded by random cutting and stitching of the program to increase the number of features at the multiscale level.

Table 2. Experiment parameters (Single Classes).

| Category          | Carrier Fre | Mod Fre | Sample Fre | SNR   | Special PAR            |
|-------------------|-------------|---------|------------|-------|------------------------|
| FM_TRI            | 1 GHz       | 100 kHz | 4 GHz      | −15~0 Step | None                  |
| FM_SAW            | 1 GHz       | 100 kHz | 4 GHz      | −15~0 Step | None                  |
| SFM               | 1 GHz       | 100 kHz | 4 GHz      | −15~0 Step | None                  |
| PULSE             | 2 GHz       | 100 kHz | 4 GHz      | −15~0 Step | Duty Ratio 50%        |
| AM                | 800 M–1.5 GHz | 100 kHz | 4 GHz      | −15~0 Step | 50% Mod Rate          |
| Fre_Agility Coding | 800 M–1.5 MHz | 100 kHz | 4 GHz      | −15~0 Step | 5 Frequencies        |
| Biphase Coding    | 1 GHz       | 100 kHz | 4 GHz      | −15~0 Step | Duty Ratio 50%        |

Figure 10 shows the YOLOv5s data-enhancement result. Four time-frequency images are cropped, scaled, and then randomly arranged into one image to improve the generalization ability of the model. To better distinguish the states of signals in the frequency-domain overlap, this paper considers the law of signal overlap in the labeling of datasets and uses different label sizes to cover multisignal features under different parameters, so that the classifier can better learn the types of signals under the overlap.

Figure 10. Data enhancement of the CWD time-frequency graph training process by input.
As shown in Figure 11, the dataset contains single category labels and mixed category labels. Since the training dataset requires a large amount of manual labeling, the unlabeled dataset is also added to this study to participate in the training. The superposition mode of multiple signals is used as the label, and only the signal types need to be labeled on the whole image. The advantages of this method are as follows: first, a small number of unlabeled data images can be added to the labeled dataset for training, and the robustness of the classifier can be improved by adjusting the number of unlabeled datasets. Second, taking the unlabeled dataset as all the training data is more in line with the application scenario, can also reduce the tag-traversal time of the model in training in terms of the number of tags and improve the response speed of the system.

![Figure 11. Range of labels in the case of multiple signal overlap.](image)

### 4.2. Experimental Analysis Based on the YOLOv5 Classification Model

Figure 12a shows the n size model, (b) the s size model, (c) the m size model, and (d) the x size model, which can be seen from the chart. The x in a CWD image-size model identifies the characteristics of the largest number, and the residual quantity is small, but the training of approximately 3000 copies of the image dataset takes a long time—here, it reached 50 h. The n-size model only takes approximately 6 h to train, but the missed detection rate and recognition accuracy cannot reach the classification standard. The m size model has a low recognition rate for small target features, so the s size model is selected for training in this study to achieve a balance between training time and comprehensive recognition effect.

![Figure 12. Training validation procedure of the four models.](image)

To ensure the real-time performance of engineering applications, the boundary conditions of recognition accuracy and training time should be considered comprehensively according to the working environment. Therefore, this paper lists the training duration and recognition accuracy of YOLOv5 models of different sizes. As shown in Table 3, the dataset...
used in each experiment is the same as the experimental environment. It can be seen that the training time of YOLOv5n is the shortest, 376 min. The training time of the YOLOv5s model is 1170 min, and the ratio of training time to prediction accuracy is second only to YOLOv5x. Although the x size model has high accuracy, the training time is as high as 3000 min, which is not conducive to updating the dataset.

| Size      | Precision (%) | GFLOPs | Parameters | Data Pictures | Training Time (min) |
|-----------|---------------|--------|------------|--------------|--------------------|
| YOLOv5n   | 42.3%         | 6.0    | 2.9 M      | 3500         | 376                |
| YOLOv5s   | 53%           | 17.1   | 7.7 M      | 3500         | 1170               |
| YOLOv5m   | 89.2%         | 43.6   | 21.4 M     | 3500         | 1982               |
| YOLOv5x   | 96.5%         | 219.7  | 219.0 M    | 3500         | 3000               |

To facilitate the selection of the hardware platform, the GFlops value is listed in this paper. It can be seen that the computing pressure caused by different size models on GPU varies greatly. In a processing platform that needs a fast response, a suitable device can be selected quickly according to the real-time processing capability of the hardware processing platform.

To compare the recognition rate (PSR) of the signal preprocessing method proposed in this paper with other time-frequency analysis methods, −15 −0 dB time-frequency analysis images are used as input to the YOLOv5s model to obtain recognition accuracy. As shown in Figure 13, the trained model can distinguish subsignals of a single category and compound category. Due to the large improvement in the noise-suppression performance of the preprocessing algorithm, the characteristics of each signal can be highlighted at the image level, which improves the model recognition rate and reduces the probability of missing detection.

Figure 13. Identification results of the YOLOv5s model.

Figure 14a lists the recognition-rate accuracy of the mainstream time-frequency processing method and the preprocessing method proposed in this paper for mixed signals under 0–15 dB. After the preprocessing method, it is imported into the YOLOv5s model for feature recognition. Compared with common time-frequency analysis methods such as CWD, STFT, WVD, and wavelet transform, the data-preprocessing method proposed in this paper has certain advantages in anti-noise, so it obtains higher recognition accuracy at low SNRs. At 0 dB, the mainstream preprocessing method can maintain high recognition
accuracy under the YOLOv5s model because of the small noise influence. It also proves the effectiveness of the YOLOv5s model.

![Figure 14](image1.png)

**Figure 14.** Recognition accuracy of different preprocessing methods in YOLOv5s: (a) the recognition accuracy of different preprocessing methods in YOLOv5s and (b) the comparison of recognition accuracy between the proposed multisignal recognition framework and LPI-NET and other frameworks.

Figure 14b shows the recognition-accuracy comparison between the proposed multisignal recognition framework and other LPI-NET frameworks. It can be seen from the figure that the CWD time-frequency analysis method has more advantages than STFT, WVD, and other methods in the multisignal preprocessing stage with a low SNR because CWD has better cross-term suppression performance. After using CWD time-frequency analysis, the network detection accuracy is more than 90% when approaching −8 dB, and the recognition accuracy is more than 50% when approaching −15 dB.

Figure 15a shows the comparison of recognition accuracy between this paper and other classification frameworks. As can be seen from the figure, under low SNR, the proposed method has higher recognition accuracy than other frameworks, and there is no phenomenon of declining accuracy. This paper conducts a 10-fold cross-validation experiment to comprehensively measure the recognition accuracy and stability of the YOLOv5S model. It can be seen from Figure 15b that the recognition accuracy of different sub-datasets fluctuates to a certain extent, but the recognition accuracy is 53%.

![Figure 15](image2.png)

**Figure 15.** (a) Comparison between the proposed multisignal recognition framework and CNNs in other papers. (b) 10-fold cross-validation of −15 dB mixed signals.
This paper compares the training time and mAP value of the YOLOv5s model and other classifiers. As can be seen from Figure 16, the training time of the YOLOv5s model is 19.5 h, about three times that of the n-size model, but the mAP value exceeds most classifiers. The YOLO5s model has a faster training speed and higher recognition accuracy. Although it does not exceed the x size model, the accuracy of the s size model is acceptable compared to the ultra-long training time. If the calculation time is prolonged, the precision can be improved again from the aspect of preprocessing.

Figure 16. (a) Training time of different object detection models. (b) mAP values of different object detection models.

Table 4 lists the single classification time consumption for different sizes of YOLO models and other classifiers. The single classification time consumption of the YOLO model is not directly related to size. Compared with the single recognition time of other classifiers, the single recognition time of the YOLO model is significantly improved compared with other classifiers, which has significant advantages in real-time performance.

### Table 4. Test time of multiple target-detection models.

| YOLOv5n | YOLOv5s | YOLOv5x | R-CNN | SPPNet | Fast-R-CNN |
|---------|---------|---------|-------|--------|------------|
| 0.37 s  | 0.4 s   | 0.45 s  | 47 s  | 2.3 s  | 0.32 s     |

To evaluate the stability and convergence effect of the training process, this paper lists the curve of the loss function and mAP value. Figure 17 shows the convergence process of the loss function and the change curve of the mAP value of the YOLO5s model in the training process. In this paper, 1000 cycles are used to complete the training process, and the loss curve tends to converge when approaching 1000 cycles without any mutation. The mAP value is continuously improved to achieve the ideal training effect.
5. Conclusions

Aiming at the identification problem of radiation source modulated signals in a low SNR environment, this paper proposes a signal preprocessing method based on the fusion of cyclic mean denoising, IFFT, and CWD time-frequency analysis. The length of the cyclic domain is adjusted by controlling the cyclic frequency. This can not only ensure the noise-suppression effect but also reduce the calculation amount of CWD time-frequency analysis and improve the real-time performance of the system. The noise-suppression effect is obvious when the SNR is above \(-15\) dB. When it is lower than \(-15\) dB, the operation time is significantly increased to ensure the noise-suppression effect.

The time-frequency images were identified using the YOLOv5s model, and a radiation source mixed signal dataset composed of 3500 CWD time-frequency images was generated. The classifier was trained by two forms of fusion, labeled and unlabeled. The classifier was able to identify independent signals and mixed-signal features and mark the position of signal features in the graph, paving the way for further parameter estimation. Compared with the single CWD, STFT, WVD, and wavelet transform time-frequency analysis methods, the proposed preprocessing method performs better in the classifier and is less affected by SNR fluctuations. The prediction accuracy with SNR changes has no mutation or decline, and the signal-classification accuracy under low SNR increases significantly. In the process of training, the loss function and mAP value mutations are lower, and an ideal training effect is achieved. In the process of signal preprocessing, the problem of low accuracy of training, the loss function and mAP value mutations are lower, and an ideal training effect is achieved. In the process of signal preprocessing, the problem of SNR fluctuations. The prediction accuracy with SNR changes has no mutation or decline, and the signal-classification accuracy under low SNR increases significantly. In the process of training, the loss function and mAP value mutations are lower, and an ideal training effect is achieved. In the process of signal preprocessing, the problem of low accuracy of training, the loss function and mAP value mutations are lower, and an ideal training effect is achieved. In the process of signal preprocessing, the problem of low accuracy of training, the loss function and mAP value mutations are lower, and an ideal training effect is achieved. In the process of signal preprocessing, the problem of low accuracy of training, the loss function and mAP value mutations are lower, and an ideal training effect is achieved. In the process of signal preprocessing, the problem of low accuracy of training, the loss function and mAP value mutations are lower, and an ideal training effect is achieved.

6. Future Work Outlook

The real-time signal processing scheme proposed in this paper can be applied to embedded systems, and we will further improve the preprocessing and classifier work in the future.

Firstly, we will port the technical framework to the deep learning framework. Secondly, we will use the output results of the classifiers on the time-frequency graph to estimate the signal parameters of the mixed emitter. This estimation method does not need to be calculated by the model. It just needs to read the size and position of the marker box in real-time.

Finally, we will use the signals collected in the real world to generate large-scale mixed-emitter signal datasets in the form of manual and automatic labeling. After the production of the dataset, we will transplant the technical framework into a high-performance embedded system with a deep learning function, which can complete online learning, real-time analysis, and database updates.
Author Contributions: Conceptualization, D.H. and X.Y.; methodology, J.D.; software, D.H. and X.W.; validation, X.Y. and X.H.; formal analysis, D.H. and J.D.; investigation, D.H.; resources, X.Y. and X.W.; data curation, X.H. and J.D.; writing—original draft preparation, D.H. and X.Y.; writing—review and editing, D.H. and X.Y.; visualization, X.H.; supervision, X.Y. and X.H.; project administration, X.Y.; funding acquisition, X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by the National Natural Science Foundation of China, grant numbers 61973037 and 61871414.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

| Acronym | Description                                      |
|---------|--------------------------------------------------|
| AM      | Amplitude modulation                             |
| BN      | Batch Normalization                              |
| CBL     | Convolution Batch Normalization Leakyrelu        |
| CMWT    | Complex Morlet Wavelet Transform                 |
| CNN     | Convolutional Neural Network                     |
| CNN-DNN | Convolutional Neural Network-Deep Neural Networks|
| CSP     | Cross Stage Partial                              |
| CWD     | Choi–Williams Distribution                      |
| DCN     | Deformable Convolution Networks                  |
| BILSTM  | Bi-Directional Long Short-Term Memory            |
| FFT     | Fast Fourier Transform                           |
| FM      | Frequency Modulation                             |
| FM_SAW  | Frequency Modulation Sawtooth wave               |
| FM_TRI  | Frequency Modulation Triangle wave               |
| FPN     | Feature Pyramid Networks                         |
| PAN     | Path Aggregation Network                         |
| IFFT    | Inverse Fast Fourier Transform                   |
| GIOU    | Generalized Intersection over Union              |
| LPI-NET | Lightweight Pyramid Inpainting Network           |
| mAP     | Mean Average Precision                           |
| NMS     | Non Maximum Suppression                          |
| PWVD    | Pseudo Wigner–Ville Distribution                 |
| RRSARNet| Radar Radio Sources Adaptive Recognition Network |
| SEP-ResNet | Channel-Separable Residual Neural Network      |
| SFM     | Sine Frequency Modulation                        |
| SiLU    | Sigmoid Linear Unit                              |
| SNR     | Signal-Noise Ratio                               |
| SPP     | Spatial Pyramid Pooling                          |
| STFT    | Short-time Fourier Transform                     |
| T-F     | Time-Frequency                                   |
| VHF     | Very High Frequency                              |
| YOLO    | You Only Look Once                               |
Appendix A Pseudocode

Pseudocode for Preprocessing Methods

Calculate the Cyclic Mean
Input: Mixed Signal Sequences.

Step 1: Setting the Value of alpha and attach, alpha = 2*pi/2000; It can be adjusted according to the desired noise-suppression effect and calculation time.
Step 2: Setting length of the cyclic sequence, length = ceil(2*pi/alpha) + attach;
Step 3: Construct the cyclic mean sequence space m
for k = 0:length
   m(k + 1)= mean(signal.*exp(-1i*k*alpha*t));

Calculate the IFFT of the Cyclic mean series
Input: The Cyclic mean sequence of mixed-signal sequences.

IFFT_signal = ifft(m_ga_add,n);
This step should adjust the IFFT sequence length parameter n to reduce the CWD image’s computation time.

Run the CWD algorithm
Input: Time domain signal sequence after IFFT operation.

cwd_signal = tfrcw(IFFT_signal);
CWD_FIG = image(cwd_signal);

Output: Mixed signal CWD image.

Pseudocode for YOLOv5 Detect Part (Important Steps)

Loading system libraries, setting the system environment.
Input: The CWD picture of Mixed signals.

Step 1: The size of the zoom image is [640 640]
Step 2: Conf_value = 0.3 ## Boxes whose confidence is higher than 30% will be reserved.
Step 3: Iou_value = 0.4 ## IOU boxes whose value is higher than this value can be reserved.
Step 4: max_det = 10; ## Maximum number of targets. This value can change depending on the testing environment.

Predicting part

Step 1: Use blank picture (0 matrix) prediction to accelerate the prediction process. Import single-label and multi-label datasets.
Step 2: im/= 255 # 0–255 to 0.0–1.0 # Normalize the image.
   if len(im.shape) == 3:
      im = im[None] ## Add the 0th dimension.
Step 3: pred = model(im, augment = augment, visualize = visualize)
   pred = non_max_suppression(pred,conf_thres,iou_thres,classes, agnostic_nms, max_det = max_det) ## Prediction box save and NMS
Step 4: if len(det):
   # Rescale boxes from img_size to im0 size.
   det[:, :4] = scale_coords(im.shape [2:], det[:, :, 4], im0.shape).round()
   # Resize the annotated bounding_box to the same size as the original image (because the original image has been scaled up and down during training)
   # Print results
Output the predicted time for each image.
   The accuracy distribution is returned according to the result of each prediction, and the lowest prediction accuracy category is output.

Output: Mixed signal CWD image with predict box.
The target category.
The Degree of confidence.
The box position.
Appendix A The Key Formula of Preprocessing Algorithm

\[ u_{AM}(t) = U_{cm}(1 + m_a \cos(f_m t)) \cos(2\pi f_c t) + n(t) \quad (A1) \]

\[ M_{uA}^a(t) = < U_{cm}(1 + m_a \cos(f_m t)) \cos(2\pi f_c t)e^{-j2\pi at} > + < n(t)e^{-j2\pi at} > \quad (A2) \]

\[ M_{uA}^a(t) = < U_{cm} \cos(2\pi f_c t)[\cos(2\pi at) - j \sin(2\pi at)] > \quad (A3) \]

\[ M_{uA}^a(t) = < \frac{1}{2} U_{cm} \{ \cos(2\pi (f_c + \alpha)t) + \cos(2\pi (f_c - \alpha)t) \} > \quad (A4) \]

\[ M_{uB}^a(t) = < U_{cm} m_a \cos(2\pi f_c t) \cos(f_m t)[\cos(2\pi at) - j \sin(2\pi at)] > \quad (A5) \]

\[ M_{uB}^a(t) = < \frac{1}{2} m_a U_{cm}[\cos(2\pi f_c t + f_m t) + \cos(2\pi f_c t - f_m t)][\cos(2\pi at) - j \sin(2\pi at)] > \quad (A6) \]

\[ M_{uB}^b(t) = < u_{fm}(t)e^{-j2\pi at} > = < U_m(\cos(2\pi f_c t + m_f \sin \Omega t)\cos(2\pi at) - j \sin(2\pi at) > \quad (A8) \]

\[ M_{uB}^a(t) = < u_{fm}(t)e^{-j2\pi at} > = < U_m(\cos(2\pi f_c t + m_f \sin \Omega t)\cos(2\pi at) - j \sin(2\pi at) > \quad (A9) \]

\[ M_{uB}^a(t) = < u_{fm}(t)e^{-j2\pi at} > = < U_m(\cos(2\pi f_c t + m_f \sin \Omega t)\cos(2\pi at) - j \sin(2\pi at) > \quad (A10) \]

\[ M_{uB}^a(t) = < u_{fm}(t)e^{-j2\pi at} > = < U_m(\cos(2\pi f_c t + m_f \sin \Omega t)(\cos(2\pi at) + j \sin(2\pi at) > \quad (A11) \]

\[ M_{uB}^a(t) = < u_{fm}(t)e^{-j2\pi at} > = < U_m(\cos(2\pi f_c t + m_f \sin \Omega t)(\cos(2\pi at) + j \sin(2\pi at) > \quad (A12) \]

\[ M_{uB}^a(t) = < u_{fm}(t)e^{-j2\pi at} > = < U_m(\cos(2\pi f_c t + m_f \sin \Omega t)(\cos(2\pi at) + j \sin(2\pi at) > \quad (A13) \]

\[ M_{uB}^a(t) = < u_{fm}(t)e^{-j2\pi at} > = < U_m(\cos(2\pi f_c t + m_f \sin \Omega t)(\cos(2\pi at) + j \sin(2\pi at) > \quad (A14) \]

\[ M_{uB}^a(t) = < u_{fm}(t)e^{-j2\pi at} > = < U_m(\cos(2\pi f_c t + m_f \sin \Omega t)(\cos(2\pi at) + j \sin(2\pi at) > \quad (A15) \]

References

1. Yu, N.; Liang, W.; Shi, L.; Liu, W. A new blind detection algorithm based on cyclic statistics. Tactical Missile Technol. 2017, 6, 94–99. [CrossRef]

2. Yang, F.; Li, Z.; Luo, Z. Research on VHF Band Signal Modulation Classification and Recognition Methods Based on Algorithm of First-Order Cyclic Moment. Telecom Sci. 2014, 30, 76–81.

3. Lin, X.; Zhang, L.; Wu, Z.; Jiang, J. Modulation recognition method based on convolutional neural network and cyclic spectrum images. J. Terahertz Sci. Electron. Inf. 2019, 21, 617–622.
4. Huynh-The, T.; Doan, V.-S.; Hua, C.-H.; Pham, Q.-V.; Nguyen, T.-V.; Kim, D.-S. Accurate LPI Radar Waveform Recognition with CWD-TFA for Deep Convolutional Network. *IEEE Wirel. Commun. Lett.* 2021, 10, 1638–1642. [CrossRef]

5. Zhang, Q.; Ji, H.; Jin, Y. Cyclostationary Signals Analysis Methods Based on High-Dimensional Space Transformation Under Impulsive Noise. *IEEE Signal Process. Lett.* 2022, 28, 1724–1728. [CrossRef]

6. Dong, P.; Wang, H.; Xiao, B.; Chen, Y.; Sheng, T.; Zhang, H.; Zhou, Y. Study for classification and recognition of radar emitter intra-pulse signals based on the energy cumulant of CWD. *J. Ambient Intell. Humaniz. Comput.* 2021, 12, 9809–9823. [CrossRef]

7. Gao, J.; Wang, X.; Wu, R.; Xu, X. A New Modulation Recognition Method Based on Flying Fish Swarm Algorithm. *IEEE Access* 2021, 9, 76689–76706. [CrossRef]

8. Zhang, L.; Liu, H.; Yang, X.; Jiang, Y.; Wu, Z. Intelligent Denoising-Aided Deep Learning Modulation Recognition With Cyclic Spectrum Features for Higher Accuracy. *IEEE Trans. Aerosp. Electron. Syst.* 2021, 57, 3749–3757. [CrossRef]

9. Liu, K.; Gao, W.; Huang, Q. Automatic Modulation Recognition Based on a DCN-BiLSTM Network. *Sensors* 2021, 21, 1577. [CrossRef] [PubMed]

10. Lang, P.; Fu, X.; Martorella, M.; Dong, J.; Qin, R.; Feng, C.; Zhao, C.X. RRSARNet: A Novel Network for Radar Radio Sources Adaptive Recognition. *IEEE Trans. Veh. Technol.* 2021, 70, 11483–11498. [CrossRef]

11. Liu, L.; Li, X. Radar signal recognition based on triplet convolutional neural network. *EURASIP J. Adv. Signal Process.* 2021, 2021, 112. [CrossRef]

12. Mao, Y.; Ren, W.; Yang, Z. Radar Signal Modulation Recognition Based on Sep-ResNet. *Sensors* 2021, 21, 7474. [CrossRef] [PubMed]

13. Wang, C.; Wang, J.; Zhang, X. Automatic radar waveform recognition based on time-frequency analysis and convolutional neural network. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 2437–2441. [CrossRef]

14. Zhang, M.; Diao, M.; Guo, L. Convolutional Neural Networks for Automatic Cognitive Radio Waveform Recognition. *IEEE Access* 2017, 5, 11074–11082. [CrossRef]

15. Kong, S.-H.; Kim, M.; Hoang, L.M.; Kim, E. Automatic LPI Radar Waveform Recognition Using CNN. *IEEE Access* 2018, 6, 4207–4219. [CrossRef]

16. Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, Montreal, BC, Canada, 11–17 October 2021; pp. 2778–2788. [CrossRef]

17. Zhao, Y.; Shi, Y.; Wang, Z. The Improved YOLOV5 Algorithm and Its Application in Small Target Detection. In Proceedings of the International Conference on Intelligent Robotics and Applications, Harbin, China, 1–3 August 2022; pp. 679–688. [CrossRef]

18. Ye, J.; Yuan, Z.; Qian, C.; Li, X. CAA-YOLO: Combined-Attention-Augmented YOLO for Infrared Ocean Ships Detection. *Sensors* 2022, 22, 3782. [CrossRef]

19. Ganesh, P.; Chen, Y.; Yang, Y.; Chen, D.M.; Winslett, M.; Soc, I.C. In YOLO-ResNet: Towards High Accuracy Real-time Object Detection on Edge GPUs. In Proceedings of the 22nd IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 4–8 January 2022; pp. 1311–1321.

20. Guo, Q.; Liu, J.; Kaliuzhnyi, M. YOLOX-SAR: High-Precision Object Detection System Based on Visible and Infrared Sensors for SAR Remote Sensing. *IEEE Sens. J.* 2022, 22, 17243–17253. [CrossRef]

21. Kosuge, A.; Suehiro, S.; Hamada, M.; Kuroda, T. mmWave-YOLO: A mmWave Imaging Radar-Based Real-Time Multiclass Object Recognition System for ADAS Applications. *IEEE Trans. Instrum. Meas.* 2022, 71, 2509810. [CrossRef]

22. Song, Y.Y.; Xie, Z.Y.; Wang, X.W.; Zou, Y.Q. MS-YOLO: Object Detection Based on YOLOv5 Optimized Fusion Millimeter-Wave Radar and Machine Vision. *IEEE Sens. J.* 2022, 22, 15435–15447. [CrossRef]

23. Hanna, S.; Dick, C.; Cabric, D. Signal Processing-Based Deep Learning for Blind Symbol Decoding and Modulation Classiﬁcation. *IEEE J. Sel. Areas Commun.* 2022, 40, 82–96. [CrossRef]

24. Kim, J.; Harne, R.L.; Wang, K.-W. Online Signal Denoising Using Adaptive Stochastic Resonance in Parallel Array and its Application to Acoustic Emission Signals. *J. Vib. Acoust.* 2021, 144, 031006. [CrossRef]

25. Xu, X.W.; Zhang, X.L.; Zhang, T.W.; Shi, J.; Wei, S.J.; Li, J.W. On-Board Ship Detection in SAR Images Based on L-YOLO. In Proceedings of the IEEE Radar Conference (RadarConf22), New York, NY, USA, 21–25 March 2022.

26. Zhang, C.; Van der Baan, M. Signal Processing Using Dictionaries, Atoms, and Deep Learning: A Common Analysis-Synthesis Framework. *Proc. IEEE* 2022, 110, 454–475. [CrossRef]

27. Xu, C.; Zhang, Y.; Fan, X.J.; Lan, X.J.; Ye, X.; Wu, T.N. An efficient fluorescence in situ hybridization (FISH)-based circulating ge-netically abnormal cells (CACs) identification method based on Multi-scale MobileNet-YOLO-V4. *Quant. Imaging Med. Surg.* 2022, 12, 2961–2976. [CrossRef] [PubMed]

28. Shimura, T.; Umehira, M.; Watanabe, Y.; Wang, X.Y.; Takeda, S. An Advanced Wideband Interference Suppression Technique using Envelope Detection and Sorting for Automotive FMCW Radar. In Proceedings of the IEEE Radar Conference (RadarConf22), New York, NY, USA, 21–25 March 2022.

29. Ma, H.; Sun, Y.; Wu, N.; Li, Y. Relative Attributes-Based Generative Adversarial Network for Desert Seismic Noise Suppression. *IEEE Geosci. Remote Sens. Lett.* 2021, 19, 8023005. [CrossRef]

30. Klintberg, J.; McKelvey, T.; Dammert, P. A Parametric Approach to Space-Time Adaptive Processing in Bistatic Radar Systems. *IEEE Trans. Aerosp. Electron. Syst.* 2021, 58, 1149–1160. [CrossRef]