Article

Multi-Label Radar Compound Jamming Signal Recognition Using Complex-Valued CNN with Jamming Class Representation Fusion

Yunyun Meng, Lei Yu * and Yinsheng Wei

School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China;
myyhit@163.com (Y.M.); weiys@hit.edu.cn (Y.W.)
* Correspondence: yu.lei@hit.edu.cn

Abstract: In the complex battlefield electromagnetic environment, multiple jamming signals can enter the radar receiver simultaneously due to the development of jammers and modulation technology. The received compound jamming signals aggravate the difficulty of recognition and subsequent counter-countermeasure. In the face of strong overlapping signals and unseen jamming signal combinations, the performance of existing recognition methods usually seriously degrades. In this paper, an end-to-end multi-label classification framework combining a complex-valued convolutional neural network (CV-CNN) and jamming class representations is proposed to automatically recognize the jamming signal components of compound jamming signals. A basic multi-label CV-CNN (ML-CV-CNN) is first designed to directly process time–domain complex signals and fully retain jamming signal information. Then, the jamming class representations are generated using prototype clustering implemented by learning vector quantization, and they are fused with the ML-CV-CNN using class decoupling implemented by the attention mechanism to construct a multi-label class representation CV-CNN (ML-CR-CV-CNN), which can better learn the class-related features required for recognition. Finally, an adaptive threshold calibration is adopted to obtain optimal recognition results by multi-threshold discrimination. Simulation results verify that the proposed method has superior recognition performance, which is reflected in the strong robustness to the varying jamming-to-noise ratio (JNR) and power ratio, faster convergence speed with high JNRs, and better generalization for unseen jamming signal combinations.

Keywords: radar jamming signal recognition; compound jamming signal; multi-label classification; complex-valued convolutional neural network; jamming class representation fusion

1. Introduction

In modern warfare, with increasing jamming sources and advanced modulation technology, the battle space is filled with a variety of dynamic electromagnetic jamming signals, which are manifested as ubiquitous in space, dense, and overlapping in the time and frequency spectrum. Automatic recognition of jamming types is important for reasonably deploying anti-jamming resources and ensuring anti-jamming performance. Relying on the development of the capability and number of jammers, the generation of jamming signals is more convenient, and the jamming strategy is more flexible. In order to make better use of jamming resources and achieve better jamming effects, the compound jamming signals with overlapping multiple jamming signals are generated by modulation technology or multi-jammer collaboration to interfere with radar systems. For the purpose of providing sufficient a priori information for anti-jamming, it is necessary to find an appropriate method that can effectively recognize not only single jamming signals and compound jamming signals, but also the specific components in compound jamming signals.

Generally, jamming recognition is mainly carried out in two aspects. At the data processing level, jamming signals are recognized by the likelihood detection and goodness-

of-fit detection according to amplitude fluctuation characteristics [1,2]. The approach is based on probability statistical models and usually requires a lot of prior knowledge. Moreover, model mismatches caused by missing parameters can lead to serious degradation of recognition performance. At the signal processing level, jamming signals are recognized by feature extraction and classifiers [3,4]. When there are many types of jamming signals, the separable feature parameters are difficult to extract with expert knowledge, and the computational complexity is also high. Recently, inspired by the powerful advantages of deep learning in feature extraction, deep neural networks (DNN) with various frameworks have been validated in a variety of radar signal recognition tasks, including SAR image recognition [5,6], automatic modulation recognition [7–9], high-resolution range profile recognition [10], and waveform recognition [11,12]. In terms of jamming signal recognition, most methods usually require data preprocessing. The one-dimensional radar echo data are converted into two-dimensional images that are more suitable for the input of DNNs [13–15]. In addition, one-dimensional jamming data are also attempted to be processed directly by DNNs. In [16], a complex-valued convolutional neural network (CNN) using raw jamming signals is constructed for fast recognition. In [17], two CNN models are adopted to extract two-dimensional time-frequency image features and one-dimensional signal features respectively, and then the two features are fused to further improve recognition performance. However, most studies on the jamming signal recognition task mainly focus on single jamming signals, and the existing recognition models are not completely applicable to compound jamming signals.

For compound jamming signal recognition, the traditional idea is to perform signal separation before recognizing signal components. First, the overlapping signals are separated into independent signal components by blind separation algorithms such as independent component analysis and natural gradient independent component analysis, and then the final classification decision is made using the feature parameters obtained by cumulants and wavelet transform [18,19]. In order to ensure good separation performance, these methods require adequate receiving channels and the accurate estimation of the number of signal sources, which have poor recognition accuracy for single-channel compound signals. The signal separation required for recognition also brings additional computational costs. A cumulant-based maximum likelihood classification method omits the signal separation process and directly uses composite cumulants for classification decisions [20]. Nevertheless, it still needs to estimate the number of sources and channels as a priori information for calculating composite cumulants. And the increasing compound signal components also lead to the degradation of recognition performance. In addition, some methods based on DNNs are proposed to recognize compound signals and avoid the signal separation process. Relying on the distribution differences of signal components in the time-frequency domain, a deep CNN model is adopted to recognize segmented time-frequency images obtained by a repeated selective strategy [21]. And a deep object detection network is used to recognize jamming types and locate position information [22]. However, these methods cannot deal well with compound signals with strongly overlapping signal components in the time-frequency domain when the suppression jamming power is greater than the power of the other jamming signal components. Some models based on multi-class classification are also used for compound signal recognition [23,24]. They define each possible combination as a separate class, which is marked with a single label during the recognition process. A Siamese-CNN model based on the original echo data [25] and a jamming recognition model based on power spectrum features [26] realize the recognition of additive and convolutional compound jamming signals, respectively. For the recognition models based on multi-class classification, the size of the output node representing classes increases exponentially with the number of candidate signal components. The feature extractors in models are required to extract the classifiable features of all combinations to ensure recognition performance. Unfortunately, the increasing and overlapping candidate signal components greatly enhance the difficulty of feature extraction, which leads to the increasing model complexity and declining recognition performance.

In addition to multi-class classification methods, another potential strategy that can be used for compound jamming signal recognition is multi-label classification [27,28]. It has been successfully applied in many fields such as automatic video annotation [29], action recognition [30], visual object recognition [31], web page categorization [32], audio annotation [33], and image recognition [34,35]. In terms of radar signal recognition, the multi-instance multi-label learning frameworks based on a CNN and a residual attention-aided U-net generative adversarial network realize the automatic recognition of the overlapping low probability of intercept radar signals [36,37]. A CNN-based multi-label framework using time-frequency images is proposed for compound jamming signal recognition [38]. It has good recognition accuracy at multiple values of the jamming-to-noise ratio (JNR). However, the time-frequency preprocessing comes with an additional computational burden. And the strong overlapping effect of time-frequency images generated by high-power suppression jamming signals is ignored, which can lead to serious degradation of recognition performance when the power of signal components is unbalanced.

Inspired by the multi-label classification methods and considering the problems mentioned above, a multi-label class representation complex-valued convolutional neural network (ML-CR-CV-CNN) with an end-to-end manner is proposed for compound jamming signal recognition. The main contributions of this paper are summarized as follows:

1. A basic ML-CV-CNN is designed to directly process the raw one-dimensional time–domain compound jamming signals. The introduction of complex-valued components reduces the possible information loss caused by data preprocessing and enhances the weak feature extraction ability for strong overlapping signals.
2. Using class decoupling implemented by the attention mechanism, the basis ML-CV-CNN is fused with the jamming class representations generated by learning vector quantization (LVQ) to construct the ML-CR-CV-CNN, which enhances the class-related feature learning of compound jamming signals and improves the recognition performance of unseen combinations in training.
3. Simulation results show that the proposed method can effectively recognize compound jamming signals, especially in the face of high-power suppression jamming signals with strong overlapping effect. The performance improvement is mainly manifested in robustness to the variation of the power ratio (PR), model convergence speed, and generalization to unseen combinations.

The rest of this paper is organized as follows: Section 2 gives the compound jamming signal model and introduces the basic theory of the recognition model. Section 3 describes the proposed recognition method in detail. In Section 4, the data description, experiment configurations, evaluation metrics and experimental results are presented and analyzed. Finally, the conclusions are summarized in Section 5.

2. Signal Model and Problem Formulation

2.1. Compound Jamming Signal Model

The complex electromagnetic environment contains various types of jamming signals with different effects. In general, suppression jamming signals with high power can completely cover real targets to destroy radar detection, and deception jamming signals can generate false targets to interfere with the determination of real targets. In the practical application, the compound jamming signals formed by enemy jammers using multiple jamming technologies can simultaneously greatly enhance the jamming ability. According to generation mechanisms, the compound jamming signals can be divided into the additive compound, multiplicative compound, and convolutional compound. In this paper, we mainly consider the most common scenario in which multiple jammers work together in a complex electronic countermeasure environment. As shown in Figure 1, many different types of jamming signals enter the radar receiver simultaneously to generate additive compound jamming signals.
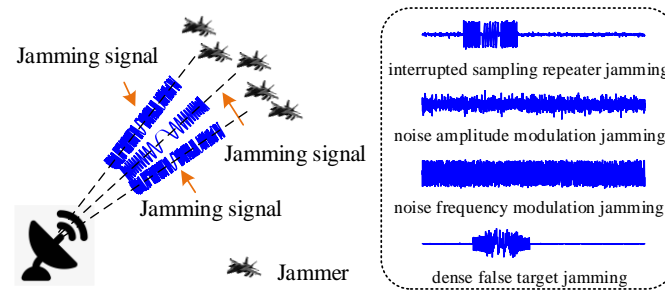
Figure 1. Schematic diagram of compound jamming signal generation and typical jamming signals.

In the observation range, assuming that m jammers emit different types of jamming signals to interfere with the radar system, the radar receiving antenna receives all jamming signals at the same time. And each type of jamming signal enters with the same probability. The formed compound jamming signal can be expressed as

$$S(n) = \sum_{i=1}^{m} \alpha_i J_i(n) + \omega(n),$$ (1)

where $J_i(n)$ denotes the discrete-time jamming signal generated by the i-th jammer, and n is the sampling point. α_i denotes the amplitude coefficient generated by the signal propagation. $\omega(n)$ represents the additive white Gaussian noise (AWGN) with zero mean.

We assume that the electromagnetic space contains C kinds of candidate jamming signals, and different types of jamming signals can be mixed randomly. Thus, there are N_c different combinations of compound jamming signals. The specific value can be calculated according to the following formula:

$$N_c = \sum_{k=1}^{C} C_C^k = 2^C - 1,$$ (2)

where C_C^k is the calculation of the combinatorial number. For a compound jamming signal S, there are k kinds of jamming signal components, and k varies in the range of $[1, 2, \cdots, C]$.

2.2. Recognition Model for Compound Jamming Signals

The compound jamming signal recognition can be regarded as a multi-label classification task. The main purpose is to identify the jamming type of each component in compound signals. Compared with the multi-class classification method, the label corresponding to a compound jamming signal is not a single value, but a vector with a length equal to the number of candidate jamming signal components. The values in the label vector are 1 and 0, which represent the presence and absence of a specific type of jamming signal, respectively.

In the end-to-end multi-label compound jamming signal recognition, it is expected to estimate the label vector y directly from S, which can be formulated as

$$y = f_{multi-label}(S, \theta),$$ (3)

where $y = \{y_c | c = 1, 2, \cdots, C\}$ and $y_c \in \{0, 1\}$. θ is the model parameter. The detailed mapping between the input signal and output label of compound jamming signal recognition models can be seen in Figure 2. It is assumed that the compound jamming signal to be tested contains three different types of jamming signal components. The trained multi-class classifier outputs a single label value at a certain node, while the trained multi-label classifier outputs a label vector with an encoding length of C. There are three values marked as 1 in the label vector, and the value of 1 for the i-th node indicates the existence of the i-th jamming signal component. Obviously, the multi-label classifier reduces the number of output nodes from $2^C - 1$ to C compared to the multi-class classifier.
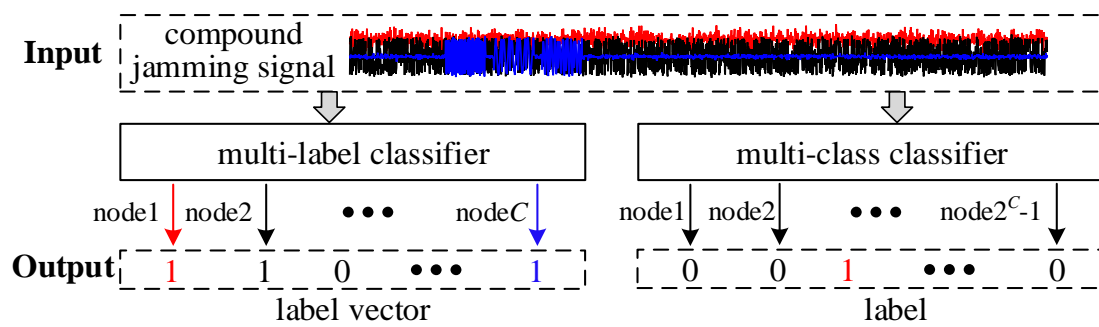
Figure 2. Mapping between the input signal and output label of compound jamming signal recognition models based on multi-label classification and multi-class classification. Three different colors represent three different types of jamming signal components.

3. Proposed ML-CR-CV-CNN for Compound Jamming Signal Recognition

Considering that radar echo signals are one-dimensional complex data, the basic ML-CV-CNN is designed for the end-to-end recognition of compound jamming signals, where each component in the model is in a complex-valued form. In order to further improve model recognition performance and generalization performance, the ML-CR-CV-CNN is constructed by fusing the jamming class representations into the ML-CV-CNN, which is equipped with a higher ability to learn class-related features required for recognition.

The overall framework of the ML-CR-CV-CNN is shown in Figure 3. It is mainly composed of four key modules: compound jamming signal feature extraction, jamming class representation generation, jamming class representation decoupling and adaptive threshold calibration. Firstly, for the single jamming signals belonging to different classes, the jamming class representation generator is constructed to obtain the class representations according to a single jamming signal recognition model and prototype clustering. Secondly, through the class representation decoupling module, the features of compound jamming signals extracted by the compound jamming signal feature extractor are fused with class representations to realize class decoupling by the attention mechanism. And then the obtained class-related feature vectors are used to calculate the existence probability of each jamming class. Finally, an adaptive threshold calibration strategy is adopted to select the optimal decision threshold for the probability of each class by maximizing the F1 value. After multi-threshold discrimination, the predicted label vectors can be determined, which are used to recognize the jamming signal components in compound jamming signals. Detailed descriptions of the main modules are given below.
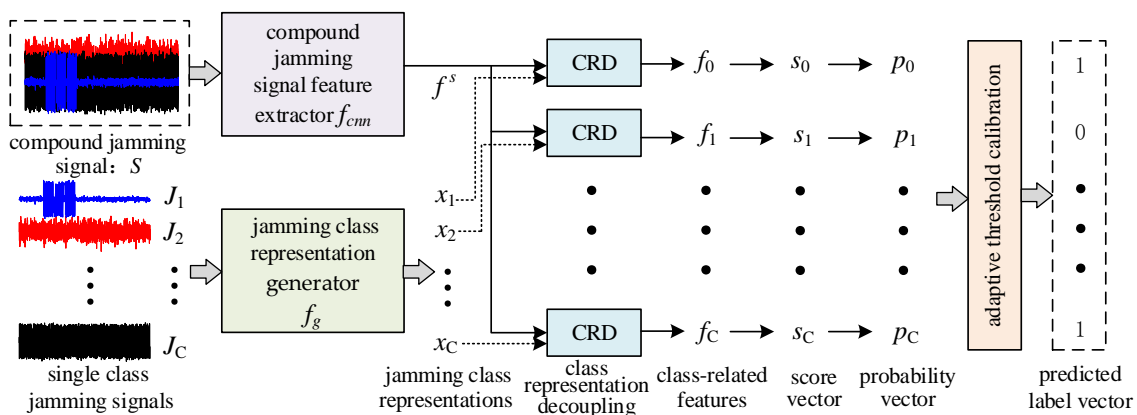

Figure 3. Framework of the ML-CR-CV-CNN model.

3.1. ML-CV-CNN Model Construction and Compound Jamming Signal Feature Extraction

The ML-CR-CV-CNN is constructed by fusing the basic ML-CV-CNN and jamming class representations, where the basic framework of the ML-CV-CNN for compound

jamming signals is shown in Figure 4. The overall structure is an end-to-end recognition model. The input is the one-dimensional time–domain compound jamming signals. After multiple convolutional layers and fully connected layers, the score vectors *s* representing the confidence level of each jamming class can be obtained. According to *s*, the probability vectors *p* indicating the existence of jamming classes can be further calculated by a sigmoid function. And the final predicted label vectors composed of 1 and 0 are estimated to recognize the jamming signal components by performing the adaptive calibration strategy on *p*.
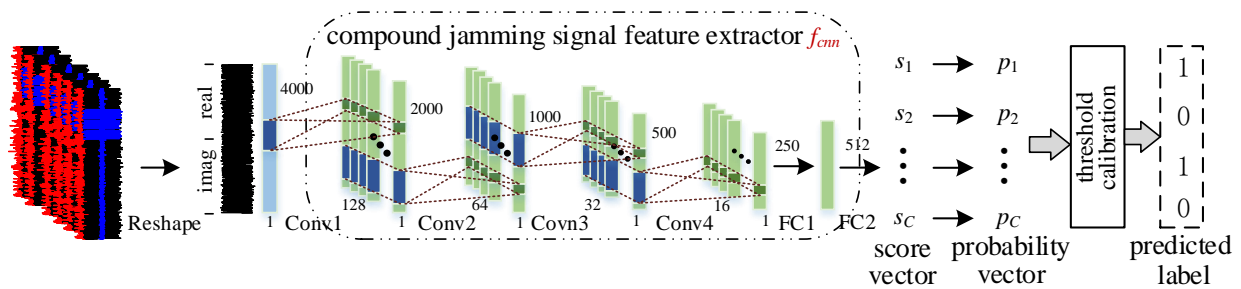

Figure 4. Framework of the ML-CV-CNN model.

The specific composition and parameters of each layer in the ML-CV-CNN are shown in Table 1. Generally, DNNs mainly use real-valued operation. For radar jamming signals, simply separating the real and imaginary parts or considering the amplitude and phase can destroy the original data relationship and lose some information. In order to effectively utilize the phase information and obtain richer signal feature representations for recognition, the layers of convolution, max pooling, activation function and dense in the ML-CV-CNN are implemented by complex-valued operation.

Table 1. Layers of the ML-CV-CNN.

Layer	Type	Output Shape
Input	Reshape	$1 \times 4000 \times 1$
Conv1	Convolution ($5 \times 1 \times 32$)-ReLU-Maxpooling (2×1)-Reshape	$32 \times 2000 \times 1$
Conv2	Convolution ($7 \times 1 \times 64$)-ReLU-Maxpooling (2×1)-Reshape	$64 \times 1000 \times 1$
Conv3	Convolution ($9 \times 1 \times 128$)-ReLU-Maxpooling (2×1)-Dropout (0.5)-Reshape	$128 \times 500 \times 1$
Conv4	Convolution ($11 \times 1 \times 256$)-ReLU-Maxpooling (2×1)-Dropout (0.5)-Reshape	$256 \times 250 \times 1$
FC1	Flatten + Linear + Concatenate-Reshape	$1 \times 512 \times 1$
FC2	Linear	$1 \times 4 \times 1$
Output	Threshold calibration	1×4

According to the mathematical expression of the complex computation, the complex-valued convolution operation can be expressed as the following [39,40]:

$$W * x = (A + iB) * (c + id) = (Ac - Bd) + i(Ad + Bc), \tag{4}$$

where $x = c + id$ is a complex vector, which represents the input of the convolutional layer. $W = A + iB$ is a complex filter matrix, which represents the weight. Among them, c, d, A, and B are all real values. $Ac - Bd$ and $Ad + Bc$ represent the real and imaginary parts of the output of the complex-valued convolutional layer, respectively. It can be inferred that a complex-valued convolutional layer can be realized by a high-dimensional real-valued convolutional layer with two filters. The specific implementation is shown in Figure 5.
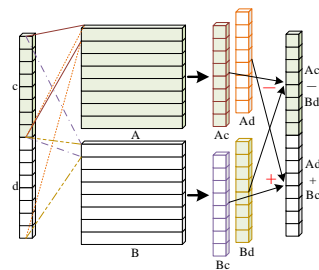
Figure 5. Implementation of a complex-valued convolutional layer.

The operation of a complex-valued dense layer is similar to the complex-valued convolutional layer. It can be implemented by a high-dimensional real-valued dense layer with two filters. In addition, a complex-valued activation function and max pooling operation can be achieved by using the rectified linear unit (ReLU) and max pooling (MaxPool) for the real and imaginary data independently. They can be defined as the following [39]:

$$
\begin{aligned}
\mathbb{C}\mathrm{ReLU}(x) &= \mathrm{ReLU}(\Re(x)) + i\,\mathrm{ReLU}(\Im(x))\\
\mathbb{C}\mathrm{MaxPool}(x) &= \mathrm{MaxPool}(\Re(x)) + i\,\mathrm{MaxPool}(\Im(x))'
\end{aligned}
\tag{5}
$$

where $\mathbb{C}$ denotes the complex-valued form. $\Re(\cdot)$ and $\Im(\cdot)$ are the real and imaginary parts.

The ML-CR-CV-CNN and ML-CV-CNN share the structure of feature extraction. The compound jamming signal feature extractor f_{cnn} in the ML-CR-CV-CNN can be found in Figure 4. For the compound jamming signal S that is input into the ML-CR-CV-CNN model shown in Figure 3, the extracted feature vector f^S can be indicated as

$$
f^S = f_{cnn}(S).
\tag{6}
$$

3.2. Jamming Class Representation Generation

The jamming class representation generator combining a feature extraction network and prototype clustering is utilized to extract the feature vectors that can represent the separability of various types of jamming singles. The feature vectors with sufficient separability are used as jamming class representations, which can help to extract the class-related information from compound jamming signals. Based on the separable feature vectors extracted by a classic end-to-end jamming signal recognition model and the prototype clustering performed by the LVQ algorithm, the jamming class representations can be determined. The specific implementation is shown in Figure 6.
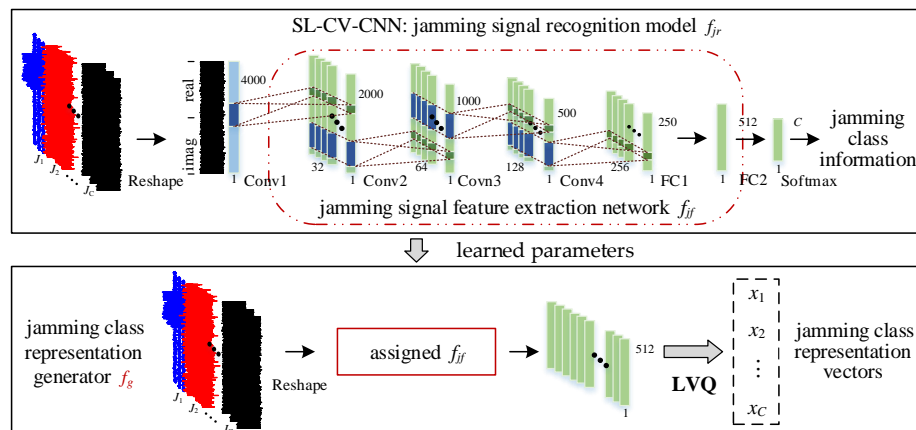

Figure 6. Implementation of the jamming class representation generator.

As shown in Figure 6, the jamming class representation generator shares the feature extraction network with the recognition model implemented by a single-label complex-valued convolutional neural network (SL-CV-CNN). The composition of the convolutional layer and FC1 layer in the SL-CV-CNN is the same as the ML-CV-CNN, which can be found in Table 1. And the mapping between the input signal and output label can be formulated as

$$y_s = f_{\text{SL-CV-CNN}}(J, \theta_s), \tag{7}$$

where J is a single time–domain jamming signal. θ_s is the learnable parameter. $y_s \in \{1, \cdots, C\}$ is a single label value indicating jamming classes.

In the process of obtaining jamming class representations, a dataset where each sample is the jamming signal with a single class is first used to train the SL-CV-CNN model. The learned parameters are saved and utilized to assign the jamming signal feature extraction network f_{jf}. For N single jamming signal samples containing C classes, the sample feature set $\{f^{J_i}, y_{si}\}_{i=1}^{N}$ can be obtained, where f^{J_i} and y_{si} are the feature vector and class label of the i-th jamming signal sample. f^{J_i} can be denoted as

$$f^{J_i} = f_{jf}(J_i). \tag{8}$$

Then, with the aid of class labels, the LVQ algorithm is used to find the prototype vectors of the obtained feature vectors as the jamming class representations by multiple iterations. Using sample mean vectors to initialize a set of prototype vectors $x^* = \{x_c^* | c = 1, 2, \cdots, C\}$, one iteration in the LVQ algorithm can be expressed as

$$\hat{c} = \text{argmin}_{c \in \{1,2,\cdots,C\}} d_{ic} = \text{argmin}_{c \in \{1,2,\cdots,C\}} \| f^{J_i} - x_c^* \|_2$$
$$x_c^* = \begin{cases} x_{\hat{c}}^* + \eta \left(f^{J_i} - x_{\hat{c}} \right), & if \ \hat{c} = y_{si} \\ x_{\hat{c}}^* - \eta \left(f^{J_i} - x_{\hat{c}} \right), & otherwise \end{cases}, \tag{9}$$

where η is the learning rate. After multiple iterations and updates, the obtained prototype vectors are regarded as the final jamming representation vectors $x = \{x_c | c = 1, 2, \cdots, C\}$, where x_c corresponds to the jamming signals J_c belonging to the class c and has the same dimension as the compound jamming signal feature vector f^S. Algorithm 1 shows the detailed process of jamming class representation generation.

Algorithm 1: Jamming class representation generation.

---

1.   **Require**: network model SL-CV-CNN; labeled training samples $\{J_i, y_{si}\}_{i=1}^{N}$
2.   **while not** done **do**
3.   select samples $\{J_i, y_{si}\}_{i=1}^{T}$ with a batch size of T;
4.   **for** $i = 1, \ldots, T$ **do**
5.   predict jamming signal sample label $\hat{y}_s = f_{\text{SL-CV-CNN}}(J, \theta_s)$;
6.   **end for**
7.   compute the cross entropy loss $loss_c$ and update the learned parameter $\theta_s \leftarrow \theta_s - \nabla_{\theta_s} loss_c$;
8.   **end while**
9.   assign the feature extraction network f_{jf} using the learned parameter θ_s; perform prototype clustering on the obtained sample feature set $\{f^{J_i}, y_{si}\}_{i=1}^{N}$; initialize prototype vectors $x^* = \{x_c^* = (1/N_c)\sum_{J_i \in J_c} f^{J_i} | c = 1, 2, \cdots, C\}$;
10. **while not** done **do**
11.   randomly select a sample $\{f^{J_i}, y_{si}\}$ and calculate the distance $d_{ic} = \| f^{J_i} - x_c^* \|_2$; update x_c^* using Equation (9);
12. **end while**
13. **Output**: jamming class representations: final prototype vectors $x = \{x_c | c = 1, 2, \cdots, C\}$.

3.3. Jamming Class Representation Decoupling

The jamming class representation decoupling module mainly uses the jamming class representation vectors to assist the ML-CR-CV-CNN to learn the class-related features required for recognition from compound jamming signals. The class decoupling is realized through the attention mechanism guided by jamming class representations.

After acquiring the jamming class representation vector x_c and f^S, the attention mechanism is adopted to guide f^S to pay more attention to the features related to the class c by fusing x_c. The class-related features are learned by the following steps.

1. The low-rank bilinear pooling method is used to fuse the compound jamming signal feature f^S and x_c, which can be formulated as [34]

$$f_c^S = P^T \left(\tanh \left(\left(U^T f^S \right) \odot \left(V^T x_c \right) \right) \right) + b,$$ (10)

where $\tanh(\cdot)$ is the hyperbolic tangent function. U, V, P, and b are the learnable parameters during training. V^T is the transpose of V. $\odot$ represents the element-wise multiplication.

2. The attention weighting coefficient obtained by an attention function $f_a(\cdot)$ can be calculated as follows:

$$a_c = f_a \left(f_c^S \right),$$ (11)

where $f_a(\cdot)$ is implemented by a fully connected network. The dimension of a_c is the same as f^S, and each value $a_{c,l}$ in a_c represents the importance of the location $(1, l)$ in the compound jamming signal feature vector f^S. The SoftMax function can normalize $a_{c,l}$ to an interval of 0–1, which is defined as

$$a_{c,l} = \frac{e^{a_{c,l}}}{\sum_{i=1}^{L} e^{a_{c,i}}}.$$ (12)

3. The normalized attention coefficient vector a_c is used to perform the weighted average pooling for all locations of f^S, and the obtained feature vector f_c related to the class c can be formulated as

$$f_c = \sum_l a_{c,l} f_l^S.$$ (13)

By performing the above steps for all types of jamming signals, the class-related feature vectors $f = \{f_c | c = 1, 2, \cdots, C\}$ can be obtained, wherein the related features of the jamming signal components that are present are strengthened, while the related features of the jamming signal components that are not present are weakened.

Based on the obtained class-related feature vectors by the jamming class representation decoupling module, the confidence score vector $s = \{s_c | c = 1, 2, \cdots, C\}$ for the existence of jamming classes can be predicted by the function $f_s(\cdot)$ implemented by a fully connected network, where the score s_c of the class c can be expressed as

$$s_c = f_s(f_c).$$ (14)

Then, the sigmoid function $\sigma(\cdot)$ is adopted to convert the predicted score vector to the probability vector $p = \{p_c | c = 1, 2, \cdots, C\}$ with values between 0 and 1, where the probability p_c of the class c can be calculated as

$$p_c = \sigma(s_c) = \frac{1}{1 + e^{-s_c}}.$$ (15)

3.4. Adaptive Threshold Calibration

Generally, only the label with the highest probability is selected as the final result for multi-class classification tasks. However, since the ultimate goal of multi-label classification tasks is to determine whether the jamming signal class represented by each value in the probability vector p exists, the decision threshold needs to be selected for every probability value p_c. Compared with selecting a fixed threshold for different classes, selecting different thresholds adaptively based on the predicted probability of multiple samples can maximize the performance of the multi-label classifier.

In order to make the classification result not biased towards a high accuracy or a high recall rate, the threshold that maximizes the F1 value can be selected as the final decision threshold ε_c for the class c according to the precision–recall curve. It can be expressed as

$$\varepsilon_c = \text{argmaxF1} = \text{argmax} \frac{2\sum_{i=1}^{N} y_{ic} t_{ic}}{\sum_{i=1}^{N} y_{ic} + \sum_{i=1}^{N} t_{ic}}, \tag{16}$$

where $y_{ic}, t_{ic} \in \{0, 1\}$ represent the predicted label and the true label of the class c of the i-th sample, respectively. N is the sample size.

Depending on the obtained multiple thresholds $\varepsilon = \{\varepsilon_c | c = 1, 2, \cdots, C\}$, the probability vector p can be converted into a predicted label vector $y = \{y_c | c = 1, 2, \ldots, C\}$ through a multi-threshold decision function $f_l(\cdot)$, where the label y_c of the class c can be formulated as

$$y_c = f_l(p_c) = \begin{cases} 1, \ if \ p_c \geq \varepsilon_c \\ 0, \ otherwise \end{cases}. \tag{17}$$

4. Experiments and Results

In this section, compound jamming signal recognition experiments are carried out to verify the superiority of the proposed method. The data description, experiment configurations, and evaluation metrics are given. And the results of several groups of comparative experiments are analyzed.

4.1. Data Description

The radar transmitting signal is a linear frequency modulation signal, and the basic signal parameters are shown in Table 2. The simulated jamming signals are single pulse signals, and the key modulation parameters can be adjusted. In the following simulation experiments, the number of candidate jamming signal components is set to 4. The jamming types cover typical suppression jamming generated by noise modulation and deception jamming generated by the full-pulse repeater and interrupted-sampling repeater, which are interrupted sampling repeater jamming, noise amplitude modulation jamming, noise frequency modulation jamming, and dense false target jamming, respectively [22,26]. These four types of jamming are marked as class1-class4 sequentially.

In Table 2, the JNR of compound jamming signals is defined as

$$\text{JNR} = 10\log_{10} \frac{\sum_{i=1}^{k} P_i}{P_{noise}}, \tag{18}$$

where P_i denotes the power of the i-th jamming signal component. P_{noise} denotes the power of the AWGN.

The PR of different jamming signal components in compound jamming signals is defined as

$$\text{PR} = 10\log_{10} \frac{P_i}{P_j}, \tag{19}$$

where P_i and P_j represent the power of the i-th and j-th jamming signal components, respectively.

Table 2. Basic simulation parameters.

	Parameter	Value
Radar transmitting signal	Pulse width	20 μs
	Pulse repetition interval	100 μs
	Bandwidth	10 mhz
	Sampling frequency	20 mhz
Compound jamming signal	JNR	[0, 2.5, 5, 10, 15, 20, 25, 30] db
	PR	[−15, −10, −5, 0, 5, 10, 15, 20, 25] db
	Number of candidate jamming signal components	4
	Number of jamming signal combinations	15

		Parameter	Value
Jamming signal component	Interrupted sampling repeater jamming	Sampling duration	1–4 μs
		Forwarding times	1–3
	Noise amplitude modulation jamming	Bandwidth	20–40 mhz
	Noise frequency modulation jamming	Sweep range	−20–20 mhz
		Sweep cycle	40–80 μs
	Dense false target jamming	Number of false targets	3–5
		False target delay	1–10 μs

Due to the different recognition tasks of the SL-CV-CNN model for single jamming signals and the ML-CR-CV-CNN model for compound jamming signals, two different datasets are required for model training separately.

SL-CV-CNN: Each sample in the dataset contains a single jamming signal class. There are 150 samples for each class and 600 samples in total.

ML-CR-CV-CNN: Since the number of signal components contained in the received compound jamming signals is usually unknown, all possible values $k = \{1, 2, 3, 4\}$ should be considered during the training phase. The compound jamming signals are composed of a total of 15 combinations in the case of 4 candidate jamming signal components. Each sample in the dataset contains k jamming signal components, and the PR defaults to 0. There are 100 samples for each combination and 1500 samples in total.

4.2. Model Training and Experiment Configurations

For the SL-CV-CNN, the training dataset contains N samples $\{J_i, y_{si}\}_{i=1}^{N}$, where J_i and $y_{si} \in \{1, \ldots, C\}$ are the i-th single jamming signal sample and the corresponding class label. The end-to-end recognition model is trained by the cross entropy loss. And the learned parameters of the jamming signal feature extraction network f_{jf} in the SL-CV-CNN are used to assign the parameters of the jamming class representation generator f_g.

For the ML-CR-CV-CNN, the training dataset contains M samples $\{S_i, t_i\}_{i=1}^{M}$, where S_i is the i-th compound jamming signal sample and $t_i = \{t_{ic}|c = 1, 2, \ldots, C\}$ is the true label vector of the i-th sample. Each value of 0 or 1 in the vector indicates the absence or presence of the class c. For the samples with a batch size of T, the training objective using the binary cross entropy is expressed as

$$-\min_{\theta} \sum_{i=1}^{T} \sum_{c=1}^{C} (t_{ic} \log p_{ic} + (1 - t_{ic}) \log(1 - p_{ic})), \tag{20}$$

where p_{ic} is the probability value of the class c of the i-th sample, and θ is the learnable parameter. The ML-CR-CV-CNN is trained by the Adagrad optimization algorithm with an epoch of 120 and a batch size of 32. The initial learning rate is 0.01, and it drops by 50% after every 30 epochs.

Since the parameters pre-trained on the single jamming signal dataset have a certain generalization ability on the compound jamming signal dataset, the parameters of the jamming class representation generator f_g learned by training the SL-CV-CNN model are

utilized to initialize the compound jamming signal feature extractor f_{cnn}. In the process of the ML-CR-CV-CNN model training, the parameters of the first three convolutional layers are fixed, and the other layers are jointly optimized.

4.3. Evaluation Metrics

In the compound jamming signal reorganization, the evaluation metrics are calculated based on the predicted label vectors of the ML-CR-CV-CNN. Due to the cases of complete and partial matching between the predicted label vectors and the real label vectors, the subset accuracy and Hamming loss are used to evaluate the overall recognition performance. In addition, the partial accuracy and label accuracy are calculated to evaluate the fine-grained performance.

Each value in the label vector corresponds to the class label of a jamming signal component in compound jamming signals. For the case of complete matching, the subset accuracy (subsetacc) is used to measure the proportion of correctly recognizing all jamming signal components, and the Hamming loss is used to measure the proportion of misclassifying jamming signal components. They can be calculated as follows [27]:

$$\text{subsetacc} = \frac{1}{N} \sum_{i=1}^{N} y_i = t_i, \tag{21}$$

$$\text{Hamming loss} = \frac{1}{NC} \sum_{i=1}^{N} \sum_{c=1}^{C} y_{ic} \neq t_{ic}, \tag{22}$$

where y_i and t_i denote the predicted label vector and the true label vector of the i-th sample, respectively. y_{ic} and t_{ic} denote the predicted label and the true label of the class c. N represents the number of testing samples. The subsetacc is recorded as 1 when $y_i = t_i$.

In addition to the complete matching of the predicted results, there are also a large number of cases of partial matching. The partial accuracy (partialacc) and label accuracy (labelacc) can be adopted to evaluate the proportion of correctly recognizing at least r jamming signal components and the proportion of correctly recognizing a certain class of jamming signal components, respectively. They can be calculated as follows [38]:

$$\text{partialacc} = \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{c=1}^{C} y_{ic} = t_{ic} \right) \geq r, \tag{23}$$

$$\text{labelacc} = \frac{1}{N} \sum_{i=1}^{N} y_{ic} = t_{ic}. \tag{24}$$

4.4. Results and Performance Analysis

The recognition performance of compound jamming signals is affected by many factors such as JNRs, PRs, the number of candidate jamming signal components, and the selection of training samples. Considering these factors, corresponding experiments are carried out to verify the robustness and effectiveness of the proposed method. In this section, the visualization of jamming class representations is first presented. And then the ML-CV-CNN and MLAMC [38] models based on multi-label classification and the 1D-CNN [25] model based on multi-class classification are adopted as baseline methods to further demonstrate the advantages of the proposed ML-CR-CV-CNN method in terms of model convergence speed, robustness to varying PRs, and generalization to unseen combinations.

4.4.1. Visualization of Jamming Class Representations

As shown in Figure 6, the output of the jamming class representation generator is complex feature vectors with a dimension of 256. The real and imaginary data are serially arranged as 512-dimensional real vectors by the reshaping operation. In order to

conveniently observe the distribution of high-dimensional feature vectors in the latent space, the tSNE [41] is used for dimensionality reduction and visualization.

The visualization result is shown in Figure 7. Observing the feature distribution of single jamming signal samples passing through the jamming signal feature extraction network in the SL-CV-CNN, the feature vectors belonging to different classes of jamming signals form four clusters. The clusters of different classes show clear separability, and the feature vectors in the same cluster show aggregation. As mentioned above, the jamming class representations are generated by finding the prototype vectors belonging to different classes using the LVQ algorithm. As shown in Figure 7, there are four prototypes corresponding to the four clusters composed of the feature vectors belonging to different classes. And the feature distribution of samples belonging to the same class is concentrated near the corresponding prototype. The prototype vectors with significant separability can be used as the jamming class representations, which can represent the class-related information of different classes of jamming signals. In the ML-CR-CV-CNN, the fusion of jamming class representations is beneficial for extracting the class-related features in compound jamming signals to improve recognition performance.
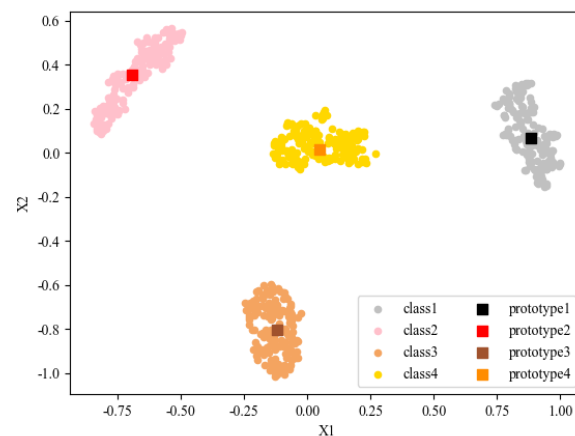

Figure 7. Visualization results of jamming signal features and jamming class representations.

4.4.2. Recognition Model Convergence Speed

In the training process, it is expected to consume the least amount of time to achieve the fast convergence of models. In addition to model structures, the selection of different datasets also affects model convergence speed. In this experiment, the datasets with $C = 4$ and JNR = 15 dB, 25 dB, and 30 dB are selected to verify the convergence performance of the three recognition models based on multi-label classification.

Figure 8 shows the subset accuracy of various models with varying epochs under different JNRs. As shown in Figure 8a, all three models can reach convergence after 70 epochs at JNR = 15 dB. Further observing the results with JNR = 25 dB and JNR = 30 dB shown in Figure 8b,c, it can be seen that the convergence speed of the ML-CV-CNN and ML-CR-CV-CNN models is stable and similar to that of JNR = 15 dB. However, the MLAMC model is more sensitive to varying JNRs. And the convergence speed decreases with increasing JNRs. At JNR = 25 dB and JNR = 30 dB, it needs about 150 and 500 epochs to achieve convergence, which takes longer than the other two models. Therefore, it can be inferred that it is more difficult for the MLAMC model to extract the time-frequency features of different jamming signal components for recognition when suppression jamming signals with high JNRs generate stronger overlap on time-frequency images. In contrast to the MLAMC model, the ML-CV-CNN and ML-CR-CV-CNN models using one-dimensional complex-valued operation are more conducive to extracting the weak features of compound jamming signals at high JNRs, and the model convergence speed is faster and more independent of the influence of JNRs. In addition, benefiting from the jamming class representation

fusion, the ML-CR-CV-CNN model takes less time to achieve the same accuracy and model convergence compared with the ML-CV-CNN model.
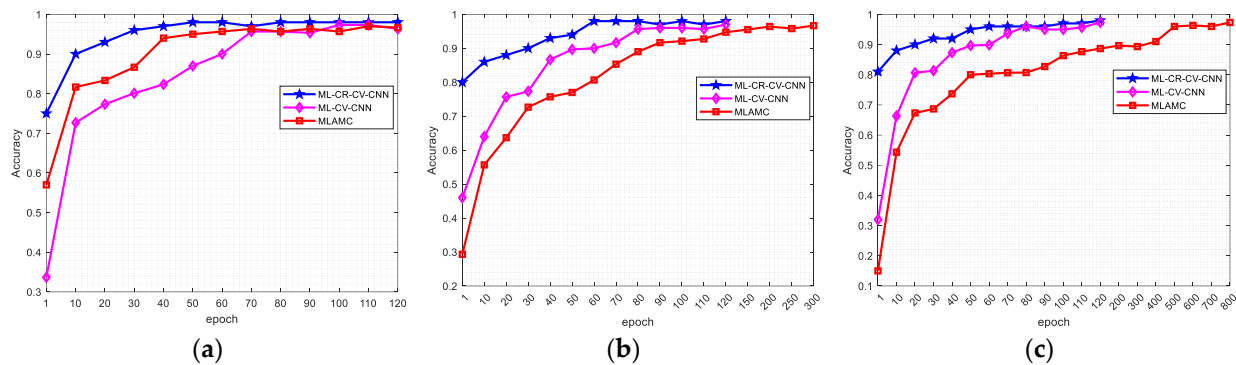

Figure 8. Convergence process at different JNRs. (**a**) JNR = 15 dB; (**b**) JNR = 25 dB; (**c**) JNR = 30 dB.

4.4.3. Recognition Performance Versus Number of Candidate Jamming Signal Components

When the number of candidate jamming signal components *C* is 1, the compound jamming signal recognition is equivalent to the single jamming signal recognition, which usually has good performance. When the value of *C* is large, the compound jamming signals with complex components and multiple combinations can lead to a certain degree of degradation in recognition performance. In this experiment, three datasets with varying values of *C* are used to evaluate the robustness of the proposed method to the varying numbers of candidate jamming signal components. The corresponding specific components are class1-class2, class1-class3, and class1-class4.

Table 3 shows the subset accuracy with varying values of *C* for various recognition models under multiple JNRs. It can be observed that the recognition accuracy decreases with the increasing values of *C*. Compared with the other baseline methods, the proposed ML-CR-CV-CNN model has the slightest performance degradation at *C* = 4 and exhibits optimal results, especially at low JNRs. At JNR = 5 dB and 25 dB, the accuracy improves by 3% and 1% compared to the suboptimal results, respectively. In addition, the three models based on multi-label classification are significantly better than the 1D-CNN model in terms of recognition accuracy and robustness to varying JNRs and values of *C*. The results demonstrate the absolute superiority of the recognition models based on multi-label classification in the problem of compound jamming signal recognition with a large number of candidate jamming signal components.

Table 3. Recognition performance at different numbers of candidate jamming signal components.

JNR (dB)	ML-CR-CV-CNN			ML-CV-CNN			MLAMC			1D-CNN		
	C = 2	*C* = 3	*C* = 4	*C* = 2	*C* = 3	*C* = 4	*C* = 2	*C* = 3	*C* = 4	*C* = 2	*C* = 3	*C* = 4
5	1	1	0.95	1	1	0.917	1	1	0.92	1	0.983	0.761
10	1	1	0.97	1	1	0.957	1	1	0.96	1	0.933	0.787
15	1	1	0.98	1	1	0.963	1	1	0.967	1	0.995	0.836
25	1	1	0.98	1	1	0.97	1	1	0.967	1	1	0.859

4.4.4. Recognition Performance Versus JNRs

Generally, the change in JNRs can affect not only the convergence performance of recognition models, but also the recognition performance of compound jamming signals. In this experiment, the robustness of the proposed recognition model to varying JNRs is validated using the compound jamming signal datasets with different JNRs.

Figure 9 shows the subset accuracy with varying JNRs for various recognition models. It can be found that the overall performance of the ML-CR-CV-CNN model is optimal at all JNRs. Compared with the 1D-CNN model, the performance is improved by an average

of more than 15% at all JNRs. Compared with the ML-CV-CNN and MLAMC models, the subset accuracy of the ML-CR-CV-CNN model has greater improvement in the case of low JNRs and similar values in the case of high JNRs. Furthermore, observing the trends of curves, the subset accuracy of all models increases with increasing JNRs. At JNR > 10 dB, the accuracy of three models based on multi-label classification reaches stable values of more than 95%. Especially, for the ML-CR-CV-CNN model, it also reaches a satisfactory 85% at JNR = 0 dB, which is better than other baseline methods. The proposed method effectively alleviates the performance decline caused by noise at low JNRs, and has stronger robustness to varying JNRs. Except for the accuracy, the proportion of the labels being misclassified is measured by the Hamming loss for the three models based on multi-label classification. The results with varying JNRs are shown in Figure 10. Contrary to the subset accuracy, the overall trend gradually decreases with increasing JNRs, and the ML-CR-CV-CNN model has the lowest classification loss.
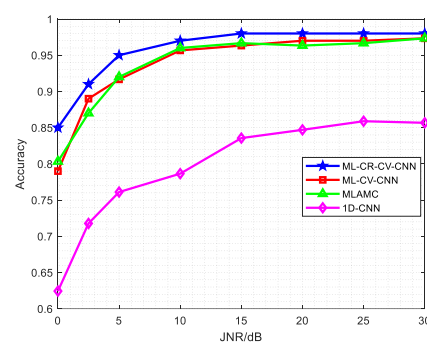

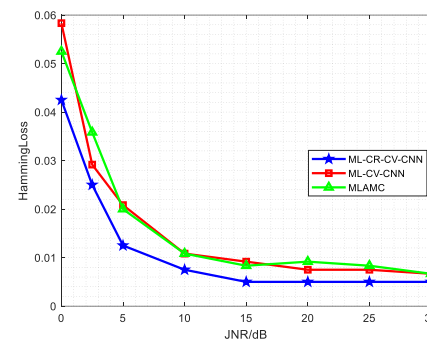**Figure 9.** Subset accuracy at different JNRs.

Figure 10. Hamming loss at different JNRs.

The 1D-CNN model based on multi-class classification can only recognize each combination of jamming signal components as a single class. There is no case where the predicted result partially matches the true label. However, in practice, the correct recognition of partial jamming signal components can also bring good gain for the subsequent countermeasure of compound jamming signals. The output results of the multi-label recognition models can be used not only to calculate the subset accuracy to evaluate the overall recognition performance in the case of complete matching, but also to calculate the partial accuracy and label accuracy to evaluate the fine-grained recognition performance of partial matching and specific jamming signal classes, respectively.

As seen in Figure 9, the subset accuracy declines significantly at JNR < 10 dB. At low JNRs, the label accuracy of various recognition models is displayed in Table 4. It can be found that the recognition performance of class2 and class4 is basically unaffected by JNRs, and the decline of the subset accuracy is mainly caused by the reduction in the recognition performance of class1 and class3. Compared with the ML-CV-CNN and MLAMC models, the ML-CR-CV-CNN model effectively improves the label accuracy of class1 and class3, and it has the best performance at almost all JNRs for different classes

of jamming signals. Meanwhile, the subset accuracy is also optimal. Furthermore, for the ML-CR-CV-CNN model with optimal performance, the partial accuracy at different JNRs is shown in Figure 11. Partialacc1, Partialacc2, Partialacc3, and Partialacc4 represent the proportion of at least 1–4 jamming signal components that are accurately recognized. For all JNRs, both Partialacc1 and Partialacc2 with the value of 1 indicate that at least two components in compound jamming signals can be correctly recognized. When JNR > 5 dB, the minimum number of correctly recognized components increases to 3.

Table 4. Label accuracy of the multi-label recognition models at low JNRs.

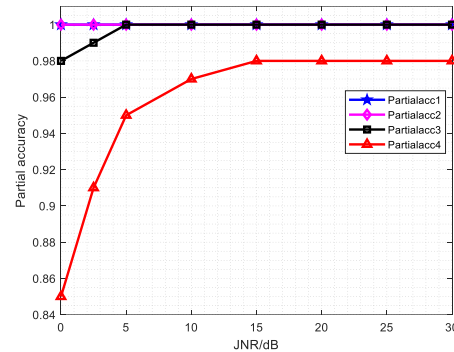| JNR (dB)	ML-CR-CV-CNN				ML-CV-CNN				MLAMC			
	class1	class2	class3	class4	class1	class2	class3	class4	class1	class2	class3	class4
0	0.92	1	0.91	1	0.873	1	0.893	1	0.877	1	0.9137	1
2.5	0.93	1	0.97	1	0.93	1	0.953	1	0.9137	1	0.947	0.997
5	0.96	1	0.99	1	0.943	1	0.973	1	0.937	1	0.9837	1
10	0.97	1	1	1	0.96	1	0.997	1	0.96	1	0.997	1

Figure 11. Partial accuracy of the ML-CR-CV-CNN model at different JNRs.

4.4.5. Recognition Performance Versus PRs

All the above experiments are carried out under the setting of PR = 0 dB. The power of each component in compound jamming signals is equal. Usually, in order to effectively achieve the jamming effect, the power of suppression jamming signals is higher than the deception jamming signals according to the generation mechanism. Therefore, considering the different energy losses caused by unequal propagation distances and transmitting power of different jammers, the intensities of the jamming signal components in compound jamming signals received by the radar are usually unequal. And the power imbalance causes fluctuations in recognition performance. In this experiment, the influence of different PRs on the overall recognition performance is verified by changing the power of one jamming signal component. The JNR is fixed at 15 dB, and the power of the noise frequency modulation jamming signal is adjusted.

Figure 12 shows the subset accuracy with varying PRs for various recognition models. It can be seen that the three models based on multi-label classification can obtain better performance at multiple PRs compared with the 1D-CNN model. The accuracy of the ML-CV-CNN and ML-CR-CV-CNN models is more robust to the variation of PRs than that of the MLAMC model. When the power of each component in compound jamming signals is equal (PR = 0 dB), the subset accuracy reaches the maximum. Using the result at PR = 0 dB as the baseline, the accuracy decreases to varying degrees with decreasing or increasing PRs. At PR > 0 dB, the degree of decline is much greater than the results at PR < 0 dB. Especially for the MLAMC model, the accuracy declines rapidly at PR > 10 dB due to the strong coverage effect on time-frequency images caused by high-power suppression jamming signals. In contrast, the ML-CV-CNN and ML-CR-CV-CNN models using one-dimensional complex data show stronger ability in mining strong overlapping signal features when the

high-power suppression jamming signal exists, and the accuracy is effectively improved at PR > 10 dB. In addition, benefiting from the fusion of jamming class representations, the performance of the ML-CR-CV-CNN model is slighter better than that of the ML-CV-CNN model at PR > 0 dB, and the decline caused by the strong overlapping effect of high-power suppression jamming signals is further alleviated. At PR = 15 dB, the accuracy of the ML-CR-CV-CNN model is close to 88%, which is 52% higher than the MLAMC model.
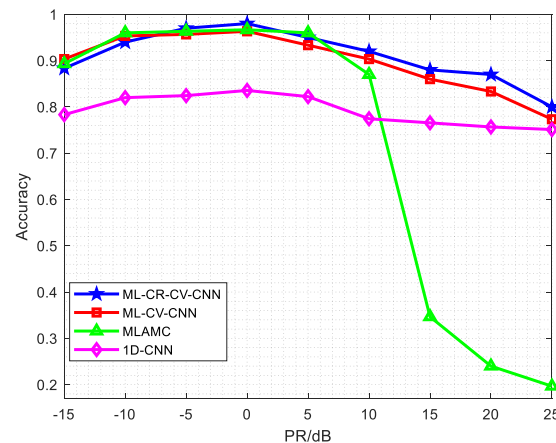

Figure 12. Recognition performance at different PRs.

4.4.6. Recognition Performance of Unseen Jamming Signal Combinations in Training

In the presence of many candidate jamming signal components, the size of the compound jamming signal dataset containing all possible combinations is large. In order to reduce the difficulty of sample acquisition and improve the training speed, it is expected that only partial samples in the dataset are used to train recognition models, and good recognition results can be achieved for all combinations simultaneously. For the jamming signal combinations that do not appear in training, the correct recognition of jamming signal components requires that the recognition models have some generalization and extensibility. In this experiment, three training sets containing partial combinations are used to evaluate the recognition performance of the trained models against unseen jamming signal combinations. For the three datasets, the JNR and PR are fixed at 15 dB and 0 dB, and the compositions of jamming signal components are $k = 1$, $k = \{1, 2\}$, and $k = \{1, 2, 3\}$. Thus, the corresponding unseen jamming signal combinations to be tested consist of $k = \{2, 3, 4\}$, $k = \{3, 4\}$, and $k = 4$.

The 1D-CNN model based on multi-class classification does not have the ability to recognize jamming signal combinations that have not been seen in training. Figure 13 shows the subset accuracy of unseen jamming signal combinations when the ML-CR-CV-CNN, ML-CV-CNN, and MLAMC models use partial combinations for model training. The black dashed line represents the average subset accuracy for all unseen combinations containing different possible values of k. It can be observed that the recognition performance of the three models has a similar trend with various training sets. With the enrichment of jamming signal combinations in training sets, the accuracy of unseen combinations increases. When the training set consists of only single jamming signals ($k = 1$), the ML-CR-CV-CNN model has the best recognition result by comparing the three models horizontally. Especially for the test samples with $k = 2$ and $k = 3$, the accuracy is significantly higher than that of the ML-CV-CNN and MLAMC models. In this case, only 26.7% of jamming signal combinations participate in model training for the dataset containing all possible jamming signal combinations. The ML-CR-CV-CNN model still has an average subset accuracy of 74.36% for the testing samples that have not been seen in training, which is 21% higher than the ML-CV-CNN model and 27% higher than the MLAMC model. The results demonstrate that the fusion of jamming class representations in the ML-CR-CV-CNN greatly enhances the generalization and extensibility for unseen jamming signal combinations.
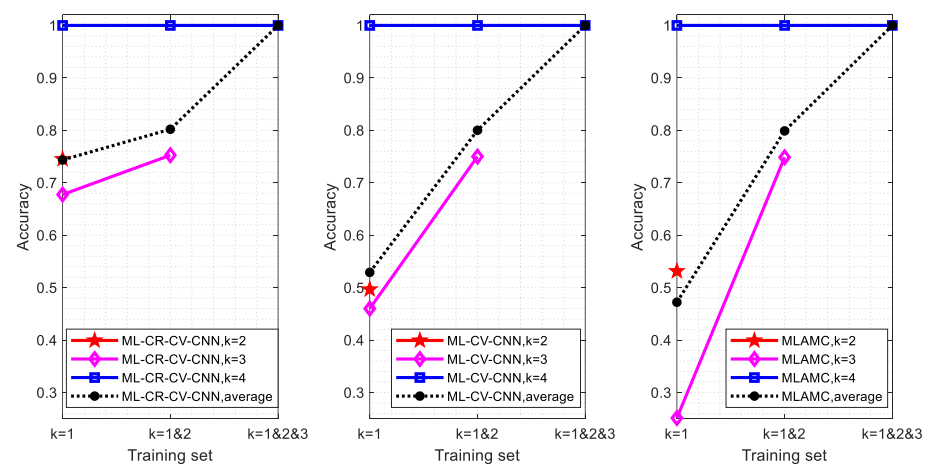
Figure 13. Recognition performance of unseen jamming signal combinations in training. From left to right are the ML-CR-CV-CNN, ML-CV-CNN, and MLAMC.

5. Conclusions

In this paper, we propose a novel multi-label classification framework based on the ML-CR-CV-CNN combining the CV-CNN and jamming class representations to solve the compound jamming signal recognition problem. In order to reduce the additional computational cost caused by the preprocessing in existing methods, the ML-CV-CNN is designed to directly process one-dimensional time–domain complex jamming signals, which alleviates information loss and enhances the feature extraction ability for strong overlapping signals. Based on the basic ML-CV-CNN, the jamming class representations obtained by the LVQ are further fused into the ML-CR-CV-CNN model by class decoupling to guide compound jamming signals to focus on the class-related features, thereby improving the extensibility and generalization ability for unseen jamming signal combinations. Moreover, the adaptive threshold calibration is used to optimize the final recognition results. Experimental results prove that the proposed method is robust to varying JNRs and PRs. Especially, when the high-power suppression jamming signal with strong overlapping effect exists, the recognition performance in the case of high PRs and model convergence speed in the case of high JNRs are significantly improved. Meanwhile, the applicability of the proposed recognition method is effectively extended in the face of unseen jamming signal combinations in training when there are fewer training samples.

Author Contributions: Conceptualization, Y.M., L.Y. and Y.W.; methodology, Y.M. and L.Y.; writing original draft preparation, Y.M.; review and editing, Y.M., L.Y. and Y.W. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the fact that it is currently privileged information.

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

References

1. Tian, X. Radar Deceptive Jamming Detection Based on Goodness-of-fit Testing. *J. Inf. Comput. Sci.* **2012**, *9*, 3839–3847.
2. Zhao, S.; Zhou, Y.; Zhang, L.; Guo, Y.; Tang, S. Discrimination between radar targets and deception jamming in distributed multiple-radar architectures. *IET Radar Sonar Navig.* **2017**, *11*, 1124–1131. [CrossRef]
3. Hao, Z.; Yu, W.; Chen, W. Recognition method of dense false targets jamming based on time-frequency atomic decomposition. *J. Eng.* **2019**, *20*, 6354–6358. [CrossRef]
4. Du, C.; Zhao, Y.; Wang, L.; Tang, B.; Xiong, Y. Deceptive multiple false targets jamming recognition for linear frequency modulation radars. *J. Eng.* **2019**, *21*, 7690–7694. [CrossRef]

5. Wang, N.; Wang, Y.; Liu, H.; Zuo, Q. Target discrimination method for SAR images via convolutional neural network with semi-supervised learning and minimum feature divergence constraint. *Remote Sens. Lett.* **2020**, *11*, 1167–1174. [CrossRef]
6. Zeng, Z.; Sun, J.; Han, Z.; Hong, W. SAR Automatic Target Recognition Method Based on Multi-Stream Complex-Valued Networks. *IEEE Trans. Geosci. Remote Sensing* **2022**, *60*, 1–18. [CrossRef]
7. Lin, R.; Ren, W.; Sun, X.; Yang, Z.; Fu, K. A Hybrid Neural Network for Fast Automatic Modulation Classification. *IEEE Access* **2020**, *8*, 130314–130322. [CrossRef]
8. Bai, J.; Yao, J.; Qi, J.; Wang, L. Electromagnetic Modulation Signal Classification Using Dual-Modal Feature Fusion CNN. *Entropy* **2022**, *24*, 700. [CrossRef]
9. Luo, J.; Si, W.; Deng, Z. Few-Shot Learning for Radar Signal Recognition Based on Tensor Imprint and Re-Parameterization Multi-Channel Multi-Branch Model. *IEEE Signal Process. Lett.* **2022**, *29*, 1327–1331. [CrossRef]
10. Tian, L.; Chen, B.; Guo, Z.; Du, C.; Peng, Y.; Liu, H. Open set HRRP recognition with few samples based on multi-modality prototypical networks. *Signal Process.* **2022**, *193*, 108391. [CrossRef]
11. Wang, Q.; Du, P.; Yang, J.; Wang, G. Transferred deep learning based waveform recognition for cognitive passive radar. *Signal Process.* **2019**, *155*, 259–267. [CrossRef]
12. Hoang, L.M.; Kim, M.; Kong, S.H. Automatic Recognition of General LPI Radar Waveform Using SSD and Supplementary Classifier. *IEEE Trans. Signal Process.* **2019**, *67*, 3516–3530. [CrossRef]
13. Wang, Y.; Sun, B.; Wang, N. Recognition of radar active-jamming through convolutional neural networks. *J. Eng.* **2019**, *21*, 7695–7697. [CrossRef]
14. Wang, F.; Huang, S.; Liang, C. Automatic Jamming Modulation Classification Exploiting Convolutional Neural Network for Cognitive Radar. *Math. Probl. Eng.* **2020**, *2020*, 9148096. [CrossRef]
15. Lang, B.; Gong, J. JR-TFViT: A Lightweight Efficient Radar Jamming Recognition Network Based on Global Representation of the Time–Frequency Domain. *Electronics* **2022**, *11*, 2794. [CrossRef]
16. Zhang, H.; Yu, L.; Chen, Y.; Wei, Y. Fast Complex-Valued CNN for Radar Jamming Signal Recognition. *Remote Sens.* **2021**, *13*, 2867. [CrossRef]
17. Shao, G.; Chen, Y.; Wei, Y. Deep Fusion for Radar Jamming Signal Classification Based on CNN. *IEEE Access* **2020**, *8*, 117236–117244. [CrossRef]
18. Michael, M.; Oner, M.; Dobre, O.; Jakel, H.; Jondral, F. Automatic Modulation Classification for MIMO Systems Using Fourth-Order Cumulants. In Proceedings of the 2012 IEEE Vehicular Technology Conference (VTC Fall), Quebec City, QC, Cananda, 3–6 September 2012.
19. Huang, S.; Yao, Y.; Wei, Z.; Feng, Z.; Zhang, P. Automatic Modulation Classification of Overlapped Sources Using Multiple Cumulants. *IEEE Trans. Veh. Technol.* **2017**, *66*, 6089–6101. [CrossRef]
20. Huang, S.; Yao, Y.; Yan, X.; Feng, Z. Cumulant based Maximum Likelihood Classification for Overlapped Signals. *Electron. Lett.* **2016**, *52*, 1761–1763. [CrossRef]
21. Liu, Z.; Li, L.; Xu, H.; Li, H. A method for recognition and classification for hybrid signals based on Deep Convolutional Neural Network. In Proceedings of the 2018 International Conference on Electronics Technology (ICET), Chengdu, China, 23–27 May 2018.
22. Zhang, J.; Liang, Z.; Zhou, C.; Liu, Q.; Long, T. Radar Compound Jamming Cognition Based on a Deep Object Detection Network. *IEEE Trans. Aerosp. Electron. Syst.* **2023**, *59*, 3251–3263. [CrossRef]
23. Tang, H.; Liu, M.; Chen, L.; Li, J. Time-frequency Overlapped Signals Intelligent Modulation Recognition in Underlay CRN. In Proceedings of the 2020 IEEE/CIC International Conference on Communications in China (ICCC), Foshan, China, 9–11 August 2020.
24. Ren, Y.; Huo, W.; Pei, J.; Huang, Y.; Yang, J. Automatic Modulation Recognition for Overlapping Radar Signals based on Multi-Domain SE-ResNeXt. In Proceedings of the 2021 IEEE Radar Conference (RadarConf21), Atlanta, GA, USA, 7–14 May 2021.
25. Shao, G.; Chen, Y.; Wei, Y. Convolutional neural network-based radar jamming signal classification with sufficient and limited samples. *IEEE Access* **2020**, *8*, 80588–80598. [CrossRef]
26. Qu, Q.; Wei, S.; Liu, S.; Liang, J.; Shi, J. JRNet: Jamming Recognition Networks for Radar Compound Suppression Jamming Signals. *IEEE Trans. Veh. Technol.* **2020**, *69*, 15035–15045. [CrossRef]
27. Zhang, M.; Zhou, Z. A Review on Multi-Label Learning Algorithms. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1819–1837. [CrossRef]
28. Liu, W.; Wang, H.; Shen, X.; Tsang, I.W. The Emerging Trends of Multi-Label Learning. *IEEE Trans. Pattern Anal.* **2021**, *44*, 7955–7974. [CrossRef]
29. Qi, G.; Hua, X.; Rui, Y.; Tang, J.; Mei, T.; Wang, M.; Zhang, H. Correlative multi-label video annotation with temporal kernels. *ACM Trans. Multimed. Comput. Commun. Appl.* **2008**, *5*, 3. [CrossRef]
30. Zhang, X.; Li, H.; Shi, C.; Li, P. Multi-Instance Multi-Label Action Recognition and Localization Based on Spatio-Temporal Pre-Trimming for Untrimmed Videos. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020.
31. Kim, D.Y.; Vo, B.N.; Vo, B.T. Online Visual Multi-Object Tracking via Labeled Random Finite Set Filtering. *arXiv* **2016**, arXiv:1611.06011.
32. Ciarelli, P.M.; Oliveira, E.; Salles, E.O.T. Multi-label incremental learning applied to web page categorization. *Neural Comput. Appl.* **2014**, *24*, 1403–1419. [CrossRef]

33. Adavanne, S.; Politis, A.; Nikunen, J.; Virtanen, T. Sound Event Localization and Detection of Overlapping Sources Using Convolutional Recurrent Neural Networks. *IEEE J. Sel. Top. Signal Process.* **2018**, *13*, 34–48. [CrossRef]

34. Chen, T.; Xu, M.; Hui, X.; Wu, H.; Lin, L. Learning Semantics-Specific Graph Representation for Multi-Label Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, 16–20 June 2019.

35. Ye, J.; He, J.; Peng, X.; Wu, W.; Qiao, Y. Attention-driven dynamic graph convolutional network for multi-label image recognition. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Springer: Berlin/Heidelberg, Germany, 2020.

36. Pan, Z.; Wang, S.; Zhu, M.; Li, Y. Automatic Waveform Recognition of Overlapping LPI Radar Signals Based on Multi-Instance Multi-Label Learning. *IEEE Signal Process. Lett.* **2020**, *27*, 1275–1279. [CrossRef]

37. Pan, Z.; Wang, S.; Li, Y. Residual Attention-Aided U-Net GAN and Multi-Instance Multilabel Classifier forAutomatic Waveform Recognition of Overlapping LPI Radar Signals. *IEEE Trans. Aerosp. Electron. Syst.* **2022**, *58*, 4377–4395. [CrossRef]

38. Zhu, M.; Li, Y.; Pan, Z.; Yang, J. Automatic modulation recognition of compound signals using a deep multi-label classifier: A case study with radar jamming signals. *Signal Process.* **2020**, *169*, 107393. [CrossRef]

39. Trabelsi, C.; Bilaniuk, O.; Zhang, Y.; Serdyuk, D.; Subramanian, S.; Santos, J.F.; Mehri, S.; Rostamzadeh, N.; Bengio, Y.; Pal, C.J. Deep Complex Networks. In Proceedings of the Sixth International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018.

40. Zhao, Z.; Vuran, M.C.; Guo, F.; Scott, S. Deep-Waveform: A Learned OFDM Receiver Based on Deep Complex Convolutional Networks. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 2407–2420. [CrossRef]

41. Maaten, L.V.D.; Hinton, G. Visualizing data using t-sne. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.