

Article

# Specific Emitter Identification through Multi-Domain Mixed Kernel Canonical Correlation Analysis

Jian Chen, Shengyong Li, Jianchi Qi and Hongke Li \*

Department of Electronic Technology, Naval University of Engineering, Wuhan 430033, China; chen\_j\_158@163.com (J.C.); lsy929@163.com (S.L.); cheeserabbit0214@163.com (J.Q.)

\* Correspondence: dotagent@163.com

**Abstract:** Radar specific emitter identification (SEI) involves extracting distinct fingerprints from radar signals to precisely attribute them to corresponding radar transmitters. In view of the limited characterization of fingerprint information by single-domain features, this paper proposes the utilization of multi-domain mixed kernel canonical correlation analysis for radar SEI. Initially, leveraging the complementarity across diverse feature domains, fingerprint features are extracted from four distinct domains including: envelope feature, spectrum feature, short-time Fourier transform and ambiguity function. Subsequently, kernel canonical correlation analysis is employed to amalgamate the correlation characteristics inherent in multi-domain data. Considering the insufficiency of a single kernel function with only interpolation or extrapolation ability, we adopt mixed kernel to improve the projection ability of the kernel function. Experimental results substantiate that the proposed feature fusion approach maximizes the complementarity of multiple features while reducing feature dimensionality. The method achieves an accuracy of up to 95% in experiments, thereby enhancing the efficacy of radar SEI.

**Keywords:** specific emitter identification; feature fusion; feature extraction; kernel canonical correlation analysis



**Citation:** Chen, J.; Li, S.; Qi, J.; Li, H. Specific Emitter Identification through Multi-Domain Mixed Kernel Canonical Correlation Analysis. *Electronics* **2024**, *13*, 1173. <https://doi.org/10.3390/electronics13071173>

Academic Editors: Yin Zhang, Deqing Mao, Yulin Huang and Yachao Li

Received: 24 January 2024

Revised: 19 March 2024

Accepted: 20 March 2024

Published: 22 March 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Specific emitter identification (SEI) by a radar involves identifying individual radars by analyzing the distinctive features or fingerprints embedded in their signals [1]. This intricate procedure comprises key stages such as data pre-processing, feature extraction, classification, and identification. The correlation of radar signals with specific radars and their associated platforms is pivotal for target intent analysis, decision support, and situational awareness [2]. Consequently, radar SEI has attracted considerable attention in the fields of electronic reconnaissance and electronic countermeasures [3].

The key component for individual radar emitter identification is the fingerprint feature of the radar transmitter, characterized by stability and uniqueness [4,5]. Fingerprint features persist in radar signals, resisting complete erasure due to their accidental modulation features arising from minute changes in radar technology [3,6]. Each radar must be assigned unique labels, considering that even radars of the same model may be affected by minor faults in electronics, operating time, and environmental conditions [7]. Extracting fingerprint features encoded in received radar signals poses a substantial challenge, particularly when these features are immersed in strong noise signals. To date, research on radar fingerprint feature extraction has predominantly focused on the time domain [8], frequency domain [9,10], time-frequency domain [11–13], ambiguity-function domain [14], and neural network domains [15–17] of the signals. While experiments have demonstrated satisfactory recognition results using the aforementioned methods, the subtle distinctions between fingerprint characteristics may be overlooked by the discussed extraction strategies.

Relying solely on a single feature in radar SEI leads to a decline in recognition accuracy when confronted with diverse radars and varying signal backgrounds. In recent years,

multimodal fusion technology has garnered considerable attention, demonstrating remarkable achievements across multiple domains. Researchers have developed the multimodal approach to radar SEI, employing the feature-level fusion method within diverse feature domains. Compared with the previously employed single-feature method, the feature fusion method leverages the differences between multiple features to preserve fingerprint information [18–20]. In one study [18], an innovative parallel feature fusion technique was developed based on the investigation of a simple series-parallel relationship between different features. However, drawbacks such as the absence of a nonlinear description of uncorrelated data and high feature dimensionality were noted. Another study [19] employed a local kernel approach based on the widely utilized dimensionality reduction method principal component analysis (PCA), enhancing dimensionality reduction efficiency. Nevertheless, PCA-based algorithms fail to harness non-linear correlations among several features since they do not consider the inherent links between features. With the rapid evolution of neural networks, deep learning has been extensively used in SEI. The multi-channel approach, capitalizing on the substantial scalability of neural networks for data fusion, is well-conceived and performs admirably in SEI with multi-feature fusion. In a pertinent study [20], a multi-channel deep learning model was employed to autonomously learn to fuse multiple features and thoroughly extract fingerprint feature information. However, the training of deep learning models often demands a substantial number of training samples. In reality, acquiring a significant quantity of actual detected radar signals proves challenging, posing an incongruity with the application of data-driven methodologies. Thus, multi-feature fusion must satisfy three essential requirements: (1) the ability to explore relationships among different features, (2) the extraction of effective feature information, and (3) the reduction of feature dimensionality.

Canonical correlation analysis (CCA) serves as a linear multivariate statistical method for examining the correlation between two sets of features, identifying canonical eigenvectors with higher correlation indicative of original features. CCA encounters challenges when extracting a valid representation from data that does not adhere to a linear distribution. Therefore, to adapt to the non-linear data characteristics, the kernel method based on CCA can be employed. The kernel function facilitates the projection of fingerprint features into a feature space, adept at handling non-linear data and extracting weak fingerprint information with greater ease. The choice of kernel function in kernel methods leads to distinct mapping spaces and diverse ways of describing the data. Jia [21] classified kernel functions into two main types: local kernels and global kernels. However, both types exhibit only one type of interpolation and extrapolation capabilities. Presently, the kernel function in radar SEI predominantly relies on a single kernel. To overcome the limitations of a single kernel, researchers explore the linear combination of kernel functions, such as multiple kernel learning (MKL) [22] and mixed kernel [23,24] approaches. In the realm of multi-kernel mapping, the high-dimensional space amalgamates multiple feature spaces. Each basic kernel optimally leverages its ability to map various features within the combinatorial space. Authors in a study [22] employed MKL to select the most appropriate base kernel at each data point, determining the optimal kernel through linear weighting of the base kernels. However, the experimental choice of the base kernel and the combination method lacks theoretical grounding and remains uncertain. In another study [23,24], addressing the absence of a single kernel with only one interpolation or extrapolation ability, authors proposed a mixed kernel by weighting the radial basis function (RBF) kernel with interpolation ability and the polynomial (poly) kernel with extrapolation ability, respectively. Essentially, mixed kernel models fall into the category of single kernel models but adopt the form of MKL. This differentiation arises from the assignment of independent weights to each kernel function in the mixed kernel, circumventing the need for solving complex optimization problems present in MKL. Mixed kernels exhibit both a robust theoretical foundation and a more convenient combination method while circumventing the necessity to learn a large number of base kernel weights.

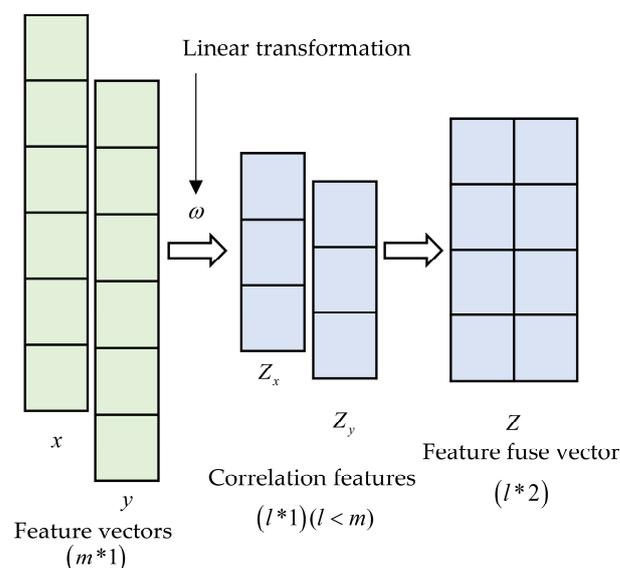
Building on the literature, this paper proposes a multi-domain mixed kernel canonical correlation analysis (MMKCCA) for radar SEI. The primary focus involves initially extracting fingerprint features in four feature domains from the radar signal with noise removed. The mixed kernel, better suited to the characteristics of multi-feature data, is then applied as the kernel function for fingerprint feature fusion using the kernel canonical correlation analysis (KCCA) technique. Additionally, the fused features are fed into the random forest classifier for classification and recognition. Experimental results indicate that the proposed method yields superior recognition outcomes, with accuracy reaching 95% under lower feature dimensions. The main contributions of this paper are threefold: (1) introducing a multi-domain feature fusion method for radar SEI based on KCCA, affirming the complementarity of different feature domains; (2) addressing the limitations of local and global kernels by proposing a mixed kernel that combines the two in a weighted composition, better adapting to the characteristics of data with multi-domain features; and (3) efficiently lowering feature dimensions while maintaining SEI recognition performance with a modest number of samples.

The remainder of this paper is organized as follows: Section 2 outlines the fundamentals of CCA; Section 3 discusses kernel function principles and selection; Section 4 describes experiments using the dataset and the aforementioned theoretical study; and Section 5 summarizes the findings of the experiments and proposes directions for future research.

## 2. Analysis of Canonical Correlation Analysis (CCA)

The numerous features obtained for the same radar signal exhibit some degree of interrelation, presenting an opportunity to fully exploit the complementing effect between these features for effective feature fusion. This work employs CCA, a technique with multivariate statistical analysis capabilities adept at uncovering subtle variations and inter-correlations among distinct features. The principles and methods of CCA will be briefly introduced below.

Hotelling introduced CCA in 1936, essentially involving finding the feature vector with the highest correlation rather than the original features. The method utilizes the correlation between features as the discriminant criterion, achieving both the reduction of original features to eliminate information redundancy and the purpose of feature fusion [25]. CCA can be seen as the problem of finding basis vectors for two sets of variables such that the correlations between the projections of the variables onto these basis vectors are mutually maximized. The simplified flow of CCA is illustrated in Figure 1:



**Figure 1.** Simplified diagram of CCA.

The symbols used by CCA and KCCA are shown in Table 1.

**Table 1.** Notations.

$\mathbb{R}$ : metric spaces	$\langle \cdot, \cdot \rangle$ : inner product
$x$ : feature vector	$L$ : Lagrange function
$\omega_x$ : the linear transformation of $x$	$K$ : kernel matrix
$\omega_y$ : the linear transformation of $y$	$\phi$ : a map into Hilbert spaces
$\rho$ : correlation coefficient	$\lambda$ : correlations
$C_{xx}$ : covariance metric	$Z$ : fusion feature vector

The two types of fingerprint feature vectors can be represented as  $x \in \mathbb{R}^p$  and  $y \in \mathbb{R}^q$ . Canonical correlation analysis seeks a pair of linear transformation  $\omega_x$  and  $\omega_y$ , one for each of the sets of feature vectors  $x$  and  $y$ , such that when the set of vectors is transformed, the corresponding vectors  $Z_x$  and  $Z_y$  are maximally correlated.  $\rho$  represents the correlation function.

The first stage of canonical correlation is to choose  $\omega_x$  and  $\omega_y$  to maximize the correlation  $\rho$  between the two feature vectors  $x$  and  $y$  [26]. The correlation function can be expressed as follows:

$$\rho = \max_{\omega_x, \omega_y} \frac{\omega_x^T C_{xy} \omega_y}{\sqrt{\omega_x^T C_{xx} \omega_x \omega_y^T C_{yy} \omega_y}} \tag{1}$$

where  $C_{xx}$  and  $C_{yy}$  denote the covariance matrices of the two types of fingerprint features, respectively, and  $C_{xy}$  represents the mutual covariance matrix between them.

To ensure a unique solution, the following constraints are applied:

$$\begin{cases} \omega_x^T C_{xx} \omega_x = 1 \\ \omega_y^T C_{yy} \omega_y = 1 \end{cases} \tag{2}$$

In order to get a linear transformation  $\omega_x$  and  $\omega_y$ . The Lagrange criterion functions is constructed by combining correlation function  $\rho$  and condition function Equation (2) [27]:

$$L(\lambda, \omega_x, \omega_y) = \omega_x^T C_{xy} \omega_y - \frac{\lambda_x}{2} (\omega_x^T C_{xx} \omega_x - 1) - \frac{\lambda_y}{2} (\omega_y^T C_{yy} \omega_y - 1) \tag{3}$$

Taking derivatives with respect to  $\omega_x$  and  $\omega_y$  and setting them to zero yields:

$$\frac{\partial L}{\partial \omega_x} = C_{xy} \omega_y - \lambda_x C_{xx} \omega_x = 0 \tag{4}$$

$$\frac{\partial L}{\partial \omega_y} = C_{yx} \omega_x - \lambda_y C_{yy} \omega_y = 0 \tag{5}$$

By multiplying Equation (4) by  $\omega_x^T$  and subtracting  $\omega_y^T$  from Equation (5),

$$\begin{aligned} 0 &= \omega_x^T C_{xy} \omega_y - \omega_x^T \lambda_x C_{xx} \omega_x - \omega_y^T C_{yx} \omega_x + \omega_y^T \lambda_y C_{yy} \omega_y \\ &= \lambda_y \omega_y^T C_{yy} \omega_y - \lambda_x \omega_x^T C_{xx} \omega_x \end{aligned} \tag{6}$$

Based on the constraints in Equation (2),  $\lambda_y - \lambda_x = 0$ , and  $\lambda = \lambda_y = \lambda_x$ . Assuming that the covariance matrix  $C_{yy}$  is singular:

$$\omega_y = \frac{C_{yy}^{-1} C_{yx} \omega_x}{\lambda} \tag{7}$$

Substituting  $\omega_y$  from Equation (7) to Equation (4),

$$C_{xy}C_{yy}^{-1}C_{yx}\omega_x = \lambda^2C_{xx}\omega_x \tag{8}$$

The projection vector of the features  $\omega_x$  can be found using the eigenvalue Equation (8). Substituting  $\omega_x$  into Equation (7) yields the projection vector  $\omega_y$ . Thus, the typical correlation features after projection can be obtained as  $Z_x = \omega_x^T x$  and  $Z_y = \omega_y^T y$ , and their combination yields the final feature fusion  $Z = [Z_x, Z_y]$ .

In situations where the original feature data deviates from Gaussian or linear distribution, effective information extraction from linearly operated CCA becomes challenging. Therefore, CCA is extended to nonlinear CCA to better handle situations where the relationship between different features is nonlinearly distributed, yielding effective features. In an attempt to increase the flexibility of the feature selection, kernelization of CCA (KCCA) has been applied to map the hypotheses to a higher-dimensional feature space. The subsequent section provides a detailed introduction to nonlinear CCA with the kernel function.

### 3. Kernel Methods

#### 3.1. Kernel CCA

Given that CCA operates linearly, it encounters limitations in effectively extracting nonlinear data features. KCCA introduces an innovative approach that employs a kernel function to nonlinearly extend the original fingerprint characteristics, projecting them into a high-dimensional feature space [27–29]. This method not only accommodates nonlinear data, transforming a nonlinear problem into a linear one, but also facilitates enhanced access to fine-grained fingerprint information. The underlying principle of KCCA is concisely described below, with a visual representation provided in Figure 2.

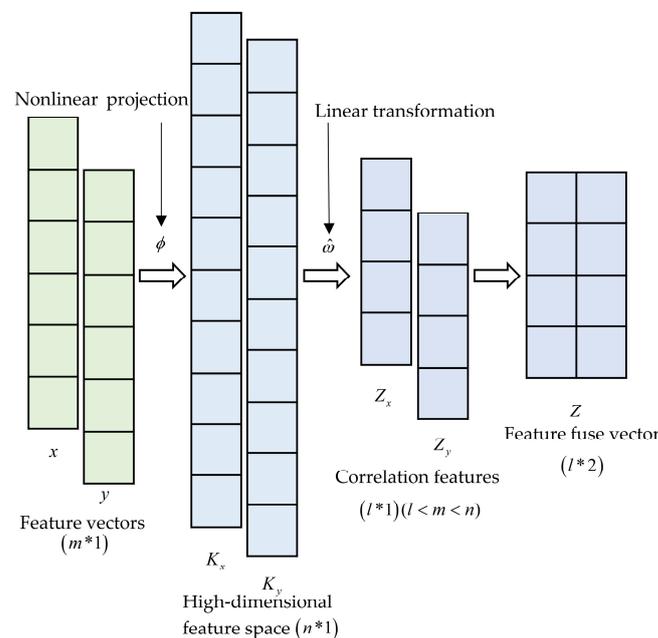


Figure 2. Simplified diagram of KCCA.

For the original fingerprint feature vectors  $x \in \mathbb{R}^p$  and  $y \in \mathbb{R}^q$ , high-dimensional feature vectors are obtained through the following kernel nonlinear transformation:

$$\phi : x = (x_1, \dots, x_m) \mapsto \phi(x) = (\phi_1(x), \dots, \phi_N(x))(m < N) \tag{9}$$

where  $\phi$  signifies the mapping of the original feature vector  $x$  to the high-dimensional feature space, and  $\phi(y)$  follows a similar process. Kernels are methods of implicitly mapping

data into a higher-dimensional feature space. The kernel function  $K(x_i, x_j)$  operation can be expressed as:

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \tag{10}$$

Using the definition of the covariance matrix in Equation (1), we can rewrite the covariance matrix  $C_{xx}$  and  $C_{yy}$ .

$$\begin{cases} C_{xx} = x'x \\ C_{yy} = y'y \end{cases} \tag{11}$$

where we use  $x'$  to denote the transpose of a vector  $x$ .

The linear transformation  $\omega_x$  and  $\omega_y$  can be rewritten as the projection of the feature vectors onto the transformation  $\hat{\omega}_x$  and  $\hat{\omega}_y$ :

$$\begin{cases} \hat{\omega}_x = x'\omega_x \\ \hat{\omega}_y = y'\omega_y \end{cases} \tag{12}$$

Substituting into Equation (1), the correlation function can be expressed as follows:

$$\rho = \max_{\hat{\omega}_x, \hat{\omega}_y} \frac{\hat{\omega}'_x x x' y y' \hat{\omega}_y}{\sqrt{\hat{\omega}'_x x x' x x' \hat{\omega}_x \hat{\omega}'_y y y' y y' \hat{\omega}_y}} \tag{13}$$

Let  $K_x = x x'$  and  $K_y = y y'$ . The maximization criterion function is reformulated as:

$$\rho = \max_{\hat{\omega}_x, \hat{\omega}_y} \frac{\hat{\omega}'_x K_x K_y \hat{\omega}_y}{\sqrt{\hat{\omega}'_x K_x K_x \hat{\omega}_x \hat{\omega}'_y K_y K_y \hat{\omega}_y}} \tag{14}$$

Once again, this criterion function must adhere to the following constraint:

$$\begin{cases} \hat{\omega}'_x K_x \hat{\omega}_x = 1 \\ \hat{\omega}'_y K_y \hat{\omega}_y = 1 \end{cases} \tag{15}$$

The subsequent computation aligns with the standard CCA procedures. Derive the projection vectors  $\hat{\omega}_x$  and  $\hat{\omega}_y$  to obtain typical correlation features  $Z_x = K_x \hat{\omega}_x$  and  $Z_y = K_y \hat{\omega}_y$ .

The resolution of nonlinear relationships between features is facilitated by projecting features into a higher-dimensional space using kernel functions. However, the various projection forms and feature descriptions offered by different kernel functions must be explored. The next subsection discusses how a suitable kernel function can be selected for the variety of radiation source features.

### 3.2. Mixed Kernel

The effectiveness of the kernel function's nonlinear fit in the feature space relies not only on its capacity for learning from neighboring data (i.e., interpolation) but also on its ability to extend beyond its observed data range (i.e., extrapolation). Kernel functions can be categorized into local and global types. The local kernel, exemplified by the RBF kernel function, excels in interpolation but lacks extrapolation capabilities. Conversely, the global kernel, exemplified by the poly kernel function, exhibits superior extrapolation but weaker interpolation capabilities.

The formula for RBF kernel function:

$$K_g(x_i, x_j) = \left\langle -\frac{\|x_i - x_j\|^2}{2\sigma^2} \right\rangle \tag{16}$$

where  $\sigma$  denotes the kernel width.

The formula for poly kernel function:

$$K_p(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d \tag{17}$$

where  $d$  is the kernel parameter that denotes the degree of the poly.

The performance of these two kernel functions is illustrated in Figure 3. The RBF kernel reaches its maximum value when the test point’s distance is zero, gradually approaching zero as the distance increases. This indicates a limited learning ability beyond a specific range. Conversely, the poly kernel exhibits increasing kernel values across all ranges as the poly degree rises, but its interpolation ability weakens. To harness both nonlinear learning capabilities simultaneously, a combined approach is considered.

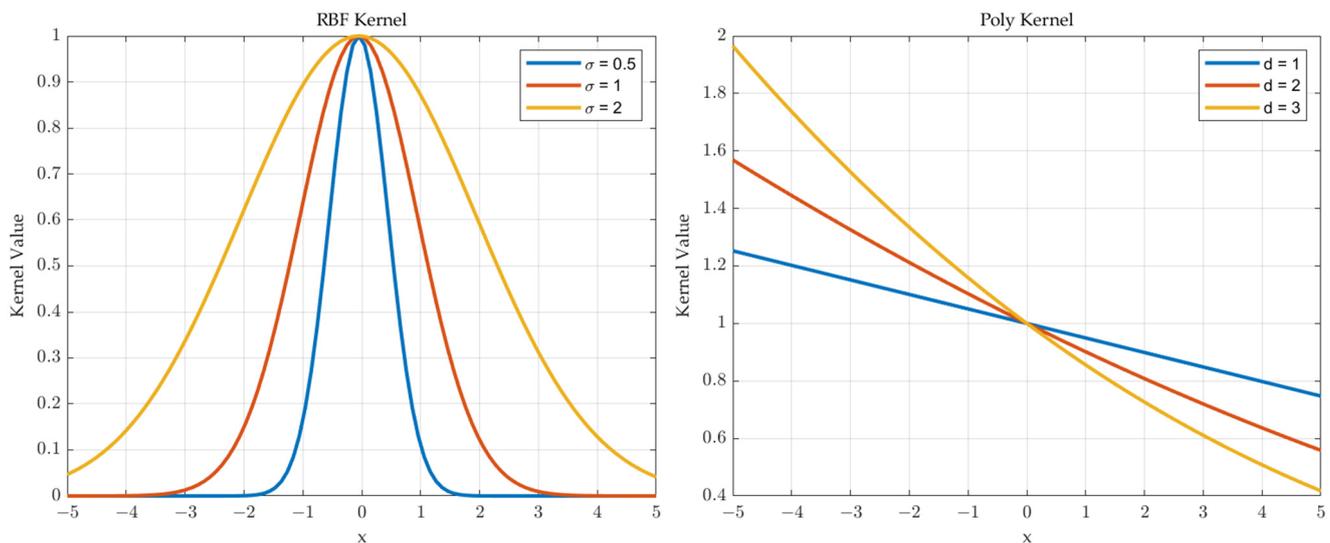


Figure 3. Kernel values of RBF kernel (left) and poly kernel (right).

Unlike single kernel models, mixed kernel models possess both interpolation and extrapolation capabilities. Unlike their single kernel counterparts, mixed kernel models offer a more expansive assumption space, making them better suited for approximating real-world problem objective functions [23,24].

Approach to combining mixed kernels:

$$K_{\min} = \omega K_p + (1 - \omega) K_g \tag{18}$$

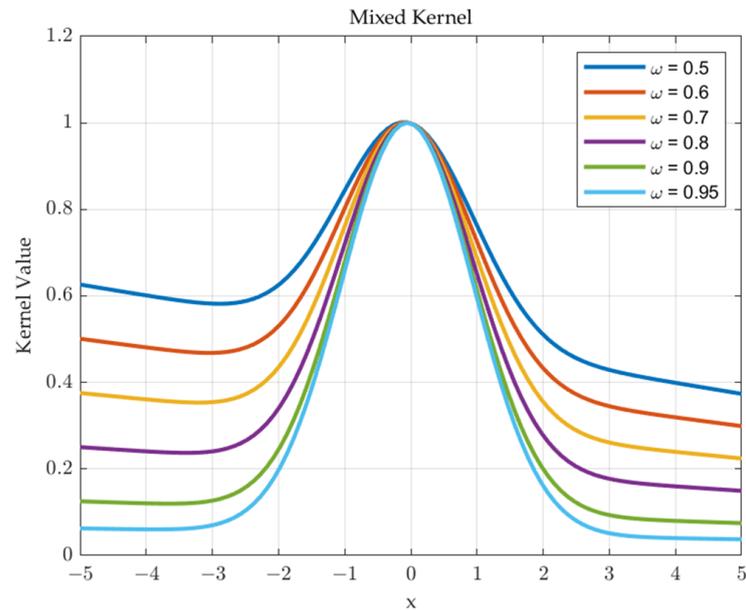
where  $\omega \in [0, 1]$  represents the mixture weight.

Figure 4 demonstrates the kernel values resulting from the fusion of the RBF and poly kernels, with assumption of parameters  $\sigma = 1$  for the RBF kernel and  $d = 1$  for the poly kernel. Adjusting the  $\omega$  parameter in Equation (15) reveals that the mixed kernel function exhibits a consistent nonlinear fitting effect under varying weights. However, the choice of parameters significantly influences algorithm performance, a topic explored in the subsequent section.

### 3.3. Parameters Optimization

The selection of parameters in the aforementioned mixed kernel method directly impacts the algorithm’s performance. Thus, employing parameter optimization becomes essential to identify the optimal parameter combination and enhance algorithm performance [21]. A genetic algorithm is used for parameter optimization, leveraging the inheritance of superior parameters from the previous generation to expedite the optimization process. While there is a possibility of falling into local optima and missing the global optimum solution, an examination of the kernel function reveals a small parameter range,

minimizing the risk of overlooking the global optimum solution. Recognition accuracy under different parameter settings serves as the fitness function, with the optimal parameter combination identified when either the highest recognition accuracy condition is met or the number of genetic iterations reaches 200.



**Figure 4.** Kernel values of mixed kernel ( $\sigma = 1, d = 1$ ).

## 4. Experimental Analysis

### 4.1. Datasets

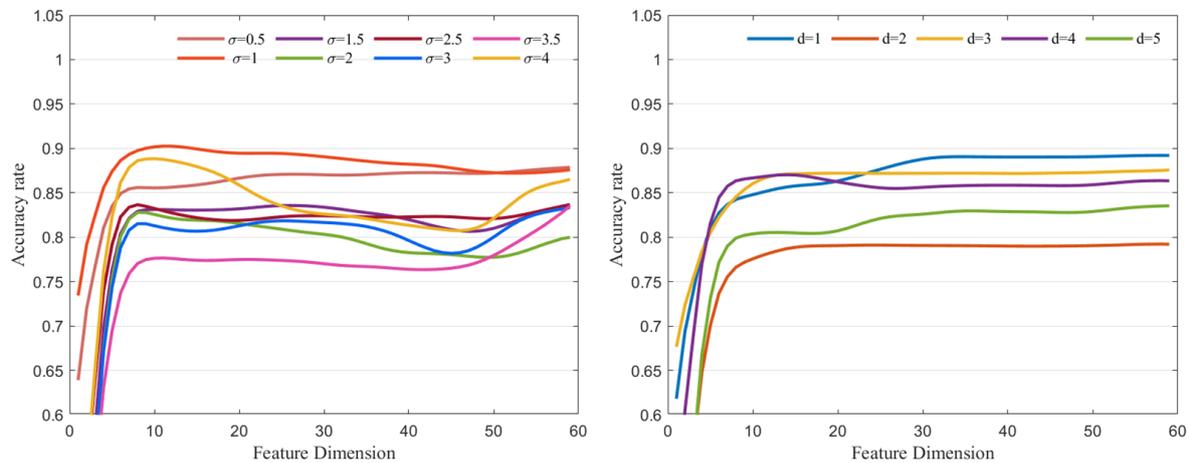
To assess the recognition effectiveness of the proposed algorithm, experimental validation is conducted using a radar dataset. The dataset comprises pulse signals emitted by eight analogue radars of three models, collected within a laboratory environment. Each simulation radar captures 200 radar pulses, with each pulse signal consisting of 1200 sample points. In evaluating the feature fusion algorithm's performance, a random forest classifier is chosen for classification and recognition. Training utilizes 80% of the collected samples, while the remaining 20% are reserved for testing.

### 4.2. Kernel Function Analysis

Examining the influence of kernel function parameters, this subsection investigates the parameter selection for the proposed mixed kernel. Key parameters include those in the RBF kernel function  $\sigma$ , the degree parameter  $d$  in the poly kernel function, and the weights  $\omega$  in the mixed kernel function. The MKCCA approach employed in this study involves critical parameter selection, as each parameter significantly affects the algorithm's performance. Given the difficulty in estimating these parameters, they often require prior information and are manually determined within an appropriate range. In practice, optimal parameters in experiments typically necessitate only a brief search within a narrow range. The parameters of the kernel function in KCCA are discussed below. First, feature vectors are input to the KCCA, which varies the parameters of the kernel function and outputs dimensionally variable feature fusion vectors, which are used as inputs to the Random Forest classifier, which outputs individual recognition accuracy. Experiments are conducted for the range of parameter values one by one to analysis the change in individual recognition accuracy for parameter values with different feature dimensions. Considering the information redundancy associated with excessively high feature dimensions, the maximum feature dimension is capped at 60.

The RBF kernel function, a focal point of recent research, represents the local kernel function. Research indicates that the parameter  $\sigma$  of the RBF kernel exhibits strong inter-

pulation ability when taking smaller values. Conversely, larger values of  $\sigma$  weaken the kernel's interpolation ability while enhancing extrapolation ability. Hence, the range for  $\sigma$  is concentrated between (0, 5). To explore the RBF kernel's performance under different parameters, the range of  $\sigma$  in this subsection is {0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4}, as illustrated in the left of Figure 5.



**Figure 5.** Performance of RBF kernel parameter  $\sigma$  (left) and poly kernel parameter  $d$  (right).

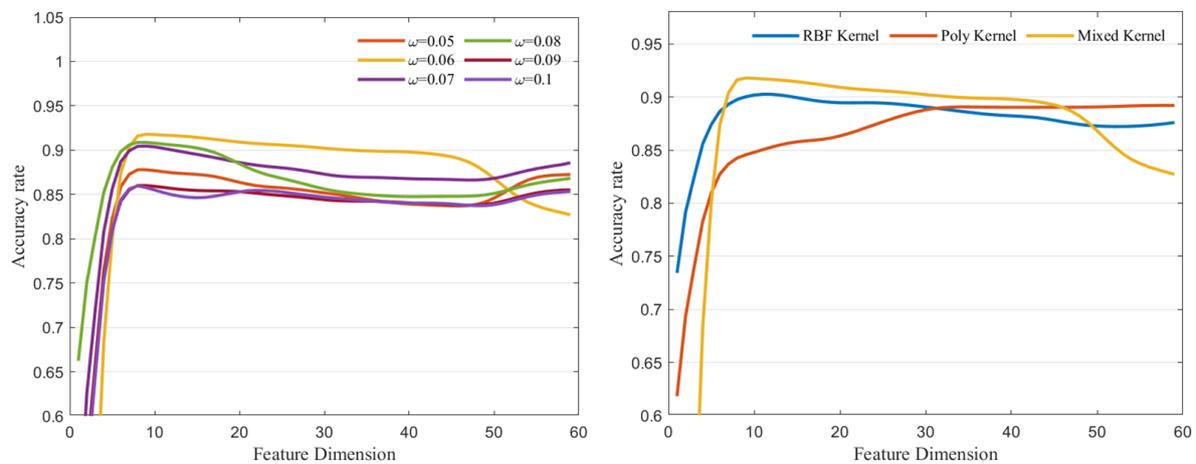
In this experiment,  $\sigma$  is selected from the specified range, with individual identification accuracy serving as the criterion as the feature dimension increases. The figure reveals that the RBF kernel function performs optimally when  $\sigma$  is set to 1. However, with  $\sigma$  set to 3.5, the recognition accuracy remains consistently low with increasing feature dimensions. For all other  $\sigma$  values within the specified range, recognition accuracy falls within the mid-range. When  $\sigma$  is set to 4, recognition rates rapidly increase with small feature dimensions but decrease as dimensions increase. Thus, determining the value of  $\sigma$  is particularly crucial for optimizing the RBF kernel function's performance.

Represented by the poly kernel function, research has demonstrated its commendable extrapolation capability. Larger values of parameter  $d$  contribute to superior interpolation capability but diminish extrapolation capability. Conversely, smaller values of  $d$  enhance extrapolation capability but weaken interpolation capability. The parameter range for the poly kernel is set to {1, 2, 3, 4, 5}, as illustrated in the right of Figure 5, showcasing the recognition accuracy comparison for varying values of parameter  $d$ . At  $d = 1$ , as feature dimensions range between (10, 30), recognition accuracy exhibits slower growth compared to other parameter values. At  $d = 3$ , recognition accuracy remains stable when the feature dimension reaches 12.

Simultaneously, the mixed kernel possesses the ability to interpolate and extrapolate, ensuring better adaptation to the characteristics of multi-featured data. According to Equation (13), the mixed kernel function requires determining the weight  $\omega$  of the two. The value of weight  $\omega$  is set in the range of {0.05, 0.06, 0.07, 0.08, 0.09, 0.10}. Considering the potential correlation between the three parameters, a genetic algorithm is employed for parameter optimization. The algorithm performs optimally when the three parameters are set to  $\sigma = 1$ ,  $d = 1$  and  $\omega = 0.06$ , respectively.

To explore the influence of the weight  $\omega$  on the algorithm's recognition effectiveness, comparative experiments with different parameter values are conducted. The left of Figure 6 presents the recognition rate comparison for the mixed kernel parameter  $\omega$ , considering values in the set range. As feature dimensions increase up to 10, the proposed algorithm exhibits a rapid increase in recognition accuracy under all parameters. Subsequently, recognition accuracy gradually declines as dimensions continue to grow, indicating information redundancy and a subsequent decrease in individual recognition accuracy. This validates the algorithm's ability to extract effective feature information with a low fea-

ture dimension. When  $\omega$  is set to 0.06, recognition accuracy is highest within the parameter set, with recognition accuracy under other parameters distributed between (85, 90).



**Figure 6.** Performance of the mixed kernel parameter  $\omega$  (left) and comparison of kernel functions (right).

Comparing the detection accuracy of the RBF and poly kernel functions mentioned earlier revealed that the mixed kernel model proposed in this paper outperforms them in radar SEI, as illustrated in the right of Figure 6. The figure displays the parameter settings of the three kernel functions with the highest recognition accuracy within their respective parameter sets. Specifically, the parameter  $\sigma$  of the RBF is set to 1, the parameter  $d$  of the poly kernel function is set to 1, and the weight  $\omega$  of the mixed kernel is set to 0.06. The accuracy comparison reveals that the RBF kernel function attains very high recognition accuracy with smaller feature dimensions below 10. However, its accuracy diminishes as dimensions increase. In contrast, the poly kernel function exhibits a gradual increase in recognition rate within the feature dimensions of 10 to 20, and after reaching 20, the accuracy rate elevates to a higher level.

By comparing the four graphs in Figures 5 and 6, it can be observed that the RBF kernel function performs well in low dimensions (around 10) but exhibits poorer performance in high dimensions. In contrast, the polynomial kernel function demonstrates highly stable performance in high dimensions. The mixed kernel function addresses the limitations of the former, showing relatively stable performance in dimensions below 40. In summary, the RBF kernel function demonstrates superior nonlinear expansion capability for a smaller number of data points, implying robust interpolation ability but weaker extrapolation capability with more data points. On the other hand, the poly kernel function excels in extrapolation ability for a greater number of data points, showing a slower increase in recognition rate initially. Combining the strengths of both, the mixed kernel function achieves higher recognition accuracy than the other two kernel functions when the feature dimension is less than 40. However, its accuracy significantly decreases beyond 40. For the method proposed in this paper, it attains superior recognition performance when the feature dimension is below 40, showcasing commendable interpolation and extrapolation abilities.

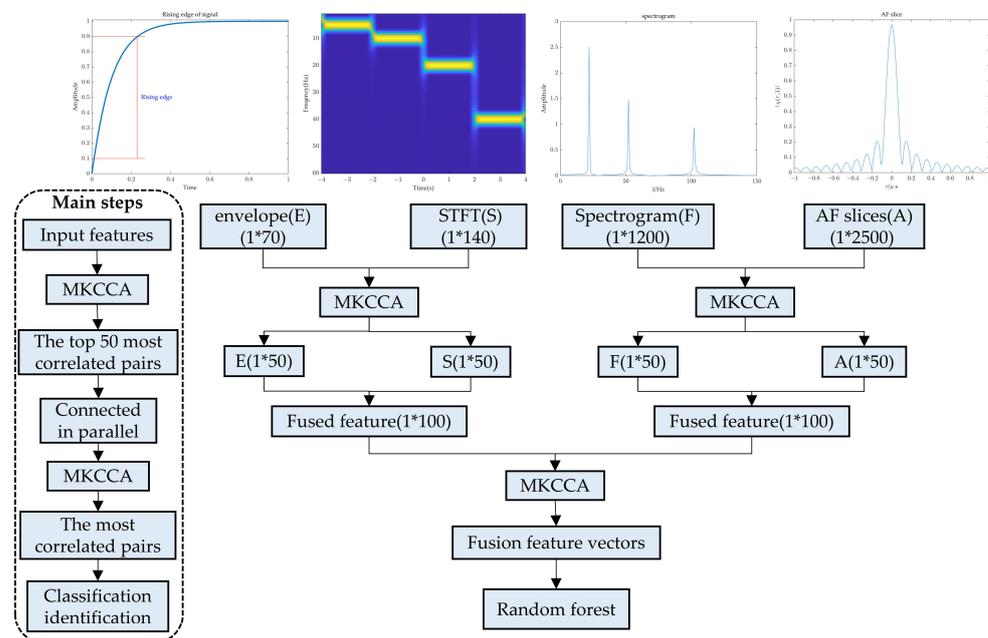
The fitting speed of different kernel functions varies, providing an additional dimension for assessing their performance. Table 2 presents the time comparison of the three kernel functions, showcasing their fitting speeds. This data reflects the average fitting time across various parameters in preceding trials, offering an accurate depiction of the kernel functions' performance. Among them, the poly kernel function exhibits the fastest fitting speed under the same dataset. The hybrid kernel model proposed in this paper demonstrates a slightly shorter time and faster fitting speed compared to the commonly used RBF kernel function.

**Table 2.** KCCA’s time using different kernel function.

Kernel Function	Mixed Kernel	RBF Kernel	Poly Kernel
Time (s)	4.1131	4.9133	3.8902

### 4.3. Multi-Domain Feature Fusion Analysis

Building upon the literature on fingerprint feature extraction, this study derives representative fingerprint features from various domains for fusion. Specifically, four fingerprint features are extracted: the envelope rising edge feature (E), the spectrum feature (F), the short-time Fourier transform (S), and the near-zero slice of the ambiguity function (A). This notation allows for clear representation: E denotes time domain features, F denotes frequency domain features, S denotes time-frequency domain features, and A denotes the ambiguity function (AF). The definitions of the four fingerprint features are given below and are shown in Figure 7.



**Figure 7.** Multi features fusion flow diagram. (The diagram of four features is shown on the top of the figure, the main steps are shown in the dashed box on the left of the figure, and the detailed process of feature fusion is shown on the right of the figure.).

The signal envelope  $A(t)$  is defined as:

$$A(t) = \sqrt{s_I^2(t) + s_Q^2(t)} = |s(t)| \tag{19}$$

where  $s_I(t)$  and  $s_Q(t)$  are orthogonal signals and the rising edge of the envelope is the front part of the envelope.

The signal spectrum  $U(f)$  is defined as:

$$U(f) = \int_{-\infty}^{+\infty} u(t)e^{-j\omega t} dt \tag{20}$$

where  $u(t)$  is the radar signal.

The short-time Fourier transform is defined as:

$$STFT = \int_{-\infty}^{+\infty} u(\tau)g^*(\tau - t)e^{-j\omega t} d\tau \tag{21}$$

where  $a$  is the window function and  $*$  is the complex conjugation.

The near-zero slice of the ambiguity function  $A_u(\tau, \xi)$  is defined as [30,31]:

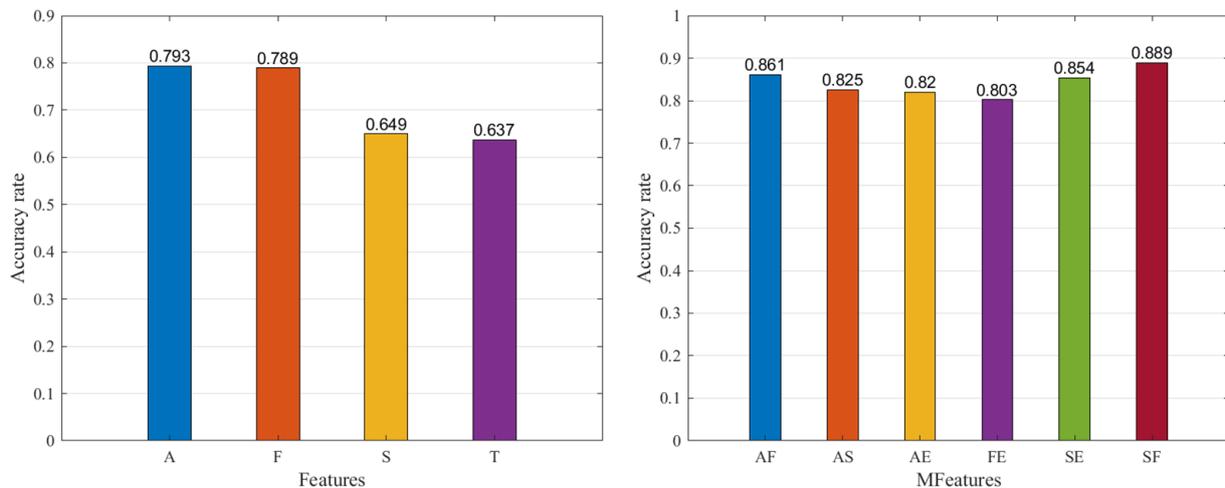
$$A_u(\tau, \xi) = \int_{-\infty}^{+\infty} U(f)U^*(f - \xi)e^{j2\pi f\tau}df \quad (22)$$

where  $\xi$  denotes the frequency shift, usually taken as 1 for values near zero.  $U(f)$  denotes the signal spectrum.  $U^*(f \pm \xi)$  denotes the conjugate frequency shift of the signal spectrum.

The fusion process of MKCCA is depicted in Figure 7, illustrating the dimensions of the four fingerprint features extracted. MKCCA takes two feature vectors as input, producing the first 50-dimensional vectors with the highest correlation. These vectors are then fused column-wise, resulting in a 100-dimensional fusion vector. This procedure is repeated for the remaining two features, and the two MKCCA outputs are fused. The optimal feature dimension is determined through a systematic increase in feature dimensions. It is then passed through a random forest classifier.

In the multi-feature fusion experiment, the number of fused features is incrementally increased—two, three, and four features are fused to assess the impact on recognition accuracy. The recognition accuracy, averaged across preserving feature dimensions 1 to 100, is displayed using a histogram.

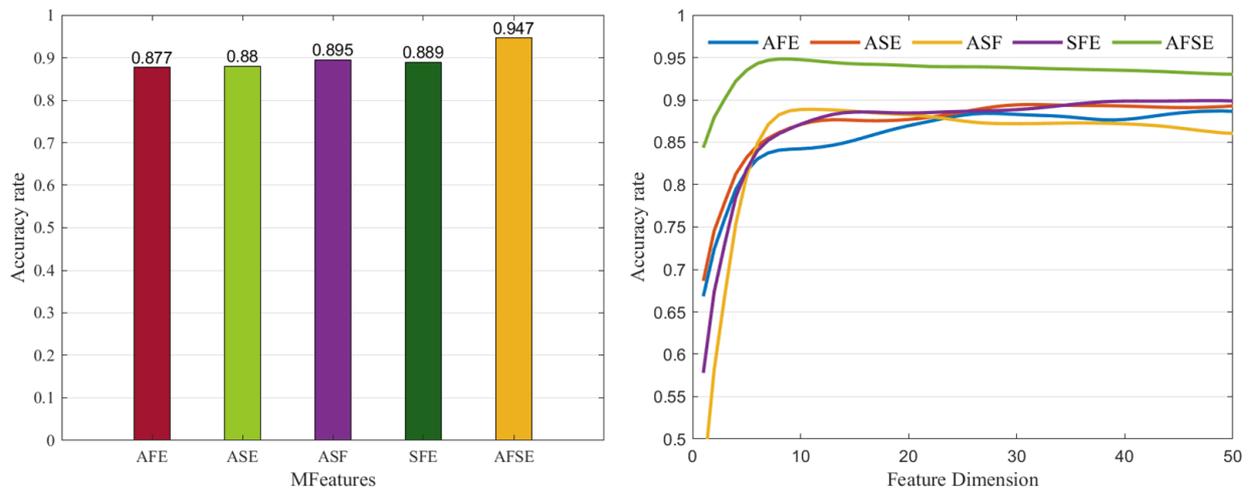
The average recognition accuracy of a single feature across all dimensions of the feature set is presented in the left of Figure 8. The recognition rate ranges from 63 to 80, indicating insufficient effectiveness for recognizing individual radar radiation sources. Recognizing the complementarity between features from different domains, the fusion of various features is explored in the subsequent analyses.



**Figure 8.** Recognition effect of single feature (left) and double features (right).

The right of Figure 8 illustrates that recognition accuracies of dual feature fusion fall within the interval (80, 89). The fusion of time-frequency features with frequency features performs best, followed by the fusion of ambiguity functions with frequency features. Although the recognition effect improves compared to a single feature, the distribution interval remains large, and the accuracy falls short of expectations.

As illustrated in the left of Figure 9, the recognition accuracy of three-feature fusion lies within the interval (87, 90), displaying reduced distribution and improved stability compared to single- and dual-feature fusion methods. The recognition result of four-feature fusion reaches 94%. Complementary use of different feature domains significantly enhances individual recognition of radar radiation sources, providing higher stability to accommodate sample diversity.

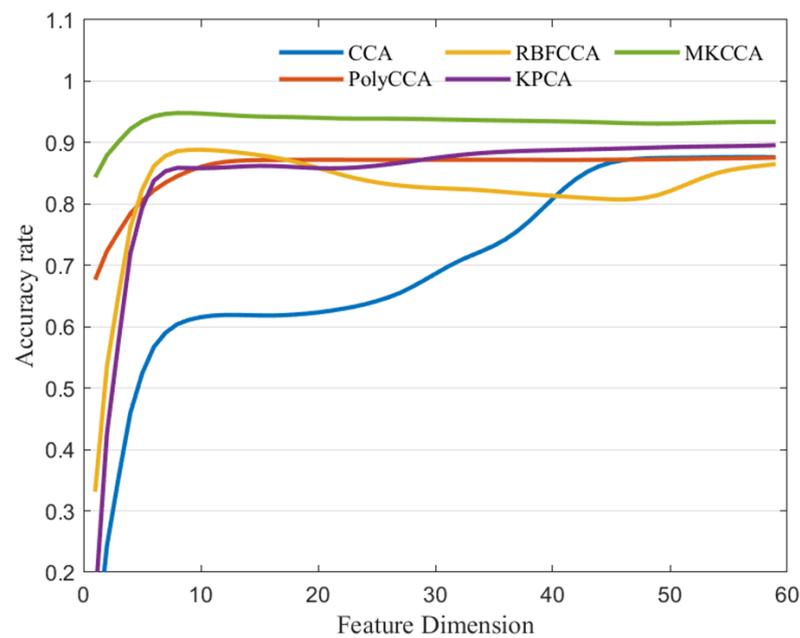


**Figure 9.** Recognition effect of multiple features (left) and influence of feature dimensions on recognition performance (right).

In the KCCA fusion algorithm, features with the largest correlation coefficient must be combined as fusion features. The right of Figure 9 illustrates that the improved recognition impact after fusion is proportional to the feature dimension. A turning point is observed at a feature dimension of 10, where the recognition impact is significantly enhanced for dimensions less than 10, while the recognition rate remains constant for dimensions larger than 10.

4.4. Performance Analysis

This section delves into the performance of various feature fusion algorithms, comparing the recognition accuracy of the proposed method with existing fusion algorithms. Furthermore, it explores the influence of different feature dimensions on accuracy. Figure 10 presents a performance comparison of the fusion algorithm.



**Figure 10.** Performance of the fusion algorithm.

In this experiment, the dimension range for all fusion algorithms is set from 0 to 60. Notably, the recognition accuracy of each fusion algorithm gradually increases with the growing dimension. The study underscores that the method proposed in this article outperforms other fusion algorithms, achieving the highest recognition accuracy of 95% while utilizing a smaller feature dimension of 10. The standard CCA algorithm exhibits inferior recognition accuracy compared to several other fusion algorithms at lower feature dimensions. Recognition outcomes become comparable to other algorithms only when the feature dimension exceeds 40, achieving a rate of approximately 87%. KCCA with the RBF kernel attains the highest recognition accuracy of 89.37% at a feature dimension of 10, but the accuracy drops as the feature dimension increases to 50. KCCA with a poly kernel achieves a peak recognition accuracy of 87.34% when increasing feature dimensionality to 20, demonstrating stable performance with dimension increase. In contrast, the KPCA algorithm shows a gradual ascent in recognition accuracy with a feature dimension increased to 10, albeit at a sluggish rate. Therefore, the algorithm introduced in this study exhibits a noticeable enhancement in recognition accuracy while significantly reducing feature redundancy compared to other fusion algorithms.

Table 3 presents the time spent by all fusion algorithms in the experiment, from loading the original dataset to obtaining the recognized results. The average time spent in the experiment is calculated for the final result. Notably, the poly CCA fusion algorithm, utilizing a poly kernel function, demonstrates the shortest algorithm time. The method proposed in this paper ranks as the second shortest. It is noteworthy that the CCA methods employed all had shorter durations compared to KPCA in terms of time. In terms of algorithm time, both the CCA method and the KCCA method extended with a kernel function outperform PCA in terms of timeliness.

**Table 3.** Time spent on the above algorithm.

Method	CCA	Poly CCA	RBF CCA	KPCA	MKCCA
Time(s)	28.543	15.116	22.797	40.237	19.729

## 5. Summary

This paper introduces a novel approach for SEI of radar radiation sources based on MMKCCA. The acquired radar signals undergo analysis to extract distinctive features from the time domain, frequency domain, time-frequency domain, and ambiguity-function domain. Subsequently, these features are amalgamated utilizing the MKCCA data fusion method, yielding a comprehensive feature set for identification purposes. The experimentation affirms the complementary nature of various feature domains, underscoring the efficacy of this fusion method in harnessing the complementarity of multiple features. The kernel function plays a pivotal role in transforming features into a high-dimensional space, facilitating the identification of nonlinear relationships among features within that space. Employing a mixed kernel—comprising a weighted combination of a local kernel and a global kernel—enhances the extraction of effective information from diverse features. This approach not only fosters an improved understanding of the nonlinear correlations among features but also significantly reduces the feature dimension. The experimental results demonstrate that the proposed feature fusion method yields a commendable recognition effect even with lower feature dimensions. It was not able to experiment with all fingerprint features in this paper, thus future research will focus on the performance of other fingerprint aspects.

**Author Contributions:** Methodology, H.L.; Formal analysis, S.L. and H.L.; Resources, J.C.; Writing—original draft, J.C. and J.Q.; Writing—review & editing, J.C., S.L. and H.L.; Supervision, S.L. and J.Q. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The experimental data in this paper are mainly obtained by mathematical simulation.

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

## References

1. Liu, Y.; Li, S.; Lin, X.; Gong, H.; Li, H. Feature Analysis and Extraction for Specific Emitter Identification Based on the Signal Generation Mechanisms of Radar Transmitters. *Sensors* **2022**, *22*, 2616. [\[CrossRef\]](#)
2. Fu, X.; Peng, Y.; Liu, Y.; Lin, Y.; Gui, G.; Gacanin, H.; Adachi, F. Semi-Supervised Specific Emitter Identification Method Using Metric-Adversarial Training. *IEEE Internet Things J.* **2023**, *10*, 10778–10789. [\[CrossRef\]](#)
3. Wang, Y.; Gui, G.; Gacanin, H.; Ohtsuki, T.; Dobre, O.A.; Poor, H.V. An Efficient Specific Emitter Identification Method Based on Complex-Valued Neural Networks and Network Compression. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 2305–2317. [\[CrossRef\]](#)
4. He, B.; Wang, F. Cooperative Specific Emitter Identification via Multiple Distorted Receivers. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 3791–3806. [\[CrossRef\]](#)
5. Zhu, M.; Feng, Z.; Stanković, L.; Ding, L.; Fan, J.; Zhou, X. A probe-feature for specific emitter identification using axiom-based grad-CAM. *Signal Process.* **2022**, *201*, 108685. [\[CrossRef\]](#)
6. Tan, K.; Yan, W.; Zhang, L.; Tang, M.; Zhang, Y. Specific Emitter Identification Based on Software-Defined Radio and Decision Fusion. *IEEE Access* **2021**, *9*, 86217–86229. [\[CrossRef\]](#)
7. Zhao, Y.; Wang, X.; Lin, Z.; Huang, Z. Multi-Classifer Fusion for Open-Set Specific Emitter Identification. *Remote Sens.* **2022**, *14*, 2226. [\[CrossRef\]](#)
8. Ali, A.M.; Uzundurukan, E.; Kara, A. Assessment of Features and Classifiers for Bluetooth RF Fingerprinting. *IEEE Access* **2019**, *7*, 50524–50535. [\[CrossRef\]](#)
9. Suski Ii, W.C.; Temple, M.A.; Mendenhall, M.J.; Mills, R.F. Radio frequency fingerprinting commercial communication devices to enhance electronic security. *Int. J. Electron. Secur. Digit. Forensics* **2008**, *1*, 301–322.
10. Ru, X.-H.; Liu, Z.; Huang, Z.-T.; Jiang, W.-L. Evaluation of unintentional modulation for pulse compression signals based on spectrum asymmetry. *IET Radar Sonar Navig.* **2017**, *11*, 656–663. [\[CrossRef\]](#)
11. Li, L.; Ji, H.-B.; Jiang, L. Quadratic time–frequency analysis and sequential recognition for specific emitter identification. *IET Signal Process.* **2011**, *5*, 568–574. [\[CrossRef\]](#)
12. Li, Y.-b.; Ge, J.; Lin, Y.; Ye, F. Radar emitter signal recognition based on multi-scale wavelet entropy and feature weighting. *J. Cent. South Univ.* **2014**, *21*, 4254–4260. [\[CrossRef\]](#)
13. Mateo, C.; Talavera, J.A. Short-time Fourier transform with the window size fixed in the frequency domain. *Digit. Signal Process.* **2018**, *77*, 13–21. [\[CrossRef\]](#)
14. Chen, J.; Xing, M.; Xia, X.G.; Zhang, J.; Liang, B.; Yang, D.G. SVD-Based Ambiguity Function Analysis for Nonlinear Trajectory SAR. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 3072–3087. [\[CrossRef\]](#)
15. Liu, Z.-M. Multi-feature fusion for specific emitter identification via deep ensemble learning. *Digit. Signal Process.* **2021**, *110*, 102939. [\[CrossRef\]](#)
16. Hou, K.; Li, N. Specific emitter identification based on CNN. *J. Phys. Conf. Ser.* **2021**, *1971*, 012014. [\[CrossRef\]](#)
17. Zhang, X.; Li, T.; Gong, P.; Zha, X.; Liu, R. Variable-Modulation Specific Emitter Identification With Domain Adaptation. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 380–395. [\[CrossRef\]](#)
18. Yang, J.; Yang, J.-Y.; Zhang, D.; Lu, J.-F. Feature fusion: Parallel strategy vs. serial strategy. *Pattern Recognit.* **2003**, *36*, 1369–1381. [\[CrossRef\]](#)
19. Tong, L.; Fang, M.; Xu, Y.; Peng, Z.; Zhu, W.; Li, K. Specific Emitter Identification Based on Multichannel Depth Feature Fusion. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 9342085. [\[CrossRef\]](#)
20. He, F.; Zhang, Z. Nonlinear Fault Detection of Batch Processes Using Functional Local Kernel Principal Component Analysis. *IEEE Access* **2020**, *8*, 117513–117527. [\[CrossRef\]](#)
21. Jia, M.; Xu, H.; Liu, X.; Wang, N. The optimization of the kind and parameters of kernel function in KPCA for process monitoring. *Comput. Chem. Eng.* **2012**, *46*, 94–104. [\[CrossRef\]](#)
22. Shiju, S.S.; Salim, A.; Sumitra, S. Multiple kernel learning using composite kernel functions. *Eng. Appl. Artif. Intell.* **2017**, *64*, 391–400.
23. Pilario, K.E.S.; Cao, Y.; Shafiee, M. Mixed kernel canonical variate dissimilarity analysis for incipient fault monitoring in nonlinear dynamic processes. *Comput. Chem. Eng.* **2019**, *123*, 143–154. [\[CrossRef\]](#)
24. Zhu, X.; Huang, Z.; Tao Shen, H.; Cheng, J.; Xu, C. Dimensionality reduction by Mixed Kernel Canonical Correlation Analysis. *Pattern Recognit.* **2012**, *45*, 3003–3016. [\[CrossRef\]](#)
25. Gao, X.; Niu, S.; Sun, Q. Two-Directional Two-Dimensional Kernel Canonical Correlation Analysis. *IEEE Signal Process. Lett.* **2019**, *26*, 1578–1582. [\[CrossRef\]](#)
26. Yoshida, K.; Yoshimoto, J.; Doya, K. Sparse kernel canonical correlation analysis for discovery of nonlinear interactions in high-dimensional data. *BMC Bioinform.* **2017**, *18*, 108. [\[CrossRef\]](#)
27. Hardoon, D.R.; Szedmak, S.; Shawe-Taylor, J. Canonical Correlation Analysis: An Overview with Application to Learning Methods. *Neural Comput.* **2004**, *16*, 2639–2664. [\[CrossRef\]](#)

28. Guo, S.; Sheng, Y.; Chai, L.; Zhang, J. PET Image Reconstruction with Kernel and Kernel Space Composite Regularizer. *IEEE Trans. Med. Imaging* **2023**, *42*, 1786–1798. [[CrossRef](#)]
29. Cai, J.; Tang, Y.; Wang, J. Kernel canonical correlation analysis via gradient descent. *Neurocomputing* **2016**, *182*, 322–331. [[CrossRef](#)]
30. Pu, Y. Morphological Feature Extraction Based on the Polar Transformation of the Slice of Ambiguity Function Main Ridge. *Yi Qi Yi Biao Xue Bao/Chin. J. Sci. Instrum.* **2018**, *39*, 1–9.
31. Wan, T.; Ji, H.; Xiong, W.; Tang, B.; Fang, X.; Zhang, L. Deep Learning-Based Specific Emitter Identification Using Integral Bispectrum and the Slice of Ambiguity Function. *Signal Image Video Process.* **2022**, *16*, 2009–2017. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.