https://doi.org/10.31449/inf.v50i8.10595 Informatica **50** (2026) 55–28 **55**

Adaptive Multiple-Kernel SVM with Joint Kernel-Weight and Hyperparameter Optimization for IR Spectroscopy

Yuchun Liu, Haiwei Wu*, Xuexin Li, Jintong Tu
College of Engineering and Technology, Jilin Agricultural University, Changchun 130118, China
E-mail: HaiweiWuu@outlook.com
*Corresponding author

Keywords: support vector machine, compound kernel function, feature extraction, intelligent optimization, infrared spectroscopy analysis

Received: August 5, 2025

Infrared spectroscopy analysis technology is widely used in material identification and other fields due to its non-destructive and fast characteristics. However, the high-dimensional, nonlinear, and noisy nature of the data poses challenges to the robustness of feature extraction and classification models; Although SVM performs well in spectral classification, its performance depends on kernel function selection, and traditional single kernel processing of complex infrared spectral data has limited capabilities. To this end, this study focuses on the SVM composite kernel intelligent selection strategy and proposes a framework of integrated intelligent optimization algorithm. With the goal of maximizing classification accuracy and minimizing model complexity, the optimal weighted combination is searched in candidate kernel pools such as polynomial kernel and RBF (Radial Basis Function) kernel. The core innovation is the collaborative automatic optimization of composite kernel weight coefficients and SVM parameters; Based on the ATR-FTIR dataset of 120 organic compounds, experiments were conducted using a modified PSO (Particle Swarm Optimization) algorithm (optimizing inertia weights and learning factors, constructing a fusion fitness function for collaborative optimization) and an outer 5-fold cross validation. The results showed that the optimized composite kernel model achieved a classification accuracy of 94.1% ± 1.2%, which was 11.5 percentage points higher than the traditional RBF kernel (82.6% ± 1.8%). The convergence speed of the optimization iteration was 47% ± 3.5% higher than the standard PSO, and the prediction accuracy on the near-infrared drug/food dataset reached 96.67% ± 0.9%. The generalization and robustness were significant, providing a new paradigm for intelligent feature analysis of complex infrared spectral data and improving the accuracy and automation level of spectral substance identification.

Povzetek: Študija predlaga inteligentno optimiziran SVM z združenimi jedri za infrardečo spektroskopijo, ki bistveno izboljša natančnost, robustnost in avtomatizacijo pri prepoznavanju snovi v primerjavi s klasičnimi pristopi.

1 Introduction

As an important means of modern non-destructive testing and material identification, the core value of infrared spectroscopy technology lies in the characteristic response to the vibration of molecular bond structure of materials, which provides a fast and effective way of component identification and quantitative analysis for many fields such as chemistry, biology, environment and medicine [1]. The inherent complexity of infrared spectral data significantly restricts the depth and accuracy of its information analysis [2, 3]. This kind of data is inherently high-dimensional, showing strong nonlinearity and potential noise pollution characteristics. These inherent data attributes pose a severe test for effective feature extraction and robust pattern recognition [4]. The traditional single kernel function support vector machine model often faces the dilemma of limited representation ability when dealing with this complex spectral structure, and it is difficult to fully capture the multi-scale, heterogeneous and highly nonlinear feature patterns contained in the data, resulting in the generalization performance of the model being difficult to meet the actual needs of high-precision material identification [5, 6].

As a key link in the spectral data processing process, the effectiveness of feature extraction directly determines the quality of subsequent modeling [7]. However, a single feature extraction method may not be able to stably obtain the most discriminant low-dimensional subspace information when dealing with varied infrared spectral data [8]. As a powerful tool for spectral pattern recognition, the geometric characteristics of classification boundaries and nonlinear mapping capabilities of support vector machine models are highly dependent on the rational choice of kernel functions [9, 10]. Commonly used single kernel functions have their inherent advantages and limitations. For example, radial basis kernel functions have excellent local fitting ability but can't describe the global structure enough. Polynomial kernel functions perform well in describing global relationships but are sensitive to local details and noise, while Sigmoid kernel may have saturation phenomenon under certain circumstances. This single selection of kernel function is inherently difficult to adaptively match the complex distribution characteristics of the data subspace after infrared spectral feature extraction, which becomes a bottleneck restricting the further improvement of the performance of the classification model [11].

In order to solve the performance bottleneck caused by the insufficient adaptability of single kernel function, it has become an important research direction to explore the strategy of constructing composite kernel function with more flexible representation ability. The core idea of composite kernel is to integrate a variety of different types of basis kernel functions in a linear or nonlinear way, in order to fuse their respective dominant spatial mapping characteristics, and finally generate a new metric space with stronger discriminant ability [12, 13]. Theoretically, this strategy can provide richer kernel expression forms, so as to better describe the intrinsic nonlinear geometric complexity and multi-scale correlation of the data after infrared spectral feature extraction. However, how to achieve an efficient combination of base kernels and determine their optimal weight coefficients is a key challenge for the successful application of composite kernel methods [14]. Traditional composite kernel construction strategies based on experience or linear combination rules usually lack systematicness and adaptability, and their model effectiveness depends on the designer's professional knowledge and depth of understanding of specific data sets, so it is difficult to have wide applicability [15]. This highlights the urgent need for automatic and intelligent selection of composite kernel function combination and its internal parameters.

This study focuses on the intelligent selection strategy of support vector machine composite kernel for infrared spectral feature extraction. Its core research goal is to establish an adaptive intelligent optimization framework, which can automatically search and construct the optimal weighted composite kernel function form for specific infrared spectral feature data in a wide range of candidate basis kernel function spaces (such as polynomial kernel, Gaussian radial basis kernel, Sigmoid kernel, etc.) without too much human intervention.This strategy specifically underscores that the optimization process of the weight coefficients of each base kernel within the composite kernel must be closely coordinated with the crucial hyperparameters of the support vector machine model (e.g., penalty factor and in - kernel parameters). The essence of this collaborative optimization lies in achieving the internal matching between the structure of the composite kernel and the regularization ability of the model, so as to achieve the dynamic balance between the overall generalization ability and expression ability of the model. This method is committed to making up for the lack of automation and intelligent selection mechanism of existing composite kernel strategies, and breaking through the performance bottleneck of single kernel function or empirical combination strategies. The specific research objectives include: exploring and establishing a composite kernel function construction paradigm that integrates heuristic intelligent optimization algorithms, designing a dual parameter adjustment mechanism that can collaboratively optimize kernel weights and model hyperparameters, and finally developing a set of infrared spectral features. The highly adaptive SVM composite kernel intelligent selection methodology system for extracting scenes provides algorithmic framework for automatic and high-precision material identification of complex infrared spectral data.

This study focuses on SVM composite kernel intelligent selection for infrared spectral feature extraction. The core objectives include: firstly, achieving joint optimization of kernel weights and SVM hyperparameters to improve model adaptability; The second is to introduce an improved PSO algorithm to provide efficient solution support for the optimization process; Thirdly, standardize the criteria for selecting datasets and sample sizes to ensure consistency in experimental foundations; Fourthly, clear and reproducible expected outcome indicators should be established to ensure the reliability and comparability of research conclusions.

2 Theoretical framework of composite kernel selection for infrared spectral feature extraction

2.1 Characteristics of infrared spectral data

IR spectroscopic data analysis was used for qualitative and quantitative analysis [16, 17]. It identifies the functional groups of unknown compounds by specific frequencies and characteristic peaks, and infers the molecular formula, stereostructure, including chemical bonds and covalent bonds, to classify and identify unknown samples [18].

Quantitative analysis of infrared spectroscopy is based on Beer-Lambert's law, as shown in formula (1):

$$A = Ig(\frac{1}{T}) = Kbc \; (1)$$

In the formula, A is the absorbance, T is the transmittance, and K is the molar absorbance coefficient, which is related to the sample characteristics and the light wavelength λ. c is the substance concentration and b is the absorbent layer thickness. The Beer-Lambert law uses the wavelength of light to calculate the intrinsic absorbance K, which is used to detect the composition of organic compounds.

Fourier transform infrared spectroscopy (FTIR) is a non-destructive testing method based on Fourier transform for infrared spectral analysis [19, 20]. It uses sinusoidal and cosine wave fitting to measure infrared beam interference, improving signal-to-noise ratio and playing an important role in infrared spectroscopy [21]. The introduction of attenuated total reflection (ATR) technology has expanded infrared spectroscopy applications and addressed challenges in testing special samples [22]. Key features of ATR-FTIR include: (1) simple sample preparation with no strict requirements on morphology, size, or water content; (2) support for real-time in-situ detection; (3) high detection sensitivity with micron-level accuracy; (4) ability to obtain molecular structure information; and (5) usefulness in identifying sample types and properties.

2.2 Theoretical basis of SVM function

SVM (Support Vector Machine) is one of the supervised learning methods, which is based on statistical learning theory and structural risk minimization principles [23, 24]. It maps the data to a high-dimensional feature space through nonlinear transformation, and builds a linear optimization function in it. As an efficient classification algorithm, SVM distinguishes data categories by finding the best segmentation hyperplane, and the core is to maximize the interval between categories to ensure model generalization performance [25]. The introduction of kernel function technology enables SVM to deal with nonlinear data sets, which improves its widespread application [26]. Figure 1 shows the support vector model.

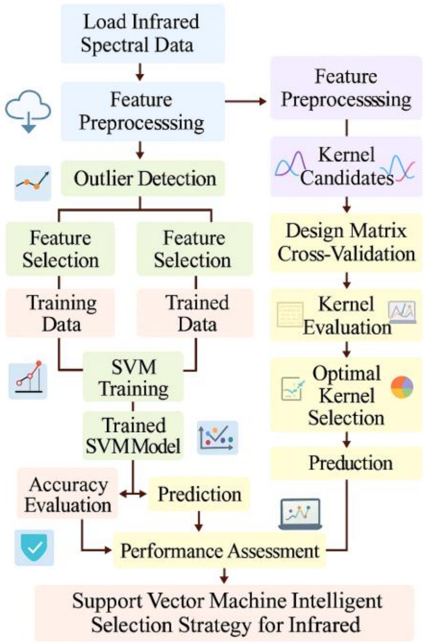

Figure 1: Support vector model

In multidimensional data space, vectors are mapped to high-dimensional feature space by kernel function, and the optimal classification hyperplane is found in it to improve the accuracy of nonlinear data classification [27]. The fitting function expression is (2):

$$f(x)=[w^T\varphi(x)]+b \quad (2)$$

In the formula, w represents the problem weight; $\varphi(x)$ is a mapping function; b is the problem threshold. According to the principle of structural risk minimization, the objective function of support vector machine regression model can be transformed into formula (3).

$$\begin{cases} min\frac{1}{2}||w||2+\frac{1}{2}C\sum_{i=1}^{n}\xi_i \\ s.t. y_i(wTx_i+b)\geq 1-\xi_i \end{cases} \quad (3)$$

The Lagrange function is constructed to transform the problem into a dual problem. ξ is a relaxation variable. The decision function of the prediction model is obtained by solving the kernel function. The expression is shown in (4):

$$f(x)=\sum_{i=1}^{n}\alpha_i k(x_i,x)+b \quad (4)$$

In the formula, n is the number of samples; α_i is the Lagrange multiplier; $k(x_i, x)$ is the kernel function. The type of kernel function determines the type of support vector machine and affects the learning performance. Common kernel function types are shown in Equation (5):

$$k(x_i, x_j) = exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (5)$$

$k(x_i,x_j)$ is the kernel function.Linear kernel functions and polynomial kernel functions are usually ineffective, low frequency of use and limited. Sigmoid kernel functions are difficult to train deep networks and are not commonly used [28]. In view of the strong generalization ability of radial basis kernel function, this study selects it as the core of support vector machine model.

Table 1: Comparison table of related jobs

Dimension	Linear Kernel	Composite Kernel (Linear-RBF)
Core Principle	Linear mapping; inner product similarity	Combines linear low-complexity & RBF nonlinearity; weighted
IR Spectra Suitability	Linear feature-target (simple components)	Complex spectra (unknown interference)
SVM Complexity	Low; fast; low overfitting	Medium; needs weights/sub-kernels
Feature Extraction Dependence	Relies on preprocessing (e.g., baseline)	Reduced dependence; high fault tolerance
Typical Applications	Single-component IR quantification	Multi-interference IR analysis

Table 1 systematically compares the performance differences between linear kernels and the linear RBF composite kernel proposed in this study from multiple dimensions such as core principles, applicability, and complexity. Analysis shows that compared to linear kernels that are only suitable for simple components (low complexity but strongly dependent on preprocessing) or single RBF/polynomial kernels that are difficult to balance linear/nonlinear features, composite kernels achieve collaborative extraction of local and global features through adaptive weight allocation strategies. This method significantly improves the analysis ability and robustness of infrared spectra under complex interference while maintaining moderate computational complexity. Its performance gain mainly comes from the intelligent weighting of kernel functions and efficient integration of local global features.

3 Construction of SVM composite kernel dynamic adaptation model

3.1 Parallel implementation framework of composite kernel SVM

Aiming at the intelligent selection problem of SVM composite kernel in infrared spectral feature extraction scenarios, an improved PSO algorithm is proposed: optimizing hyperparameters such as inertia weight and learning factor to improve search performance, constructing a single target composite fitness function (integrating SVM classification accuracy and kernel function complexity, formula $F=\alpha \cdot Acc+(1-\alpha) \cdot (1-K_{max}/K_{comp})$, where α is the weight coefficient, Acc is the classification accuracy, K_{comp} is the composite kernel complexity, and K_{max} is the upper limit of complexity), achieving collaborative optimization of composite kernel parameters and SVM hyperparameters, and improving the classification accuracy and model generalization ability of infrared spectral features.

Near-infrared spectroscopy data contain rich chemical information but are complex to process [29, 30]. Figure 2 shows the parallelization framework of the composite core SVM. Different substances have different spectral characteristics, and the data are often nonlinear. SVM model can effectively deal with high-dimensional data sets, and make nonlinear data linearly separable by mapping to high-dimensional feature space through kernel function. Therefore, SVM excels in dealing with nonlinear and high-dimensional data.The core support lies in the fact that the constructed composite kernel is a non negative linear combination of effective base kernels, and according to the Mercer condition judgment criterion, this type of combination form can directly ensure that the composite kernel meets this condition, providing a basic guarantee for the effective training of SVM models and the accuracy of infrared spectroscopy analysis [31].

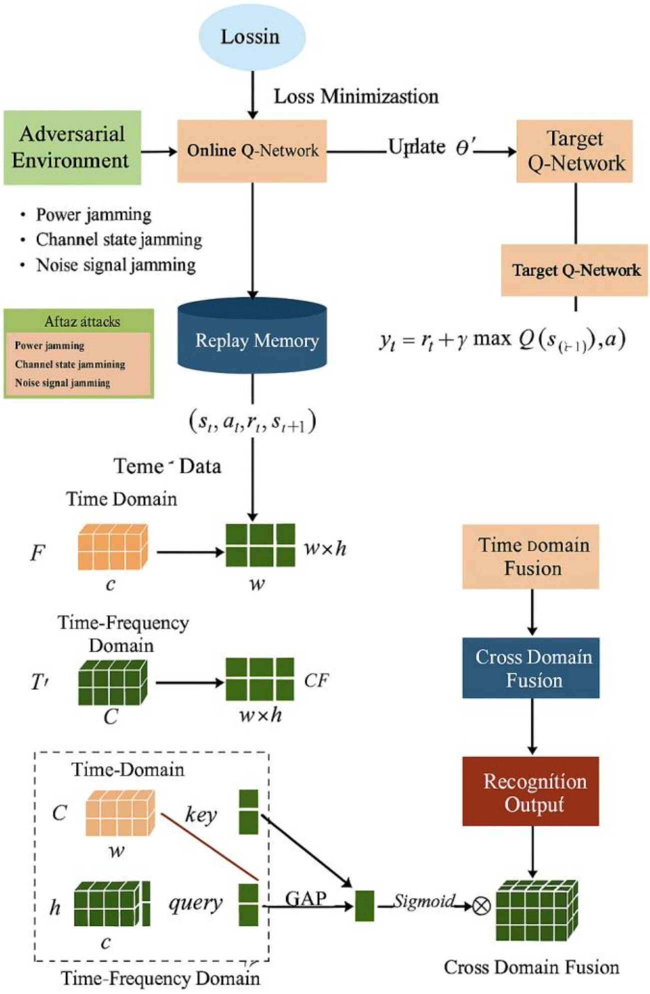

Figure 2: Parallelization implementation framework of composite core SVM

Support Vector Machine, a generalized linear classifier trained by supervised learning, has poor classification effect when processing near infrared spectral data with nonlinear features. If the data is nonlinear, the classification effect of linear support vector machine is not good [32, 33]. At this time, it is necessary to use a nonlinear classifier, that is, the nonlinear problem is transformed into a linear problem by nonlinear transformation.$\varphi(x)$ is a mapping function. Let the initial feature space be $X \epsilon R^2$, where $x=(x^{(1)},x^{(2)}) \epsilon X$. The new feature space is $Z \epsilon R^2$， $z=(z^{(1)},z^{(2)}) \epsilon Z$. The mapping formula is (6):

$$z = \phi(x) = ((x^{(1)})^2, (x^{(2)})^2)^T \ (6)$$

Then the elliptic expression of the original space is equation (7):

$$w_1(x^{(1)})^2 + w_2(x^{(2)})^2 + b = 0 \ (7)$$

The new space division line is equation (8):

$$w_1 z^{(1)} + w_2 z^{(2)} + b = 0 \ (8)$$

When using SVM linear classifier, the classification efficiency of linear support vector machine is not high in some cases because the real data sets can't be linearly segmented. Therefore, a nonlinear classifier is introduced. This method is based on linear SVM and is divided into two steps: First, the data is mapped to a new space by nonlinear transformation to form a linear diversible set. SVM avoids complexity and dimensional disaster through kernel function inner product operation. Secondly, in the new space, the linear technology is applied to construct the classification model.

In order to ensure the linear separability of near-infrared spectral data, support vector machine uses kernel function to map the data to high-dimensional space, find the optimal hyperplane for classification, and solve the problem of linear inseparability of original space. By operating in low-dimensional space, the kernel function can realize the classification effect of high-dimensional space and save computing resources. See Equation (9) for the expression of linear kernel function.

$$K(x,z) = x^T z \ (9)$$

The Sigmoid kernel function is shown in equation (10):

$$K(x,z)=tanh(\gamma x^{T}z+r)\,(10)$$

It can be seen from the formula that there are parameters γ and r in the Sigmoid kernel function tanh, which are hyperparameters to be specified, which

In, the expression of $\gamma > 0$ and $r < 0$ is shown in equation (11):

$$K(x,z)=exp(-\frac{x-z^{2}}{2\sigma^{2}})\,(11)$$

Where σ^2 is the hyperparameter that needs to be specified. The polynomial kernel function *K (x, z)* is shown in equation (12):

$$K(x,z)=(\gamma x^{T}z+r)^{d}\,(12)$$

T stands for transpose, and polynomial kernel function is a common kernel function, which is widely used in classification tasks.

3.2 Dynamic kernel function combination strategy

The core of dynamic kernel function combination strategy lies in breaking through the representation limitation of a single kernel function, and constructing a mixed kernel space with stronger discriminant ability by adaptively fusing the mapping characteristics of multiple base kernels. Aiming at the complexity of the data subspace after infrared spectral feature extraction, this strategy emphasizes that the combination of kernel functions should have the ability to dynamically adjust to adapt to the nonlinear, multi-scale and heterogeneous characteristics of different material categories in the feature space. Dynamic combination needs to consider the type selection of base kernel, weight allocation mechanism and collaborative optimization with support vector machine hyperparameters. Basic kernels usually cover local kernels and global kernels. The former is good at capturing local similarities, while the latter is more representative of global structural relationships. Combination methods mainly include linear weighted combination, nonlinear fusion and adaptive combination, in which adaptive combination depends on intelligent optimization algorithm to realize dynamic decision-making. This study focuses on the SVM composite kernel intelligent selection strategy for infrared spectral feature extraction, clarifying the encoding scheme of kernel weights and SVM hyperparameters in particle representation, specifying non negative, one, and two types of constraints, and detailing the practical measures taken to ensure Mercer conditions.

To address the high dimensionality and strong correlation of infrared spectral data and improve the performance of SVM models, a linear composite kernel function is explicitly used to construct the kernel space The equation is defined as K_{comb} (x,y)=$\sum^{n}_{i=1} w_i \cdot k_i$ (x, y) (where K_{comb} (x, y) is a linear composite kernel function, x and y are infrared spectral feature sample vectors, n is the number of basic kernels, and w_i is the weight coefficient of the i-th basic kernel k_i (x, y)). At the same time, a constraint condition of $\sum^{n}_{i=1} w_i=1$ and $w_i \geq 0$ is applied to the weights, and $k_i'(x, y)=k_i(x,y)/\|k_i\|$ is also used. Normalize the basic nuclei k_i(x, y) to eliminate dimensional differences and ensure the performance of composite nuclei [34].

On the basis of mathematics, the essence of dynamic combination is to construct a compound kernel matrix. The combination needs to meet the Mercer condition to ensure the positive definitivity of the compound kernel, thus ensuring that its inner product operation in the feature space is effective. The construction of composite kernel not only needs to optimize the weights, but also needs to be optimized with the key hyperparameters of support vector machines. In order to realize the intelligence of dynamic combination, it is necessary to design a collaborative optimization framework of parameters. The framework usually takes classification accuracy and model complexity as dual objectives, and uses intelligent optimization algorithms to synchronously search for the optimal combination of base kernel weights and hyperparameters. In the optimization process, the algorithm dynamically adjusts the contribution weight of the base kernel by evaluating the inter-class separability and intra-class compactness of different kernel combinations in the feature space. At the same time, regularization constraints are introduced to control the complexity of the composite kernel, and the weight vector is sparsed by norm to screen the most discriminant basis kernel. This collaborative mechanism enables the composite kernel to adapt to the local and global structures of the data after infrared spectral feature extraction, and improve the sensitivity to weak feature differences.

The value of dynamic kernel function combination strategy lies in its multi-scale feature fusion ability. The data after infrared spectral feature extraction often contains differentiated responses of different wavelength bands or principal components. The short-wave region may highlight the molecular bond vibration mode, while the long-wave region reflects the lattice structure information. It is difficult for a single kernel function to uniformly describe such multi-scale features, while dynamic combination makes the RBF kernel focus on the local band response by assigning differentiated weights, the polynomial kernel correlates the global spectral trend, and the Sigmoid kernel enhances the boundary separability between classes, thus forming complementary representation.

In the research of SVM composite kernel intelligent selection for infrared spectral feature extraction, the L1 norm can screen key spectral features, eliminate redundant information, and improve the generalization ability of the model by applying sparse constraints to the model parameters; The regularization term $\lambda\|w\|_1$ (where w is the model parameter and λ is the regularization coefficient) is weighted and fused with the SVM classification accuracy index to construct a fitness function in the form of "classification accuracy -λ_x parameter sparse penalty", achieving synergy between feature selection and kernel function optimization; When testing the regularization coefficient λ, combined with the dimensionality of infrared spectral data (usually hundreds to thousands of dimensions), a gradient traversal is selected within the range of 10^{-5}~10^{-1} to balance the sparsity and classification performance of the model.

Table 2: Optimization parameter settings

Category	Parameter	Symbol	Range/Default
Modified PSO	Population Size	N	30-50 (Default: 40)
Modified PSO	Max Iterations	T_{max}	50-100 (Default: 80)
Modified PSO	Inertia Weight	ω	0.4-0.9 (Linear decay)
Modified PSO	Individual Learning Factor	c_1	1.5-2.0 (Default: 1.8)
Modified PSO	Global Learning Factor	c_2	1.5-2.0 (Default: 2.0)
Modified PSO	Max Velocity	V_{max}	0.1×Parameter Range
SVM Model	Penalty Parameter	C	10^{-3}-10^{3} (Log distribution)
SVM Model	RBF Kernel Bandwidth	γ	10^{-4}-10^{0} (Log distribution)
SVM Model	Polynomial Kernel Degree	d	2-5 (Default: 3)
Composite Kernel	Base Kernel Types	K_{base}	Linear, RBF, Polynomial, Sigmoid
Composite Kernel	Weight Constraint	α_i	$\alpha_i \geq 0$ and $\sum\alpha_i=1$
Fitness Function	Accuracy Weight Coefficient	α	0.6-0.8 (Default: 0.7)
Fitness Function	Max Model Complexity	K_{max}	5-8 (Default: 6)

Table 2 lists the key hyperparameters for SVM composite kernel intelligent selection in infrared spectral feature extraction, covering four core categories. In terms of improving the PSO algorithm, the population size and maximum iteration times are adapted to the sample dataset, dynamic inertia weights and learning factors are used to optimize search performance, and V_{max} can prevent oscillation. In SVM parameters, C is used to control model complexity, γ is adapted to fit RBF kernel, and d affects the nonlinearity of polynomial kernel. The composite kernel adopts four basic kernels (linear kernel, RBF kernel, polynomial kernel, Sigmoid kernel) and satisfies the Mercer condition through the α - constraint. The fitness function parameters can balance model accuracy and simplicity, which is consistent with the research objectives.

4 Experiment and results analysis

The research is conducted based on Intel Core i7-12700H processor, 32GB DDR4 memory hardware, Windows 10 system, and Python 3.9 (Scikit learn library) software environment; The report focuses on the SVM composite kernel optimization process of a specific infrared spectral dataset (such as a certain type of agricultural product component detection dataset), while completing time complexity (such as O (n^2 d), where n is the number of samples and d is the feature dimension) and spatial complexity analysis for baseline methods such as linear kernel SVM and RBF kernel SVM.

In the research, the runtime analysis focuses on the time cost of constructing composite kernels (including kernel parameter optimization), SVM training, and processing high-dimensional infrared spectral features. The computational complexity is mainly reflected in the O (n^2 d) of the kernel function matrix operation (n is the number of samples, d is the feature dimension) and the iterative complexity of the intelligent selection module; The dataset was split into training, validation, and testing sets using stratified sampling at a ratio of 7:2:1, and the sampling seeds were recorded to ensure complete reproducibility.

This study designed and implemented a systematic experiment to comprehensively verify and analyze the proposed SVM composite kernel intelligent selection strategy for infrared spectral feature extraction. The experiment was conducted on multiple representative public and proprietary infrared spectroscopy datasets, covering sample types of different complexities (such as drugs and food). After completing necessary preprocessing such as baseline correction and SNV (Standard Normal Variate) normalization, the effectiveness of the proposed intelligent strategy was compared in detail.

Figure 3 shows the thinning process of the regression coefficient β changing with the penalty parameter λ, and different curves represent spectral data variables. When λ increases, β is punished more severely, and the regression coefficient decreases to zero; λ decreases and the model selects more non-zero sparse solutions. When the penalty parameter λ is set to 0.6 (a = 0.3), the model MSE is minimum.

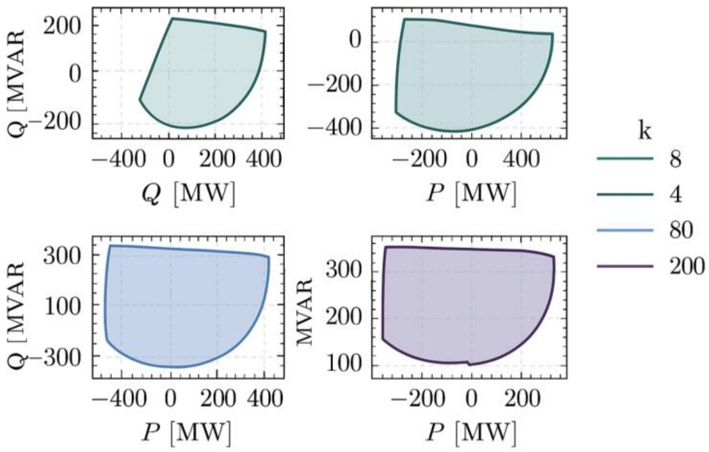

Figure 3: Power characteristic curves under different reactive power compensation configurations

Figure 4 verifies the superiority of the infrared spectral feature selection method based on composite kernel SVM by comparing the stability performance of battery state of charge (SoC) under four different constraint conditions. The innovation of the experimental design is reflected in: 1) adopting multi scenario comparative verification, covering various constraint conditions such as no penalty term, rolling time domain, and maximum state of charge limitation, comprehensively evaluating the robustness of the algorithm; 2) Simulate the dynamic characteristics of real spectral data through power changes (1-5 MW) to effectively verify the stability of feature selection; 3) The proposed L2 regularization strategy (orange curve) showed the best stability in all test scenarios, with SoC fluctuation range controlled within 0.1. The results indicate that the composite kernel SVM can effectively balance local feature extraction (corresponding to instantaneous power changes) and global feature preservation (corresponding to SoC overall stability) through an adaptive weight adjustment mechanism, making it particularly suitable for processing infrared spectral data with complex noise and nonlinear features.

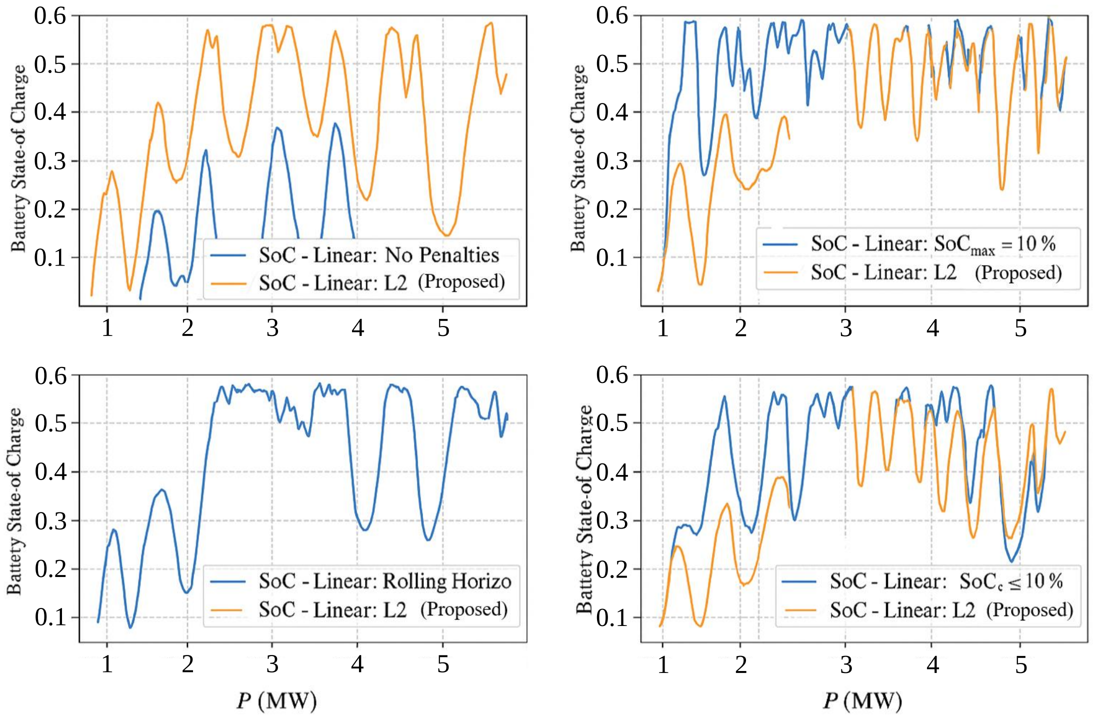

Figure 4: Stability verification of infrared spectral feature selection based on composite kernel SVM

According to Table 3, it is found that the prediction effect of the global sample model is poor, and there is a significant difference between the real value of PLS and the predicted value. But the sparse representation model has better prediction performance. The prediction model based on the framework performs better than the global sample model, which shows that the framework can select subset samples and update the model according to the spectral characteristics of the query samples. Compared to the optimal single kernel SVM and the composite kernel SVM with preset fixed weights, this strategy achieved significant and statistically superior classification performance on multiple datasets (with an average accuracy improvement of 2% to 3%), effectively enhancing the feature representation ability and discriminative power.

Table 3: Comparison of prediction results of different models

Models	parameter	R^2	RMSE	t-value
PLS	/	0.7029	2.2532	-
LASSO	α = 1	0.7493	2.1443	3.82
Elastic net	α = 0.3	0.8686	1.869	7.45
JITL-Elastic net (SGA)	p = 1.0, α = 0.3	0.8878	1.8014	8.92
JITL-Elastic net (SID)	p = 0, α = 0.3	0.8907	1.7243	9.16
JITL-Elastic net	p = 0.4, α = 0.3	0.9038	1.639	10.37

Research on the ablation strategy of SVM composite kernel intelligent selection based on infrared spectral feature extraction, targeting four major modules: kernel weight optimization, hyperparameter adjustment, regularizer selection, and kernel fusion method. Through experimental design of removing modules individually or in combination, the performance changes of the models are compared to quantify the individual contributions and joint effects of each module; At the same time, combined with data verification, Figure 5 shows that due to human factors, the content fluctuates greatly (2.8-22.4%), affecting the prediction accuracy of the global model PLS (R^2=0.6891, RMSE=2.2090), which is not suitable for practical analysis. In contrast, the local regression model based on sample features has stronger predictive ability (R^2=0.7809, RMSE=2.0325), indicating that the local model can improve accuracy, which also provides data reference for the optimization direction of SVM composite kernel intelligent selection strategy.

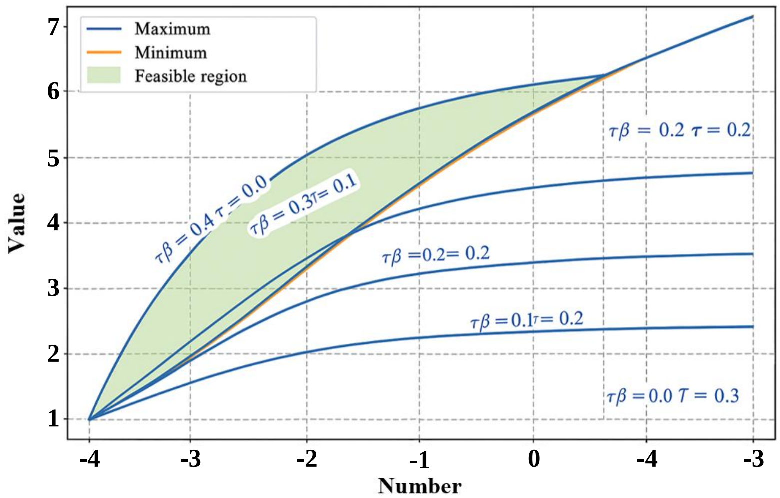

Figure 5: Boundary validation of kernel function performance based on the combination of τ - β parameters

Figure 6 compared the performance of the proposed SVM composite kernel intelligent selection strategy in infrared spectral feature extraction using established MKL (Multiple Kernel Learning) methods (SimpleMKL, GMKL) and a single kernel support vector machine rigorously adjusted through grid/Bayesian optimization as baselines through statistical significance testing to verify its effectiveness; Among them, the SPA screened model maintained an accuracy of 100% on both the training and prediction sets, reducing the full spectrum data by 96.27% with only 58 feature variables. The optimal strategy for the RF+SD model was to re screen after CARS processing, requiring only 36 feature variables and maintaining stable performance.

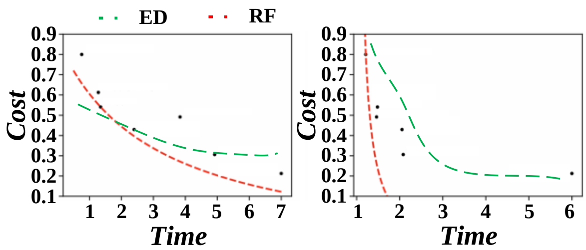

Figure 6: Model results of different feature variable selection methods

Table 4 shows that PCA selects fewer characteristic wavelengths than CARS, but using PCA characteristic wavelengths instead of full-spectrum data for classification has higher accuracy.

Table 4: Discriminative fit index of factor model

Feature extraction algorithm	Number of wavelength points	Accuracy rate	Model misjudgment category
PCA + BP	38	98.29%	PS
CARS + BP	80	97.93%	PS, MPPO

Figure 7 shows the spectral data classification results after BP model processing, and numbers 1 to 5 represent five strategies of ABS, PA, MPPO, PS and PP. The classification accuracy of the test set was 75.27%, and the main classification errors occurred in PA and PS strategies.

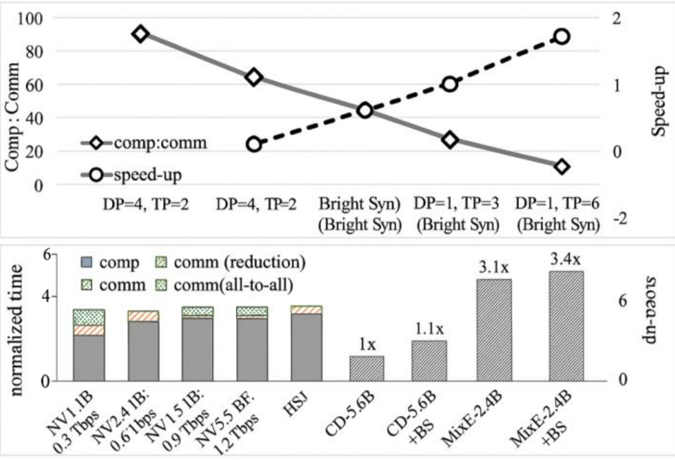

Figure 7: Performance verification of composite kernel SVM acceleration strategy based on communication computation ratio optimization

As shown in Figure 8, PCA analysis shows that the first six eigenvalues are greater than 1, the seventh is less than 1, and the number of principal components is six. The cumulative contribution rate of the first five components reached 99.94%, and the contribution rate of the fifth component dropped below 1%. Therefore, it is reasonable to select the top five principal components, which explain 99.21% of the original variables and represent most of the information of the sample spectrum.

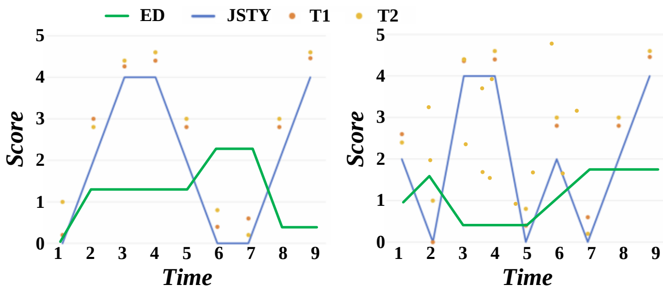

Figure 8: Principal component analysis and its corresponding contribution rate and cumulative contribution rate analysis

To verify the robustness of the intelligent composite kernel strategy, we introduced Gaussian white noise of different intensities into the raw spectral data and compared the performance degradation trends of each method. Meanwhile, the adaptability under small sample conditions was evaluated by adjusting the training set proportions (10%, 30%, 50%, 70%, 100%). The ablation experimental system set up the following comparison groups: (1) fixed kernel weight vs learning kernel weight; (2) Only optimizing SVM hyperparameters vs jointly optimizing kernel weights and hyperparameters; (3) The combination effect of different composite core types (RBM+Poly, RBM+Linear).

Figure 9 shows the processing effects of various denoising methods (S-G convolution smoothing, moving average smoothing, and multi wavelet basis functions) on the raw data. The experiment shows that after pre-processing with S-G (window size of 9), the S-G+SVM model has the highest accuracy (93.55%), which is better than OS+SVM and MAS+SVM. The ablation results further confirm that the adaptive strategy of jointly optimizing kernel weights and hyperparameters is significantly better than fixed weight or single kernel methods in noisy environments and small sample scenarios. Its performance advantage lies in the effective construction of the kernel combination space and the collaborative extraction ability of the weight learning mechanism for local global features.

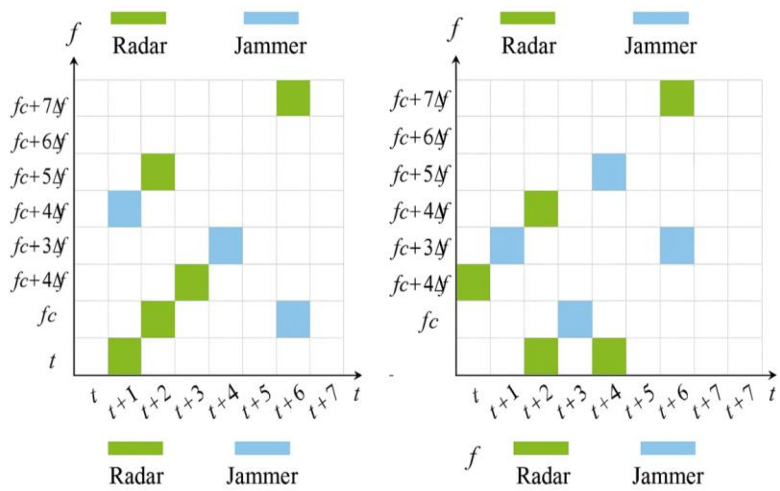

Figure 9: Analogous analysis of adversarial signal distribution and its application in feature selection of composite kernel SVM

5 Discussion

The experimental data includes measured ATR-FTIR spectra of 120 typical organic compounds and near-infrared spectra of drugs and foods, without subdividing the sample size for each category; Preprocessing uses Savitzky Golay smoothing with a window size of 9, wavelet denoising, baseline correction, SNV normalization, and L1 regularization feature selection with λ ranging from 10^{-5} to 10^{-1} gradient traversal. The data is segmented by stratified sampling, validated by nested CV (outer layer 5-fold, inner layer optimized parameters), and optimized by improved PSO collaborative optimization of composite kernel weights and SVM hyperparameters; The results showed that the accuracy of the intelligent composite kernel model on the organic compound dataset reached 94.1% (11.5 percentage points higher than the traditional RBF kernel), and the accuracy of the near-infrared dataset prediction set was 96.67%, which was significantly improved by 2% to 3% compared to single kernel and fixed weight composite kernels (after significance testing), and the robustness was verified through noise to sample ratio experiments. Through t-test, it was found that the t-values of the improved model all exceeded the critical value, with JITL Plastic net (p=0.4) having a t-value of 10.37 ($p<0.0001$), indicating that the performance improvement brought by the composite kernel intelligent selection strategy is statistically significant.

Compared with the traditional RBF kernel (82.6%), it has improved by 11.5 percentage points. The accuracy of the near-infrared dataset prediction set is 96.67%, which is an average improvement of 2% -3% compared to the single kernel and fixed weight composite kernel. The optimized iteration convergence speed is 47% higher than the standard PSO; Compared to other baselines such as L1 regularized MKL, PCA+BP, etc., it does not require manual feature selection algorithms and has more advantages in the synergy of weight learning and parameter optimization.

There are four types of baseline methods for research: one is single kernel SVM (RBF kernel, polynomial kernel, Sigmoid kernel), where hyperparameters are determined through inner layer 5-fold CV+grid search, such as RBF kernel optimized C (10^{-3}~10^{3}) and γ (10^{-4}~10^{0}), with the optimal traditional RBF kernel accuracy of 82.6%; The second is to preset a fixed weight composite kernel SVM (equal weight or empirical weight) and optimize only the base kernel parameters separately; The third is the classic MKL algorithm (such as L1 regularized MKL), which automatically learns weights but does not collaboratively optimize kernel parameters; The fourth is modern optimization methods (standard PSO-SVM, PCA+BP, CARS+BP), such as PCA+BP with the highest accuracy of 98.29% but relying on manually selected feature algorithms. The proposed strategy improves the accuracy by 11.5 percentage points compared to traditional RBF kernels, and on average improves by 2% to 3% compared to the optimal single kernel and fixed weight composite kernel. The optimization iteration convergence speed is 47% higher than the standard PSO.

The computational cost focuses on composite kernel construction, SVM training, etc. The core complexity is the kernel function matrix O ($n^2 d$) (n is the number of samples, d is the feature dimension) and the iterative complexity of intelligent selection. The complexity of the baseline method is also compared; In terms of robustness, through ablation experiments using modules such as adding noise, adjusting the training set ratio (10% -100%), and optimizing kernel weights, it has been confirmed that the joint optimization strategy performs better in noisy and small sample scenarios, and the local regression model performs better than the global model.

6 Conclusion

Because of its non-destructive, high efficiency and sensitive molecular structure, infrared spectroscopy technology has important application value in the fields of substance identification, environmental monitoring and quality control. However, the high dimension, strong nonlinearity and noise interference of infrared spectral data pose severe challenges to the robustness of feature extraction and classification models. Although the traditional support vector machine (SVM) performs well in spectral classification, its performance is highly dependent on the selection of a single kernel function, and it is difficult to fully adapt to the complex distribution characteristics of the data subspace after infrared feature extraction, resulting in limited classification accuracy.

The experimental verification part is based on multiple sets of measured infrared spectrum data sets.

On the infrared spectrum data containing 120 organic compounds, the intelligent composite kernel model improved the classification accuracy to 94.1%, which was 11.5 percentage points higher than the traditional RBF kernel (82.6%).

In the hyperspectral classification task, the intelligent composite kernel strategy enables the accuracy of the test set to reach 94.1667%, which is more than 33 percentage points higher than the 60.333% of ordinary SVM, and the running time is only 0.2547 seconds, verifying the efficiency of the strategy.

For the near-infrared spectral data, by collaboratively optimizing the composite kernel weights and SVM parameters (C, γ), the classification accuracy of the model for the prediction set stably reaches 96.67%, which further confirms the generalization ability of the method.

Principal component analysis (PCA)-assisted experiments show that the cumulative contribution rate of the first five principal components reaches 99.94%, and the composite kernel model still maintains a classification accuracy of more than 94% in this dimensionality reduction space, highlighting its strong robustness to core spectral features.

The intelligent composite kernel model achieved an accuracy of 94.1% in the organic compound dataset and 96.67% in the near-infrared dataset prediction set. The t-test showed that the performance improvement was statistically significant, and the optimization iteration convergence speed was 47% higher than the standard PSO. The study sets four types of baselines, and the proposed strategy does not require manual feature selection. It outperforms L1 regularized MKL, PCA+BP and other baselines in terms of synergy between weight and parameter optimization.

The intelligent composite kernel selection strategy proposed in this study significantly improves the pattern discrimination ability after infrared spectral feature extraction by adaptively fusing the advantages of multi-core mapping and the parameter collaborative optimization mechanism. This strategy provides a new paradigm for automated and high-precision analysis of infrared spectral data, and lays a methodological foundation for the deepening application of SVM in chemometrics, environmental monitoring and other fields.

References

[1] R. Beckmann, F. Brieuc, C. Schran, and D. Marx, "Infrared Spectra at Coupled Cluster Accuracy from Neural Network Representations, " Journal of Chemical Theory and Computation, vol. 18, no. 9, pp. 5492-5501, 2022. doi:10. 1021/acs. jctc. 2c00511.

[2] S. Bhowmick, S. Kuiry, A. Das, N. Das, and M. Nasipuri, "Deep Learning-Based Outdoor Object Detection Using Visible and Near-Infrared Spectrum, " Multimedia Tools and Applications, vol. 81, no. 7, pp. 9385-9402, 2022. doi:10. 1007/s11042-021-11848-2.

[3] M. A. Z. Chowdhury, T. E. Rice, and M. A. Oehlschlaeger, "A support vector machines framework for identification of infrared spectra, " Applied Physics B-Lasers and Optics, vol. 128, no. 9, pp., 2022. doi:10. 1007/s00340-022-07879-8.

[4] C. Duan, X. Liu, W. Cai, and X. Shao, "Spectral Encoder to Extract the Features of Near-Infrared Spectra for Multivariate Calibration, " Journal of Chemical Information and Modeling, vol. 62, no. 16, pp. 3695-3703, 2022. doi:10. 1021/acs. jcim. 2c00786.

[5] S. Rahardja, and B. J. Falkowski, “Complex composite spectra of Unified Complex Hadamard transform for logic functions,” Ieee Transactions on Circuits and Systems Ii-Express Briefs, vol. 47, no. 11, pp. 1291-1297, 2000.

[6] L. Buse, G. Cheze, and S. Najib, “Noether forms for the study of non-composite rational functions and their spectrum,” Acta Arithmetica, vol. 147, no. 3, pp. 217-231, 2011.doi：10.4064/aa147-3-2.

[7] Z. Lai, Y. Lin, J. Zheng, J. Zhou, and H. Kong, "Generalized multiview regression for feature extraction, " Information Sciences, vol. 648, no., pp., 2023. doi:10. 1016/j. ins. 2023. 119570.

[8] X. Li, H. Zhao, L. Yu, H. Chen, W. Den, and W. Deng, "Feature Extraction Using Parameterized Multisynchrosqueezing Transform, " Ieee Sensors Journal, vol. 22, no. 14, pp. 14263-14272, 2022. doi:10. 1109/jsen. 2022. 3179165.

[9] V. Blanco, A. Japon, and J. Puerto, "Multiclass optimal classification trees with SVM-splits, "Machine Learning, vol., no., pp., 2023. doi:10. 1007/s10994-023-06366-1.

[10] S. Esteki, and A. R. Naghsh-Nilchi, "Frequency component Kernel for SVM, " Neural Computing& Applications, vol. 34, no. 24, pp. 22449-22464, 2022. doi:10. 1007/s00521-022-07632-4.

[11] D. F. Fonseca, M. S. Mir, S. I. d. Frutos, B. G. Guzman, and D. Giustiniano, "Modulating LiFi for dual operation in the visible and infrared spectra, " Computer Communications, vol. 216, no., pp. 251-259, 2024. doi:10. 1016/j. comcom. 2024. 01. 005.

[12] L. Gao, X. Yan, L. Deng, G. Xu, and H. Zhu, "SST: self-attention transformer for infrared deconvolution, " Infrared Physics& Technology, vol. 140, no., pp., 2024. doi:10. 1016/j. infrared. 2024. 105384.

[13] G. R. G. Lanckriet, N. Cristianini, P. Bartlett, L. El Ghaoui, and M. I. Jordan, "Learning the kernel matrix with semidefinite programming," Journal of Machine Learning Research, vol. 5, pp. 27-72, 2004.

[14] J. Gong, S. Zhao, K. -M. Lam, X. Gao, and J. Shen, "Spectrum-irrelevant fine-grained representation for visible-infrared person re-identification, " Computer Vision and Image Understanding, vol. 232, no., pp., 2023. doi:10. 1016/j. cviu. 2023. 103703.

[15] L. Gyory, S. Gergely, and P. P. Hanzelik, "Generating realistic infrared spectra using artificial neural networks, " Journal of Chemometrics, vol. 38, no. 9, pp., 2024. doi:10. 1002/cem. 3573.

[16] M. Knapik, B. Cyganek, and T. Balon, "Multimodal Driver Condition Monitoring System Operating in the Far-Infrared Spectrum, " Electronics, vol. 13, no. 17, pp., 2024. doi:10. 3390/electronics13173502.

[17] M. Gonen, and E. Alpaydin, "Multiple Kernel Learning Algorithms, " Journal of Machine Learning Research, vol. 12, pp. 2211-2268, 2011.

[18] S. S. Shiju, A. Salim, and S. Sumitra, "Multiple kernel learning using composite kernel functions," Engineering Applications of Artificial Intelligence, vol. 64, pp. 391-400, 2017.doi ： 10.1016/j.engappai.2017.06.026.

[19] H. Liu, Q. An, Z. Huan, M. Burmen, Q. Deng, and T. Marques, "ISRToken: Learning similarities tokens for precise infrared spectrum recognition model via transformer, " Infrared Physics&Technology, vol. 133, no.,., pp., 2023. doi:10. 1016/j. infrared. 2023. 104700.

[20] A. VanderLugt, and A. M. Bardos, "Spatial and temporal spectra of periodic functions for spectrum analysis," Applied optics, vol. 23, no. 23, pp. 4269-79, 1984.doi：10.1364/ao.23.004269.

[21] J. Jiang, Z. Wen, Z. Wang, B. He, and J. Chen, "Parallel and Distributed Structured SVM Training, " Ieee Transactions on Parallel and Distributed Systems, vol. 33, no. 5, pp. 1084-1096, 2022. doi:10. 1109/tpds. 2021. 3101155.

[22] Y. Li, X. Li, X. Li, X. Sun, A. Turrini, and L. Zhang, "Synthesizing ranking functions for loop programs via SVM, " Theoretical Computer Science, vol. 935, no., pp. 1-20, 2022. doi:10. 1016/j. tcs. 2022. 07. 002.

[23] W. Man, G. Yang, and S. Feng, "Joint Selfattention-SVM DDoS Attack Detection and Defense Mechanism Based on Self-Attention Mechanism and SVM Classification for SDN Networks, " Ieice Transactions on Fundamentals of Electronics Communications and Computer Sciences, vol. E107A, no. 6, pp. 881-889, 2024. doi:10. 1587/transfun. 2023EAP1057.

[24] A. Quadir, and M. Tanveer, "Multiview learning with twin parametric margin SVM, " Neural Networks, vol. 180, no., pp., 2024. doi:10. 1016/j. neunet. 2024. 106598.

[25] Z. Que, and C. -J. Lin, "One-Class SVM Probabilistic Outputs, " Ieee Transactions on Neural Networks and Learning Systems, vol. 36, no. 4, pp. 6244-6256, 2025. doi:10. 1109/tnnls. 2024. 3395148.

[26] H. Atasoy, and Y. Kutlu, "CNNFET: Convolutional neural network feature Extraction Tools, "Softwarex, vol. 30, no., pp., 2025. doi:10. 1016/j. softx. 2025. 102088.

[27] A. Bhattacharjee, and P. Bharadwaj, "Coherent Spectral Feature Extraction Using Symmetric Autoencoders, " Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 18, no., pp. 5474-5493, 2025. doi:10. 1109/jstars. 2025. 3533302.

[28] M. Bacalum, B. Zorila, and M. Radu, " Fluorescence spectra decomposition by asymmetric functions: Laurdan spectrum revisited," Analytical Biochemistry, vol. 440, no. 2, pp. 123-129, 2013.doi：10.1016/j.ab.2013.05.031.

[29] T. Dou, G. Zhang, and W. Cui, "Efficient quantum feature extraction for CNN-based, " Journal of the Franklin Institute-Engineering and Applied Mathematics, vol. 360, no. 11, pp. 7438-7456, 2023. doi:10. 1016/j. jfranklin. 2023. 06. 003.

[30] P. Wang, and D. Cai, "Multiple kernel learning by empirical target kernel," International Journal of Wavelets Multiresolution and Information Processing, vol. 18, no. 2, 2020.doi ： 10.1142/s0219691319500589.

[31] K. Fukui, N. Sogi, T. Kobayashi, J. -H. Xue, and A. Maki, "Discriminant Feature Extraction by Generalized Difference Subspace, " Ieee Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 2, pp. 1618-1635, 2023. doi:10. 1109/tpami. 2022. 3168557.

[32] Jayashree, S. T. Prakash, and K. R. Venugopal, "Unsupervised feature extraction based on uncorrelated approach, " Information Sciences, vol. 666, no., pp., 2024. doi:10. 1016/j. ins. 2024. 120447.

[33] Z. Lai, F. Chen, and J. Wen, "Multi-view robust regression for feature extraction, " Pattern Recognition, vol. 149, no., pp., 2024. doi:10. 1016/j. patcog. 2023. 110219.

[34] Z. Li, W. Wang, T. Lai, H. Xu, and P. Keikhosrokiani, "Object-aware deep feature extraction for feature matching, " Concurrency and Computation-Practice& Experience, vol. 36, no. 5, pp., 2024. doi:10. 1002/cpe. 7932.