Research Article

A Reliable Method for Identification of Antibiotics by Terahertz Spectroscopy and SVM

Jin Guo,[1,2] **Hu Deng,**[1] **Quancheng Liu,**[1] **Linyu Chen,**[1] **Zhonggang Xiong,**[1,3] **and Liping Shang** [1]

[1]*School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China*
[2]*Tianfu College of Southwestern University of Finance and Economics, Mianyang 621000, China*
[3]*School of Mechanical Engineering, Guilin University of Aerospace Technology, Guilin 541004, China*

Correspondence should be addressed to Liping Shang; 121992278@qq.com

Given the extensive use of antibiotics at present, the identification of antibiotics and production quality monitoring are of high importance. However, conventional antibiotic identification methods have a low sensitivity and a long detection time. Here, we propose an identification method that combines terahertz (THz) spectroscopy and chemometric technology. THz time-domain spectroscopy (THz-TDS) was performed for sixteen types of antibiotics, including β-lactam, cephalosporins, macrolides, and tetracyclines. The absorption spectra within the frequency range of 0.2–1.5 THz were calculated. For dimensionality reduction, principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) were implemented, respectively. The data after dimensionality reduction were input into a support vector machine (SVM). The model parameters were optimized through grid search (GS), genetic algorithm (GA), and particle swarm optimization (PSO) methods, and the optimal identification results were obtained after comparison across these methods. Experiments indicate a differentiation of the THz absorption spectra among the sixteen types of antibiotics. After dimensionality reduction, the training time of the model significantly decreased. The use of the t-SNE-PSO-SVM model achieved the highest average accuracy on the prediction set, which was 99.91%. Thus, our study does not only confirm that the t-SNE-PSO-SVM model proves to be a reliable method for antibiotics identification, but also confirms that the combination of THz-TDS and chemometric pattern recognition has great potential for drug detection.

1. Introduction

Antibiotics are a large class of antibacterial chemical substances that occur naturally or are semisynthetic or synthetic. There are a great variety of antibiotics, which are further divided into seven major classes, namely, tetracyclines, macrolide antibiotics, aminoglycosides, peptide antibiotic, lincosamides, streptogramins, and β-lactam antibiotics [1]. Nearly every bacterium has a specific antibiotic against it. Antibiotics are mainly used to treat various types of bacterial infections in humans or in livestock to promote their growth. However, the problem of antibiotic residues has become increasingly severe due to excessive antibiotics use. Therefore, antibiotics detection and identification is of high importance [2]. In the past few decades,

numerous efforts have been made to develop analytical methods for qualitative or quantitative determination of antibiotics. Conventional methods for antibiotics detection mainly include high-performance liquid chromatography (HPLC) [3] and gas chromatography mass spectrometry (GC-MS) [4]. Although these chromatography-based techniques are sensitive and reliable, they are usually time-consuming. Capillary electrophoresis (CE) [5], immunochemistry [6], and enzyme-linked immunosorbent assay (ELISA) [7] can achieve high-accuracy detection of antibiotics. However, these procedures usually involve complex sample preprocessing, which needs to be done by well-trained professionals. The expensive costs of surface plasmon resonance (SPR) sensors [8] and Raman spectroscopy [9] have restricted their extensive application. Given the

above, it is necessary to establish a sensitive, fast, and reliable method for antibiotics detection [10].

The THz band has a wave frequency ranging between 0.1 and 10 THz, which is between the infrared and microwave frequencies. The THz waves are transient, safe (single photon energy, 4.1 meV) and highly penetrating, and have fingerprinting properties [11]. THz-TDS has already found extensive applications in biological tissue identification [12, 13], food and drug detection [14, 15], and explosive detection [16]. The vibration-rotation energy levels of such macromolecules as antibiotics are located within the THz band. As compared with other spectral detection methods, THz spectroscopy exhibits unique advantages. THz spectroscopy can detect not only molecular spinning and lattice vibration but also the inner structure and organizational features of the drugs. Limwikrant et al. [17] obtained the THz spectra of ofloxacin and complex of oxalic acid. Zhang et al. [18] analyzed the molecular vibration modes of piracetam and 3-hhydroxybenzoicacid. Zhang et al. [19] obtained the THz fingerprinting spectra of metronidazole, tinidazole, and ornidazole. Xie et al. [20] showed through DFT calculation that tetracycline had definitive THz absorption spectra at certain frequencies. Many studies have demonstrated the feasibility of applying THz spectroscopy to antibiotics detection. Qin et al. [21–24] applied THz spectroscopy to the detection of tetracycline hydrochloride and achieved good results. Massaouti et al. [25], Wang et al. [26], and Long et al. [27] used a similar method, the quantitative detection of antibiotics in the samples.

The THz technology offers extensive applications in the research fields of pesticide and antibiotic identification and residual pesticide detection [28]. Many studies have shown that THz spectroscopy is a feasible detection technique for antibiotics. Most of the studies focus on quantitative detection of antibiotics, though the use of THz spectroscopy to identify antibiotics has been rarely reported. Yan et al. [29] applied three-layer BP neural networks to identify absorption spectra of nine illicit drugs and six antibiotics, but the average identification rate was low. In this study, THz-TDS was applied to the detection of sixteen types of antibiotics, including penicillins, cephalosporins, macrolides, and tetracyclines. Then, the THz absorption spectra of these antibiotics were calculated. Dimensionality reduction was performed using principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE). Next, pattern recognition was performed using the GS-SVM, GA-SVM, and PSO-SVM models, and the best identification model was found by comparison. Thus, a novel method for fast and reliable antibiotics identification was established.

2. Materials and Methods

2.1. Terahertz Spectroscopy System.
A Z-3 Time-Domain Spectrometer (Zomega, USA) was used for the experiments. The system was located within a closed hood during the measurement process to reduce the influence of water vapor. The ambient temperature was controlled at 23°C, and humidity was below 2%. The THz-TDS parameters were set as follows: wavelength of the femtosecond laser system 800 nm, frequency 80 MHz, pump light intensity 100 mW, probe light intensity 20 mW, scan stroke 50 ps, useful spectral range 0.2–2 THz, and dynamic range above 70 db. All the experiment process was shown in Figure 1.

2.2. Sample Preparation.
Sixteen types of antibiotics, including β-lactam, tetracyclines, macrolides, and cephalosporins, were used. The name, class, and main ingredient of these antibiotics are shown in Table 1. First, the drug samples were ground in an agate mortar to avoid scattering of the THz waves caused by particle heterogeneity and also to increase the signal-to-noise ratio. Then a certain amount of the sample was weighed and placed on the automatic tablet press, with the pressure set to 2 tons and pressure maintenance time of 1 min. A digital caliper (precision 0.02 mm) was used to measure the thickness of the sample tablets. Thus, 40 samples were prepared for each of the 16 types of antibiotics and used to detect THz absorption spectra. Finally, the 640 samples were randomly divided into a training set (16 ∗ 30 samples) and a test set (16 ∗ 10 samples).

2.3. Data Processing.
The THz time-domain spectral information of the samples was obtained. The reflection peaks were removed by empirical mode decomposition [30]. Denoising was done by Savitzky–Golay filtering, followed by Fourier transform to convert the time-domain information to the frequency-domain information. The model was extracted based on the optical parameters proposed by Dorney et al. [31] and Duvillaret et al. [32], and the absorption coefficient of the sample $\alpha(\omega)$ was calculated.

$$\rho(\omega) = \frac{A_s}{A_r}, \tag{1}$$

$$\varphi(\omega) = \varphi_s - \varphi_r, \tag{2}$$

where $\rho(\omega)$ is the amplitude ratio; A_s is the signal amplitude of the sample; A_r is the signal amplitude of the reference; $\varphi(\omega)$ is the phase difference; φ_s is the phase of the sample; and φ_r is the phase of the reference.

The index of refraction and absorption coefficients are calculated using the formulae below:

$$n(\omega) = \varphi(\omega)\frac{c}{\omega d} + 1, \tag{3}$$

$$\alpha(\omega) = \frac{2}{d}\ln\left\{\frac{4n(\omega)}{\rho(\omega)[n(\omega)+1]^2}\right\}, \tag{4}$$

where $n(\omega)$ is the index of refraction; $\alpha(\omega)$ is the absorption coefficient; c is the speed of light in a vacuum; ω is the angular frequency; and d is the sample thickness.

3. Results and Discussion

3.1. Spectral Analysis.
THz-TDS was performed for the sixteen types of antibiotics shown in Table 1. The THz time-domain spectra thus obtained are shown in Figure 2(a), and
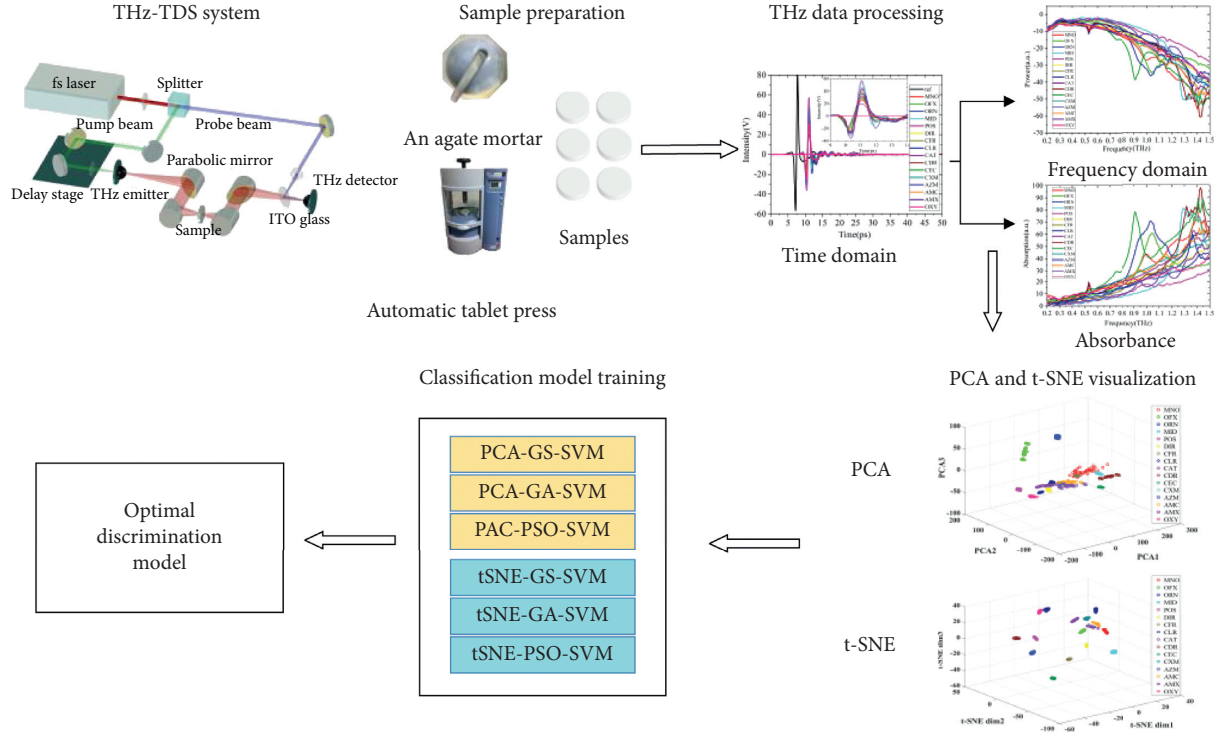
FIGURE 1: Experimental layout of terahertz spectroscopy system.

TABLE 1: Information for the sixteen types of antibiotics.

Drug name	Class	Main ingredient	Abbreviation	Molecular formula of the main ingredient
Minocycline hydrochloride capsule	Tetracyclines	Minocycline	MNO	$C_{23}H_{27}N_3O_7$
Oxytetracycline tablets	Tetracyclines	Oxytetracycline	OXY	$C_{22}H_{24}N_2O_9$
Ofloxacin tablets	Other	Ofloxacin	OFX	$C_{18}H_{20}FN_3O_4$
Ornidazole capsule	Other	Ornidazole	ORN	$C_7H_{10}ClN_3O_3$
Fosfomycin calcium and trimethoprim capsule	Other	Fosfomycin	POS	$C_3H_7O_4P$
Metronidazole Tablets	Other	Metronidazole	MID	$C_6H_9N_3O_3$
Azithromycin enteric-coated capsule	Macrolides	Azithromycin	AZM	$C_{38}H_{72}N_2O_{12}$
Dirithromycin enteric-coated tablet	Macrolides	Dirithromycin	DIR	$C_{42}H_{78}N_2O_{14}$
Clarithromycin sustained-release tablet	Macrolides	Clarithromycin	CLR	$C_{38}H_{69}NO_{13}$
Cefadroxil and trimethoprim capsule	Cephalosporins	Cefadroxil	CFR	$C_{16}H_{17}N_3O_5S$
Cefetamet pivoxil hydrochloride tablet	Cephalosporins	Cefetamet	CAT	$C_{20}H_{25}N_5O_7S_2$
Cefdinir dispersible tablet	Cephalosporins	Cefdinir	CDR	$C_{14}H_{13}N_5O_5S_2$
Cefixime dispersible tablet	Cephalosporins	Cefixime	CEC	$C_{16}H_{15}N_5O_7S_2$
Cefuroxime axetil tablet	Cephalosporins	Cefuroxime	CXM	$C_{16}H_{16}N_4O_8S$
Potassium clavulanate	β-lactam	Clavulanic acid	AMC	$C_8H_9NO_5$
Amoxicillin capsule	β-lactam	Amoxicillin	AMX	$C_{16}H_{19}N_3O_5S$

on this basis, the frequency-domain spectra and absorption spectra were calculated. The spectra corresponding to the frequency from 0.2 to 1.5 THz are shown in Figures 2(b) and 2(c). The sixteen types of antibiotics were barely differentiated by the time-domain and frequency-domain spectra. Some of the antibiotics shared the same absorption peaks, and the antibiotics could not be differentiated by the spectral features alone. To solve this problem, we introduced chemometric pattern recognition and established identification models.

3.2. Visualization of the Sample Classification. PCA can reduce a large number of intercorrelated indicators into a group of fewer and nonintercorrelated synthetic indicators. PCA usually consists of the following steps [33]. First, calculate the covariance matrix of the sample data, and then calculate the eigenvalues of the covariance matrix and the corresponding orthogonal unit eigenvectors. Sort the eigenvalues, and choose the maximum eigenvalues and the corresponding eigenvectors. Convert the data to the new space constructed by these eigenvectors. PCA can effectively
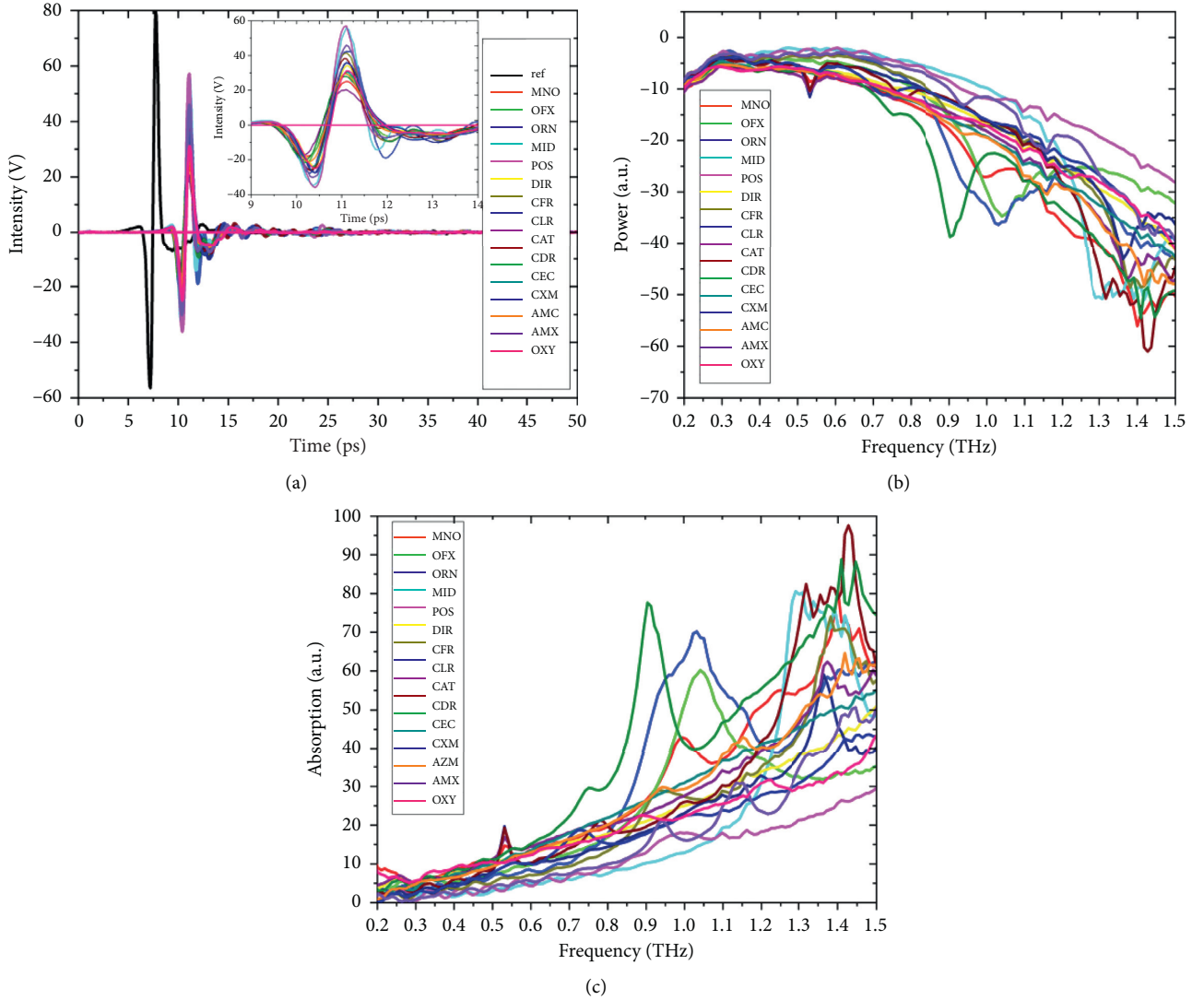
(a)

(b)

(c)

FIGURE 2: THz spectra of different types of antibiotics. (a) THz time-domain spectra. (b) THz frequency spectra. (c) THz absorbance spectra.

restore the original data and solve the problems of information overlap and multicollinearity while reducing the dimensionality of data.

t-SNE is a method that introduces a t-distribution to optimize the crowding problem suffered by the original SNE algorithm [34]. The core principle of t-SNE is to perform similarity modeling of the data points by using a normalized Gaussian kernel in the high-dimensional space and by using a t-distribution in the low-dimensional space. Following this principle, there will be a higher probability of similar points being selected and a lower probability of nonsimilar points being selected.

This algorithm consists of the following steps [35]:

First, represent the similarity between the two data points by conditional probability.

$$p_{j|i} = \frac{\exp\left(-\left\|x_i - x_j\right\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq i} \exp\left(-\left\|x_i - x_k\right\|^2 / 2\sigma_i^2\right)}. \tag{5}$$

Then, represent the joint probability distribution of the low-dimensional data by a t-distribution with a degree of freedom of 1.

$$q_{j|i} = \frac{\left(1 + \left\|y_i - y_j\right\|^2\right)^{-1}}{\sum_{k \neq i}\left(1 + \left\|y_i - y_k\right\|^2\right)^{-1}}. \tag{6}$$

Finally, obtain the optimal simulation points by gradient descent that minimizes the KL divergence of all points. Thus, samples in the low-dimensional subspace are obtained. Make sure that the probability distribution $q_{j|i}$ of data mapped into the low-dimensional space can effectively simulate the probability distribution in the high-dimensional space $p_{j|i}$.

$$C = \sum_i KL\left(P_i \| Q_i\right) = \sum_i \sum_j p_{j|i} \log \frac{p_{j|i}}{q_{j|i}}. \tag{7}$$

For the selected frequency band, the number of dimensions of data from the absorption spectra was as high as 143. In order to reduce the training time of the model and to increase the accuracy of the identification models, dimensionality reduction was performed using PCA and t-SNE, which was followed by pattern recognition using different models. Then, different methods were compared to find the optimal dimensionality reduction method for antibiotics identification. PCA was applied to the absorption spectra of 640 samples (16 ∗ 40). Figure 3(a) shows the 3D distribution of the principal components of the absorption spectra for different antibiotics. Three principal components (PC1, PC2, and PC3) were identified, and their contribution rates were 86.62%, 10.23%, and 1.13%, respectively. The sum of the contribution rates of the three principal components was 97.98%. Therefore, these three principal components could sufficiently represent the original absorption spectra. Figure 3(b) shows the 3D distribution of the different antibiotics visualized by t-SNE. It is clear to see that the divergence of the samples in Figure 3(b) is far higher than that in Figure 3(a). The samples were well clustered together, with few overlaps between different classes.

3.3. Identification Analysis. After dimensionality reduction by either PCA or t-SNE, the new data matrix (640 samples × 3 dimensions) was used to replace the original spectral data matrix (640 samples × 143 dimensions). The 640 samples were randomly divided into a training set (16 ∗ 30 = 480 samples) and a test set (16 ∗ 10 = 160 samples). The parameters of the SVM model were trained using the training set. Then, SVM was, respectively, combined with GS, GA, and PSO to optimize the model parameters [36, 37]. Finally, the prediction accuracy of the model was evaluated using the test set. The optimal combination of dimensionality reduction method and model parameter optimization was determined by comparison. Thus, the optimal identification model for the THz spectra of the antibiotics was established.

Here, the identification model was built based on an SVM. An SVM is a supervised machine learning algorithm. In SVMs, the optimal decision hyperplane is found that maximizes the distance from the two sides of the hyperplane to the two classes of samples nearest to the hyperplane. In this way, good generalization is achieved for identification. The performance of an SVM mainly depends on the penalty factor c and kernel parameter g of the model. The model should be trained to achieve the optimal identification result, and the optimal model parameters should be chosen. To do this, parameters c and g were first optimized by grid search (GS). Then, GS was combined with different dimensionality reduction methods to establish the No-GS-SVM, PCA-GS-SVM, and t-SNE-GS-SVM models. The optimal cross-validation accuracy (CVAccuracy) of each model was determined using 5-fold cross-validation, along with the prediction accuracy of this model on the training set and test set. The results are shown in Table 2. Figure 4 shows the results of parameter optimization by GS-SVM. Figure 4(a) shows the 3D results of parameter selection by No-GS-SVM,

with the CVAccuracy being 99.5833%. Figure 4(b) shows the 3D results of parameter selection by PCA-GS-SVM, with the CVAccuracy being 99.7917%. Figure 4(c) shows the 3D results of parameter selection by t-SNE-GS-SVM, with a CVAccuracy of 100%. It is clear to see that the recognition accuracy was the highest after dimensionality reduction with t-SNE.

A genetic algorithm (GA) and particle swarm optimization (PSO) were introduced to find the optimal combination of parameters c and g to further improve the prediction accuracy. The initial population size was set to 20, and the number of iterations was 50. The CVAccuracy, training set accuracy, and prediction set accuracy of No-GA-SVM, PCA-GA-SVM, and t-SNE-GA-SVM under 5-fold cross-validation are shown in Table 2. The fitness curves of the three models are presented in Figure 5. Figure 5(a) shows the fitness curve of No-GA-SVM, with a CVAccuracy of 99.7917%; Figure 5(b) shows the fitness curve of PCA-GA-SVM, with a CVAccuracy of 100%; Figure 5(c) shows the fitness curve of t-SNE-GA-SVM, with a CVAccuracy also of 100%. As compared to the GA, PSO does not include crossover and mutation operations, and the global optimum is searched by tracking the current optimal value. For this reason, the accuracy of PSO is higher. The initial population size was set to 20, and the number of iterations was 50. The CVAccuracy, training set accuracy, and prediction set accuracy of No-PSO-SVM, PCA-PSO-SVM, and t-SNE-PSO-SVM under 5-fold cross-validation are shown in Table 2. The fitness curves of the three models are shown in Figure 6. Figure 6(a) is the fitness curve of No-PSO-SVM, with a CVAccuracy of 100%; Figure 6(b) is the fitness curve of PCA-PSO-SVM, with a CVAccuracy of 100%; Figure 6(c) is the fitness curve of tSNE-PSO-SVM, with a CVAccuracy also of 100%. The optimal recognition accuracy was reached after PSO.

As shown in Table 2, under 5-fold cross-validation, the CVAccuracy of the same SVM model combined with t-SNE was higher than that of the SVM model combined with PCA or the SVM model without dimensionality reduction. When the same dimensionality reduction method was used, PSO-SVM exhibited the best identification performance compared to GA-SVM and GS-SVM.

PCA and t-SNE were, respectively, combined with GS-SVM, GA-SVM, and PSO-SVM. The samples were randomly divided into a training set and test set. Each model was run 100 times to calculate the average training accuracy, average prediction accuracy, and average time consumption. The comparison results are shown in Table 3. For the same recognition model, both the average training accuracy and prediction accuracy were higher with dimensionality reduction than without dimensionality reduction. t-SNE was consistently superior to PCA for dimensionality reduction for the THz spectra of antibiotics and also better than no use of dimensionality reduction. Additionally, the training time of the model was significantly shortened after dimensionality reduction. The time of a single training run after dimensionality reduction with PCA was shorter than that with t-SNE. This comprehensive comparison indicated that of 9 recognition models, t-SNE-PSO-SVM had the highest
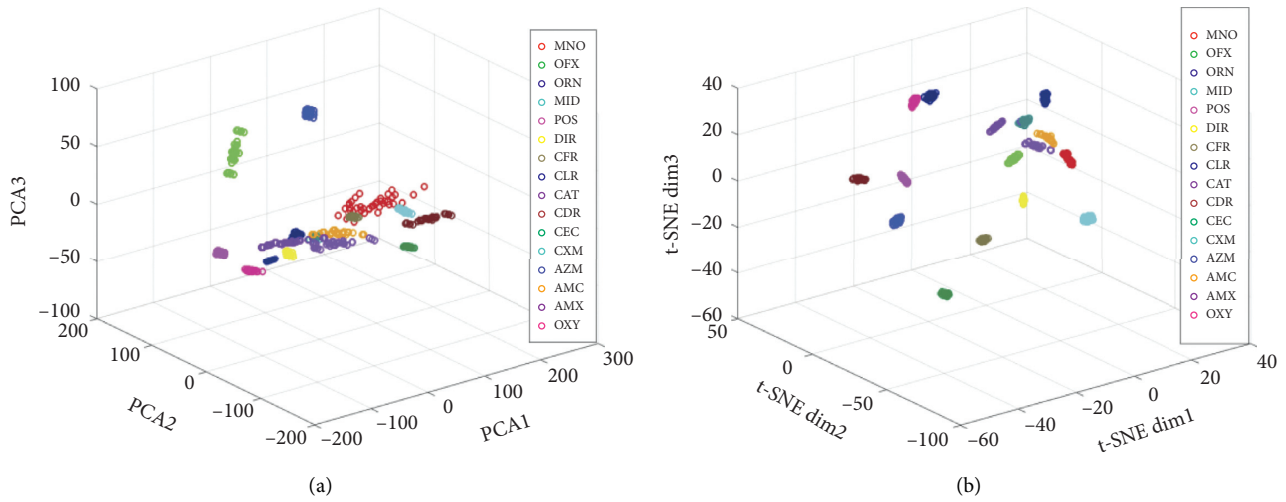
(a)

(b)

FIGURE 3: The visualization plots of different types of antibiotics by PCA and t-SNE. (a) Three-dimensional map of the first three PCs. (b) The visualization plots by t-SNE.

TABLE 2: Parameters of grid search, GA and PSO.

Chemometric method	Features selection method	Best c	Best g	Population	Iteration	Cross-validation	CVAccuracy (%)	Training set accuracy (%)	Test set accuracy (%)
GS-SVM	No	0.32988	1	—	—	5	99.5833	99.7917	98.75
	PCA	0.32988	0.32988	—	—	5	99.7917	99.7917	100
	t-SNE	0.18946	0.011842	—	—	5	100	100	98.125
GA-SVM	No	6.8678	0.58479	20	50	5	99.7917	99.7917	99.375
	PCA	13.1126	1.5594	20	50	5	100	100	99.375
	t-SNE	0.92001	0.72413	20	50	5	100	100	100
PSO-SVM	No	0.1	0.1	20	50	5	100	100	99.375
	PCA	10.2413	0.60183	20	50	5	100	100	99.375
	t-SNE	12.591	0.1	20	50	5	100	100	100

The result of SVM parameter (3D) [grid search method] Best $c = 0.32988$, $g = 1$, CV accuracy = 99.5833%

(a)

The result of SVM parameter selection (3D) [PCA + grid search method] Best $c = 0.32988$, $g = 0.32988$, CV accuracy = 99.7917%
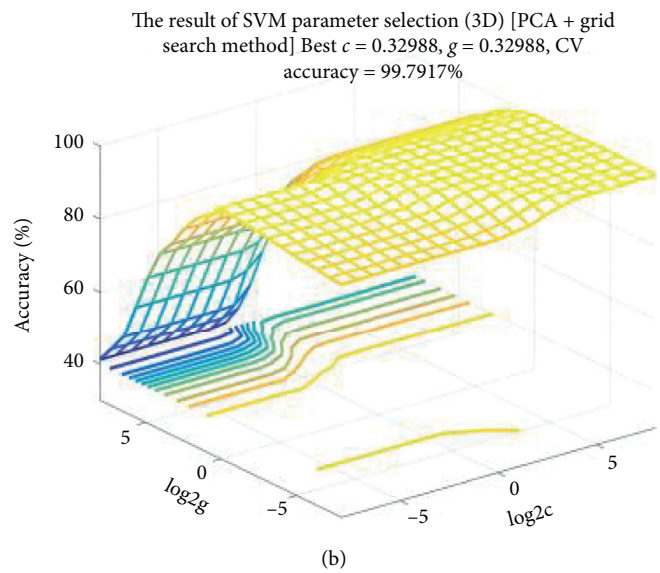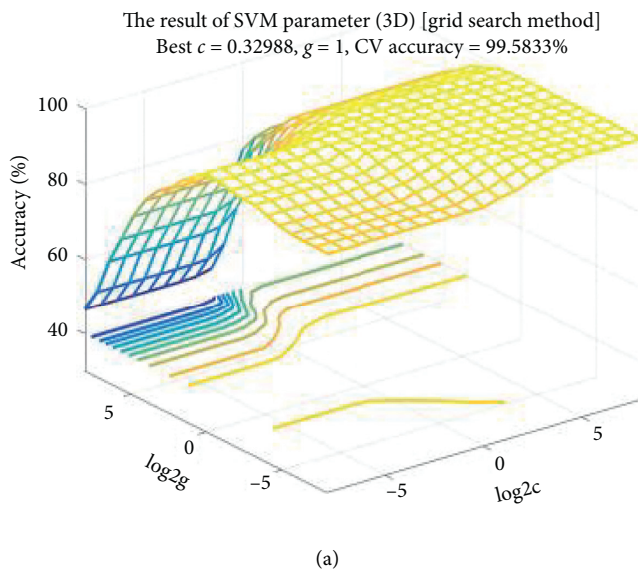

(b)

FIGURE 4: Continued.

The result of SVM parameter selection (3D) [t-SNE + grid
search method] Best $c = 0.18946$, $g = 0.011842$, CV accuracy = 100%
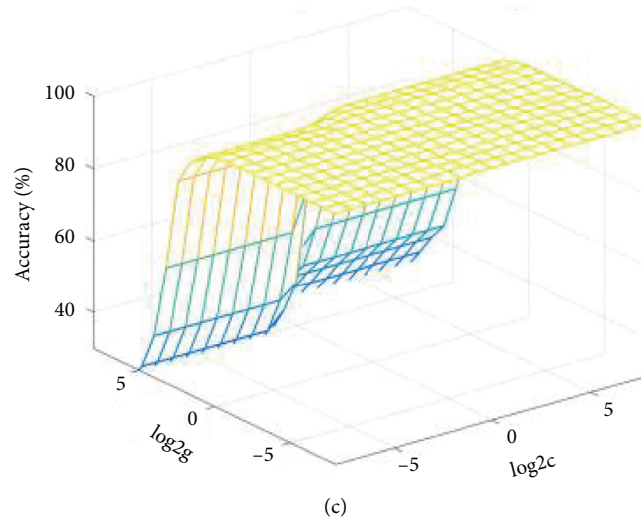

(c)

FIGURE 4: Result of GS-SVM parameter selection. (a) Result of No-GS-SVM parameter selection (best $c = 0.32988$, best $g = 1$, and CV Accuracy = 99.5833%). (b) Result of PCA-GS-SVM parameter selection (best $c = 0.32988$, best $g = 0.32988$, and CV Accuracy = 99.7917%). (c) Result of t-SNE-GS-SVM parameter selection (best $c = 0.18946$, best $g = 0.011842$, and CV Accuracy = 100%).
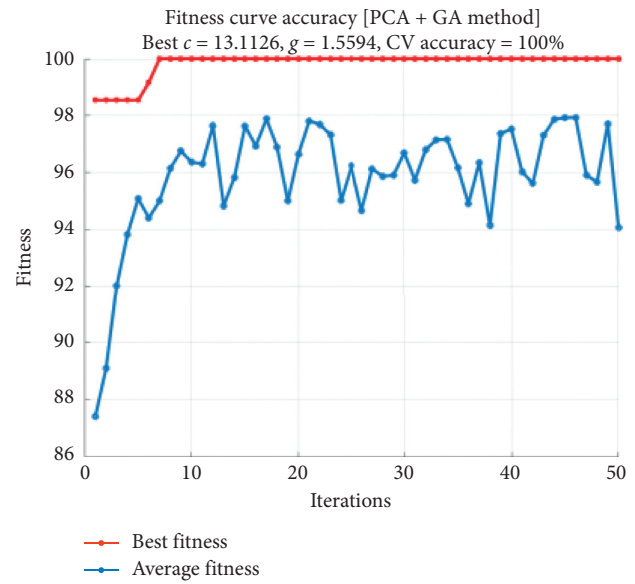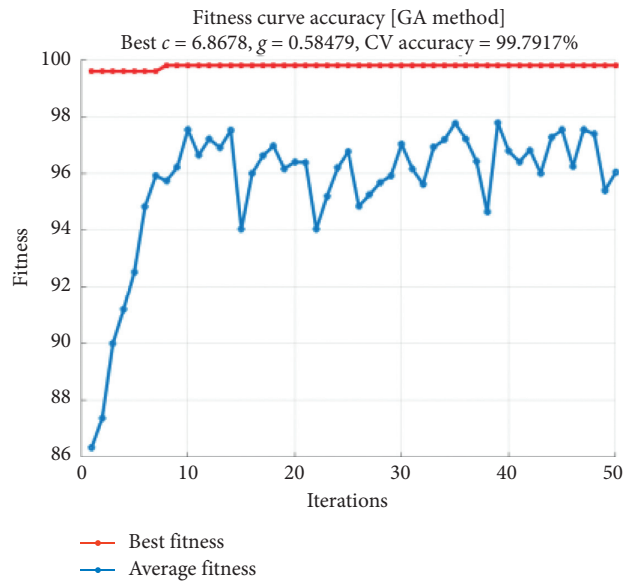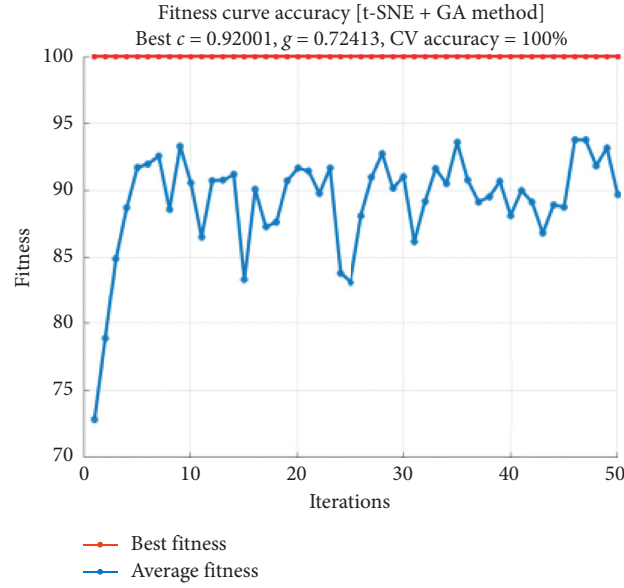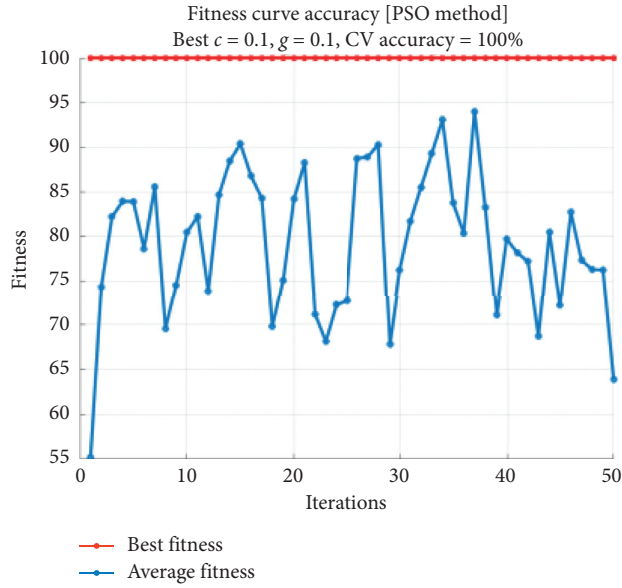

Fitness curve accuracy [GA method]
Best $c = 6.8678$, $g = 0.58479$, CV accuracy = 99.7917%

(a)

Fitness curve accuracy [PCA + GA method]
Best $c = 13.1126$, $g = 1.5594$, CV accuracy = 100%

(b)

FIGURE 5: Continued.

Fitness curve accuracy [t-SNE + GA method]
Best $c = 0.92001$, $g = 0.72413$, CV accuracy = 100%
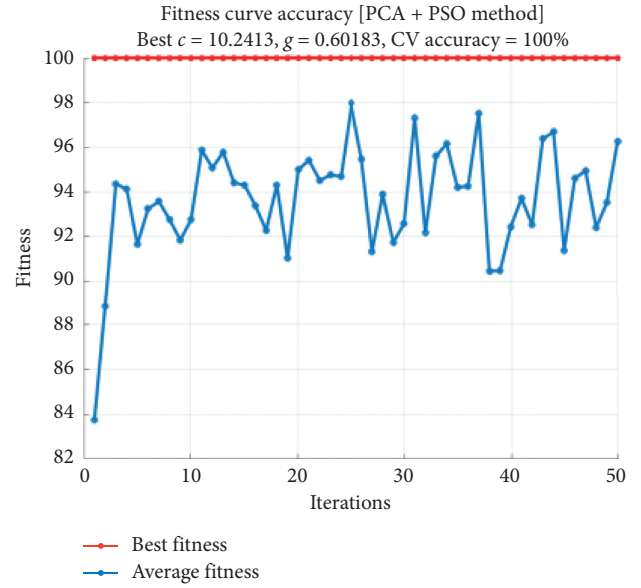
— Best fitness
— Average fitness

(c)

FIGURE 5: Fitness curve of GA. (a) Fitness curve of No-GA-SVM (best $c = 6.8678$, $g = 0.72413$, and CVAccuracy = 99.7917). (b) Fitness curve of PCA-GA-SVM (best $c = 13.1126$, $g = 1.5594$, and CVAccuracy = 100%). (c) Fitness curve of t-SNE-GA-SVM (best $c = 0.92001$, $g = 0.72413$, and CVAccuracy = 100%).

Fitness curve accuracy [PSO method]
Best $c = 0.1$, $g = 0.1$, CV accuracy = 100%

— Best fitness
— Average fitness

(a)

Fitness curve accuracy [PCA + PSO method]
Best $c = 10.2413$, $g = 0.60183$, CV accuracy = 100%

— Best fitness
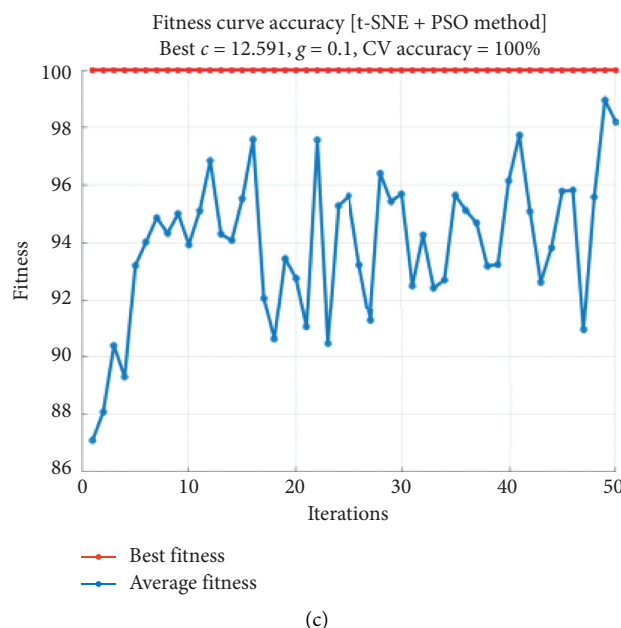— Average fitness

(b)

FIGURE 6: Continued.

(c)

Figure 6: Fitness curve of PSO. (a) Fitness curve of No-PSO-SVM (best $c = 0.1$, $g = 0.1$, and CVAccuracy = 100%). (b) Fitness curve of PCA-PSO-SVM (best $c = 10.2413$, $g = 0.60183$, and CVAccuracy = 100%). (c) Fitness curve of t-SNE-PSO-SVM (best $c = 12.591$, $g = 0.1$, and CVAccuracy = 100%).

Table 3: Identification results of all chemometric methods.

Chemometric method	Features selection method	Mean accuracy in calibration set (%)	Mean accuracy in prediction set (%)	Training time (s)
GS-SVM	No	99.15	98.19	96.61
	PCA	99.48	99.71	19.77
	t-SNE	99.85	99.79	22.76
GA-SVM	No	99.34	99.14	118.29
	PCA	99.64	99.76	28.79
	t-SNE	99.90	99.81	55.83
PSO-SVM	No	99.77	99.68	243.59
	PCA	99.81	99.67	47.04
	t-SNE	99.82	99.91	53.26

average prediction accuracy of 99.91%. Therefore, t-SNE-PSO-SVM was better for the recognition of THz spectra of antibiotics and had higher practical application value.

4. Conclusions

The present study was mainly concerned with antibiotics identification based on THz-TDS. Antibiotics come in many forms, and their direct differentiation may be impossible. We found that the THz time-domain spectra and absorption spectra only displayed minor differences between different antibiotics, which made direct differentiation difficult. Therefore, chemometric pattern recognition was introduced to build recognition models for antibiotics. PCA and t-SNE were, respectively, used for feature selection and dimensionality reduction. Then, these two methods were combined with GS-SVM, GA-SVM, and PSO-SVM to build the identification models. The optimal model was chosen after parameter

optimization and comparative analysis. The experiments showed that the training time of the identification model was significantly shortened after dimensionality reduction, and the recognition accuracy was higher with t-SNE than with PCA. The comprehensive comparison indicated that t-SNE-PSO-SVM had the highest average prediction accuracy among all models, which was 99.91%. Therefore, t-SNE-PSO-SVM was more suitable for antibiotics identification. Our study also confirmed that the combination of THz-TDS and chemometric pattern recognition has great potential for drug detection.

Data Availability

The data used to support the findings of this study have not been made available because the experimental data involved in the paper are all obtained based on our own designed experiments and need to be kept confidential, we are still using it for further research.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of the paper.

Acknowledgments

References

[1] L. Lan, Y. Yao, J. Ping, and Y. Ying, "Recent advances in nanomaterial-based biosensors for antibiotics detection," *Biosensors and Bioelectronics*, vol. 91, pp. 504–514, 2017.

[2] A. Mehlhorn, P. Rahimi, and Y. Joseph, "Aptamer-based biosensors for antibiotic detection: a review," *Biosensors*, vol. 8, no. 2, 2018.

[3] C. Kim, H.-D. Ryu, E. G. Chung, Y. Kim, and J.-k. Lee, "A review of analytical procedures for the simultaneous determination of medically important veterinary antibiotics in environmental water: sample preparation, liquid chromatography, and mass spectrometry," *Journal of Environmental Management*, vol. 217, pp. 629–645, 2018.

[4] S. Yang, X. Zhu, J. Wang et al., "Combustion of hazardous biological waste derived from the fermentation of antibiotics using TG-FTIR and Py-GC/MS techniques," *Bioresource Technology*, vol. 193, pp. 156–163, 2015.

[5] C. Long, B. Deng, S. Sun, and S. Meng, "Simultaneous determination of chlortetracycline, ampicillin and sarafloxacin in milk using capillary electrophoresis with electrochemiluminescence detection," *Food Additives & Contaminants: Part A*, vol. 34, no. 1, pp. 24–31, 2016.

[6] Z. Wang, R. C. Beier, and J. Shen, "Immunoassays for the detection of macrocyclic lactones in food matrices-a review," *TrAC Trends in Analytical Chemistry*, vol. 92, pp. 42–61, 2017.

[7] R. Mahmoudi, R. Norian, and M. Pajohi-Alamoti, "Antibiotic residues in Iranian honey by elisa," *International Journal of Food Properties*, vol. 17, no. 10, pp. 2367–2373, 2014.

[8] Q. Luo, N. Yu, C. Shi, X. Wang, and J. Wu, "Surface plasmon resonance sensor for antibiotics detection based on photo-initiated polymerization molecularly imprinted array," *Talanta*, vol. 161, pp. 797–803, 2016.

[9] Y. Du, Q. Cai, J. Xue, and Q. Zhang, "Raman and terahertz spectroscopic investigation of cocrystal formation involving antibiotic nitrofurantoin drug and coformer 4-aminobenzoic acid," *Crystals*, vol. 6, no. 12, 2016.

[10] Q. Wang and W.-M. Zhao, "Optical methods of antibiotic residues detections: a comprehensive review," *Sensors and Actuators B: Chemical*, vol. 269, pp. 238–256, 2018.

[11] S. Zhong, "Progress in terahertz nondestructive testing: a review," *Frontiers of Mechanical Engineering*, vol. 14, no. 3, pp. 273–281, 2019.

[12] L. H. Eadie, C. B. Reid, A. J. Fitzgerald, and V. P. Wallace, "Optimizing multi-dimensional terahertz imaging analysis for colon cancer diagnosis," *Expert Systems with Applications*, vol. 40, no. 6, pp. 2043–2050, 2013.

[13] K. Lee, K. Jeoung, S. H. Kim et al., "Measuring water contents in animal organ tissues using terahertz spectroscopic imaging," *Biomedical Optics Express*, vol. 9, no. 4, pp. 1582–1589, 2018.

[14] L. Afsah-Hejri, P. Hajeb, P. Ara, and R. J. Ehsani, "A comprehensive review on food applications of terahertz spectroscopy and imaging," *Comprehensive Reviews in Food Science and Food Safety*, vol. 18, no. 5, pp. 1563–1621, 2019.

[15] B. Cao, H. Li, M. Fan, W. Wang, and M. Wang, "Determination of pesticides in a flour substrate by chemometric methods using terahertz spectroscopy," *Analytical Methods*, vol. 10, no. 42, pp. 5097–5104, 2018.

[16] W. Shi, C. Dong, L. Hou, Z. Xing, Q. Sun, and L. Zhang, "Investigation of aging characteristics in explosive using terahertz time-domain spectroscopy," *International Journal of Modern Physics B*, vol. 33, no. 24, 2019.

[17] W. Limwikrant, K. Higashi, K. Yamamoto, and K. Moribe, "Characterization of ofloxacin-oxalic acid complex by PXRD, NMR, and THz spectroscopy," *International Journal of Pharmaceutics*, vol. 382, no. 1-2, pp. 50–55, 2009.

[18] H.-l. Zhang, Y. Xia, Z. Hong, and Y. Du, "Spectral analysis about the pharmaceutical cocrystal formation of piracetam and 3-hydroxybenzoic acid," *Spectroscopy and Spectral Analysis*, vol. 35, no. 7, pp. 1854–1859, 2015.

[19] Z.-w. Zhang, J. Zuo, and C.-l. Zhang, "Far-IR and THz absorption spectra studies of metronidazole, tinidazole and ornidazole," *Spectroscopy and Spectral Analysis*, vol. 32, no. 4, pp. 906–909, 2012.

[20] L. Xie, C. Wang, M. Chen et al., "Temperature-dependent terahertz vibrational spectra of tetracycline and its degradation products," *Spectrochimica Acta Part A-Molecular and Biomolecular Spectroscopy*, vol. 222, 2019.

[21] J. Qin, L. Xie, and Y. Ying, "Feasibility of terahertz time-domain spectroscopy to detect tetracyclines hydrochloride in infant milk powder," *Analytical Chemistry*, vol. 86, no. 23, pp. 11750–11757, 2014.

[22] J. Qin, L. Xie, and Y. Ying, "Determination of tetracycline hydrochloride by terahertz spectroscopy with PLSR model," *Food Chemistry*, vol. 170, pp. 415–422, 2015.

[23] J. Qin, L. Xie, and Y. Ying, "A high-sensitivity terahertz spectroscopy technology for tetracycline hydrochloride detection using metamaterials," *Food Chemistry*, vol. 211, pp. 300–305, 2016.

[24] J. Qin, L. Xie, and Y. Ying, "Rapid analysis of tetracycline hydrochloride solution by attenuated total reflection terahertz time-domain spectroscopy," *Food Chemistry*, vol. 224, pp. 262–269, 2017.

[25] M. Massaouti, C. Daskalaki, A. Gorodetsky, A. D. Koulouklidis, and S. Tzortzakis, "Detection of harmful residues in honey using terahertz time-domain spectroscopy," *Applied Spectroscopy*, vol. 67, no. 11, pp. 1264–1269, 2013.

[26] Y. Wang, Q. Wang, Z. Zhao, A. Liu, Y. Tian, and J. Qin, "Rapid qualitative and quantitative analysis of chlortetracycline hydrochloride and tetracycline hydrochloride in environmental samples based on terahertz frequency-domain spectroscopy," *Talanta*, vol. 190, pp. 284–291, 2018.

[27] Y. Long, B. Li, and H. Liu, "Analysis of fluoroquinolones antibiotic residue in feed matrices using terahertz spectroscopy," *Applied Optics*, vol. 57, no. 3, pp. 544–550, 2018.

[28] L. Yang, T. Guo, X. Zhang, S. Cao, and X. Ding, "Toxic chemical compound detection by terahertz spectroscopy: a review," *Reviews in Analytical Chemistry*, vol. 37, no. 3, 2018.

[29] J. I. A. Yan, C. Sijia, L. I. Ning et al., "Identification of terahertz absorption spectra of illicit drugs using back propagation neural networks," *Chinese Journal of Lasers*, vol. 34, no. 5, pp. 719–722, 2007.

[30] R. Zhang, T. Wu, and Y. Zhao, "Terahertz signal analysis for biological tissues based on empirical mode decomposition," *Optik*, vol. 183, pp. 906–911, 2019.

[31] T. D. Dorney, R. G. Baraniuk, and D. M. Mittleman, "Material parameter estimation with terahertz time-domain spectroscopy," *Journal of the Optical Society of America A*, vol. 18, no. 7, pp. 1562–1571, 2001.

[32] L. Duvillaret, F. Garet, and J.-L. Coutaz, "A reliable method for extraction of material parameters in terahertz time-domain spectroscopy," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 2, no. 3, pp. 739–746, 1996.

[33] H. Yu, R. Chen, and G. Zhang, "A SVM stock selection model within PCA," in Proceedings of the 2nd International Conference on Information Technology and Quantitative Management, Itqm 2014, Procedia Computer Science, Y. Shi, A. Lepskiy, and F. Aleskerov, Eds., pp. 406–412, Moscow, Russia, June 2014.

[34] L. van der Maaten and G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.

[35] L. van der Maaten, "Accelerating t-SNE using tree-based algorithms," *Journal of Machine Learning Research*, vol. 15, pp. 3221–3245, 2014.

[36] X. Wang, S. Guan, L. Hua, B. Wang, and X. He, "Classification of spot-welded joint strength using ultrasonic signal time-frequency features and PSO-SVM method," *Ultrasonics*, vol. 91, pp. 161–169, 2019.

[37] T. Zhou, H. Lu, W. Wang, and X. Yong, "GA-SVM based feature selection and parameter optimization in hospitalization expense modeling," *Applied Soft Computing*, vol. 75, pp. 323–332, 2019.