Research Article

A Novel THz Differential Spectral Clustering Recognition Method Based on t-SNE

Tie-Jun Li,¹ Chih-Cheng Chen ^{1,2} Jian-jun Liu,³ Gui-fang Shao ⁴,
and Christopher Chun Ki Chan⁵

¹College of Information Engineering, Jimei University, Fujian Province, Xiamen 361021, China

²Department of Aeronautical Engineering, Chaoyang University of Technology, Taichung, Taiwan

³Shaoguan University, Guangzhou Province, Shaoguan 512005, China

⁴Department of Automation, Xiamen University, Fujian Province, Xiamen 361005, China

⁵Department of Information Management, Chaoyang University of Technology, Taichung, Taiwan

Correspondence should be addressed to Chih-Cheng Chen; 3343033397@qq.com and Gui-fang Shao; gfshao@xmu.edu.cn

Received 6 March 2020; Accepted 6 July 2020; Published 1 September 2020

Guest Editor: Fuqiang Gu

Copyright © 2020 Tie-Jun Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We apply time-domain spectroscopy (THz) imaging technology to perform nondestructive detection on three industrial ceramic matrix composite (CMC) samples and one silicon slice with defects. In terms of spectrum recognition, a low-resolution THz spectrum image results in an ineffective recognition on sample defect features. Therefore, in this article, we propose a spectrum clustering recognition model based on t-distribution stochastic neighborhood embedding (t-SNE) to address this ineffective sample defect recognition. Firstly, we propose a model to recognize a reduced dimensional clustering of different spectrums drawn from the imaging spectrum data sets, in order to judge whether a sample includes a feature indicating a defect or not in a low-dimensional space. Second, we improve computation efficiency by mapping spectrum data samples from high-dimensional space to low-dimensional space by the use of a manifold learning algorithm (t-SNE). Finally, to achieve a visible observation of sample features in low-dimensional space, we use a conditional probability distribution to measure the distance invariant similarity. Comparative experiments indicate that our model can judge the existence of sample defect features or not through spectrum clustering, as a predetection process for image analysis.

1. Introduction

Nondestructive testing is one of the most significant applications of terahertz technology, and terahertz time-domain spectroscopy (THz—TDS) system is a commonly used technique [1, 2]. Information from each pixel of spectrum analysis can be obtained by terahertz determinant scanning and analysis of a large number of high-dimensional spectral image data to detect superficial damage and internal defects such as bubbles, cracks, and impurities of samples. However, when applying terahertz nondestructive testing, a terahertz wave is usually subject to a diffraction limit and a spatial optical resolution limit of the system, which makes it difficult to identify target microdefect structures with an optical resolution lower than terahertz. To solve this kind of

problem, improving the precision of optical hardware and image processing [3, 4] is a common method but with limited effectiveness. Since the spectrum of the defect feature of a detected target sample is a “differential spectrum,” (different than the normal structure), the cluster of the abnormal spectrum can be identified in a low-dimensional space through unsupervised clustering using the terahertz spectral data set from the perspective of spectral clustering recognition.

To observe the distribution of sample points from high-dimensional spectral data, it is necessary to reduce the dimensionality of it [5]. We can then perform terahertz spectral data redundancy and noise removal through the extraction of the main spectral features of each local point. The sample points cluster in two or three dimensions. At

present, the manifold learning method is usually used in dimension reduction for hyperspectral data clustering recognition, along with principal components analysis (PCA), multidimensional scaling analysis (MDS) [6], isometric feature mapping (ISOMAP) [7], locally linear embedding (LLE) [8], and spectrum embedding (SE) [9]. However, these data dimensionality reduction methods have problems such as unclear classification interface, poor visual classification effect, and slow convergence speed.

Maaten proposed a t-distribution SNE (t-SNE) algorithm [10, 11] based on the stochastic neighbor embedding (SNE) algorithm [12], which is a nonlinear dimension reduction method based on manifold learning. According to the principle that data points with similar distances in high-dimensional space are mapped to low-dimensional space with similar distances, a method of subspace analysis is adopted to measure the similarity of such spatial distances with conditional probability distribution. t-SNE changes the idea of similarity based on Gaussian distance or Euclidean distance measure in MDS and ISOMAP algorithms. It maps high-dimensional space sample points to low-dimensional space; meanwhile, the distribution probability between them remains unchanged as far as possible. Because the t-SNE algorithm has excellent visual classification effect, clear classification interface, and high algorithm efficiency, the method has been widely used in biomedical data analysis, fault diagnosis, spectrum analysis, artificial intelligence, and many other fields. For example, applied t-SNE dimensional reduction techniques are used to classify disease cells [13, 14], human genetic patterns [15], and RNA sequences [16]. By utilizing the t-SNE technique, other literatures [16–18] also applied it to classify multifaults in mechanical systems. In order to realize the classification and visualized diagnosis on the spectral information, the authors in [17–19] also made relevant research progress in their respective fields. In recent years, with the rapid development of artificial intelligence, the t-SNE technique was also applied to different AI-related applications [20–22].

This paper proposes a spectral identification model based on the t-SNE algorithm for the “differential spectrum” of sample defect features as a nondestructive method of testing. Based on the unsupervised clustering of “differential spectrum” from terahertz spectral data, we can achieve super-resolution identification of sample defect features at the pixel level, thus performing predetection analysis for further terahertz spectral imaging.

2. Terahertz Spectral Recognition Model Based on t-Distributed Stochastic Neighbor Embedding

The establishment of t-SNE terahertz spectral recognition model consists of the following four steps:

- (1) Define the data set, calculate the confusion cost function, and initialize the optimization parameters of the model
- (2) Set the low-dimensional data representation of optimized results

- (3) Obtain target results from stochastic gradient descent optimization training
- (4) Iterate through the pipeline until the number of iterations is reached

The algorithm flowchart is shown in Figure 1.

2.1. Basic Parameters Definition. To build the t-SNE model, we first define some basic parameters.

Let X be the spectral data set in a higher dimensional space, X_i represents the sample point and $X = \{x_1, x_2, \dots, x\}$, and the dimension of the sample is D . Low-dimensional spatial data sets are represented by $Y = \{y_1, y_2, \dots, y_n\}$, and the dimension of the sample is d with value of 2 or 3 to visualize the cluster analysis. The conditional profile distribution matrix P_i of a high-dimensional data set is defined as follows:

$$P_{j|i} = \frac{\exp\left(-\|x_i - x_j\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq i} \exp\left(-\|x_i - x_k\|^2 / 2\sigma_i^2\right)}, \quad (1)$$

in which $P_{j|i}$ represents the probability that the i th sample is distributed around sample j , $P_{j|i} = 0$; σ denotes the variance of the Gaussian distribution centered on x_i and is determined according to the principle of maximum entropy. The entropy $H(P_i)$ in which P_i increases with the increase in σ_i , is defined as

$$H(P_i) = - \sum_j P_{j|i} \log_2 P_{j|i}. \quad (2)$$

To evaluate the number of effective nearest neighbors around a point, we introduce the concept of perplexity, which is a global parameter and defined as follows:

$$\text{Perp}(P_i) = 2^{H(P_i)}. \quad (3)$$

In order to make the adjustment of perplexity more robust, the perplexity is usually chosen between 5 and 50, and the binary search method is used to find the best σ .

The conditional probability distribution matrix Q_i in low-dimensional space is defined as follows:

$$q_{j|i} = \frac{\exp(-y_i - y_j^2)}{\sum_{k \neq i} \exp(-y_i - y_k^2)}. \quad (4)$$

2.2. Symmetric t-SNE. We let the probability distribution matrix in high-dimensional and low-dimensional space be symmetric and construct the joint probability distribution P and Q so that for any i and j , $p_{ij} = p_{ji}$ and $q_{ij} = q_{ji}$.

We redefine q_{ij} in low-dimensional space by t-distribution:

$$q_{ij} = \frac{(1 + y_i - y_j^2)^{-1}}{\sum_{k \neq i} (1 + y_k - y_k^2)^{-1}}. \quad (5)$$

Then, define q_{ij} in higher dimensions:

FIGURE 1: The flowchart of spectrum recognition based on the t-SNE algorithm.

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n}, \quad (6)$$

in which n is the total number of sample points in the data set.

2.3. Cost Function and Training. We use Kullback–Leibler divergence (KLD) to measure the similarity of two spatial distributions in high and low dimensions, and the SEN algorithm aims to minimize the KL distance for all data points in the sample set.

We then use gradient descent to minimize cost function:

$$C = KL(PQ) = \sum p_{ij} \log_2 \frac{p_{ij}}{q_{ij}}. \quad (7)$$

The gradient descent is also used for training the model and its formula is as follows:

$$\frac{\delta C}{\delta y_i} = 4 \sum_j (p_{ij} - q_{ij})(y_i - y_j) \left(1 + \|y_i - y_j\|^2\right)^{-1}. \quad (8)$$

In addition, to accelerate the optimization process and avoid falling just obtaining a local optima, a relatively large momentum should be used in the gradient; that is, in addition to the current gradient, the exponential decay term accumulated by the previous gradient should also be introduced in the parameter update. The formula is as follows:

$$Y^{(t)} = Y^{(t-1)} + \eta \frac{\delta C}{\delta y} + \alpha(t)(Y^{(t-1)} - Y^{(t-2)}), \quad (9)$$

where $Y^{(t)}$ is the solution for iteration t , η represents the learning rate, and $\alpha(t)$ denotes the momentum for iteration t . The random normal distribution of the initial value $Y^{(0)}$ is usually set to $N(0, 10^{-4}I)$.

2.4. Implementation Steps. When the t-SNE algorithm is adopted to reduce high-dimensional data, if the dimension of data points is too large, then the algorithm will take a long time. In order to improve the efficiency of t-SNE, the PCA method is usually introduced first to reduce the dimension of a high-dimensional sample point data set to 50 dimensions, and then, t-SNE is used for cluster recognition. The specific pseudocode is shown in Algorithm 1.

3. Experiment and Analysis

3.1. Types of Samples. In this paper, two kinds of ceramic matrix composites (CMCs) and a silicon slice were selected as the detection objects for terahertz spectral imaging. The samples are 1 piece of alumina (Al_2O_3) ceramic sheet, 2 pieces of beryllium oxide (BeO) ceramic sheets, and 1 piece of monocrystalline silicon. For convenient comparison and analysis, the sample sheet is prepared for defect treatment. The specific specifications and defects are shown in Table 1.

3.2. Terahertz Spectrum of Samples. Nondestructive testing (NDT) method of spectral imaging was used to image the samples. The transmission time-domain spectrum of the sample is shown in Figure 2.

It can be seen from Figure 2(a) that the spectrum of alumina crack defect is significantly different from the normal spectrum, and the electric field intensity value is smaller than the peak intensity of the normal spectrum; also, the time delay is smaller than the normal spectrum. In Figure 2(b), the beryllium oxide defect part scatters a large amount of terahertz wave, resulting in severe attenuation of terahertz waves. The peak of the field strength of the differential spectrum is smaller than that of the normal spectrum, and the time delay is slightly smaller than that of the normal spectrum. In Figure 2(c), the spectrum of the monocrystalline silicon and the spectrum of the background reference signal have obvious differences in field strength and time delay.

3.3. Model Discriminant Analysis. The time-domain spectral data for each scan pixel of the samples are obtained directly by two-dimensional spectral scanning. To simplify data analysis, the original terahertz time-domain spectrum was directly used in this paper to establish the spectral data set of samples, and t-SNE was used to obtain the sample spectral data set. The scanning background spectrum was also included in the differential spectrum for cluster analysis.

Due to the high number of sample points and the spectral dimension, a random sampling method was adopted for spectral data set to reduce the time and complexity of model calculation. A certain number of sample points were randomly selected for model discrimination, and spectral clustering effect of the model was investigated

```

Algorithm: t-SNE
Input data: the sample terahertz spectral data set  $X = \{x_1, x_2, \dots, x_n\}$ 
Cost function:  $C = KL(P\|Q) = \sum_i \sum_j p_{ij} \log_2 p_{ij}/q_{ij}$ 
Output the result: low-dimensional spatial data representation  $Y(t) = \{y_1, y_2, \dots, y_n\}$ 
Optimize training process
begin
  Set iteration times  $T$ , learning rate  $\eta$ , and momentum  $\alpha(t)$ 
  Calculate the perplexity  $Perp$  and the conditional probability according to  $p_{ij} = p_{ji} + p_{ij}/2n$ 
  Randomly initialize  $Y(0) = \{y_1, y_2, \dots, y_n\}$  with a normal  $N(0, 10^{-4}, I)$  distribution
  For  $t = 1$  to  $T$ , do
    Calculate  $q_{ij}$  in lower dimensions with formula (5)
    Compute gradient  $\delta C/\delta y$  according to formula (8)
    Update  $Y^{(t)} = Y^{(t-1)} + \eta \delta C/\delta y + \alpha(t)(Y^{(t-1)} - Y^{(t-2)})$ 
  end
end
return  $Y$ 

```

ALGORITHM 1: The pseudocode of t-SNE algorithm.

TABLE 1: The sample information of CMC and silicon.

The sample name	Specification	Description
Aluminum oxide (Al_2O_3) sheet with defects	Length $\times$ width $\times$ thickness (mm): 101.0 $\times$ 101.0 $\times$ 3.0 mm Bulk density ≥ 3.65 g/cm ³ Surface roughness $Ra \geq 0.5$ μ m Principal component content $\geq 95\%$	Gray-white, defects are internal irregularities cracks
Beryllium oxide (BeO) sheet with defects	Length $\times$ width $\times$ thickness (mm): 50.8 $\times$ 50.8 $\times$ 1.5 Bulk density ≥ 2.85 g/cm ³ Surface roughness $Ra \geq 0.1$ μ m Principal component content $\geq 99\%$	Pure white, highly toxic powder, the defect is the shape of the regular set of holes 4 Outside diameter: 4.2–5.9 mm Inner diameter: 2.6–4.2 mm Hole depth: 0.8 mm
Beryllium oxide (BeO) sheet with zero defects	Length $\times$ width $\times$ thickness (mm): 50.8 $\times$ 50.8 $\times$ 1.5 Bulk density ≥ 2.85 g/cm ³ Surface roughness $Ra \geq 0.1$ μ m Principal component content $\geq 99\%$	Pure white, highly toxic powder, zero defects
Monocrystalline silicon with zero defects	Thickness: 450 $\pm$ 10 μ m Diameter: 76.3 $\pm$ 0.2 mm N dopant (phosphorus) Crystal orientation: <111>	Crystal growth: CZ resistivity: 14–45 Ω cm zero defects

under different perplexities and iteration times. The specific model parameter settings are shown in Table 2, and the model clustering results are shown in Figures 3–6, respectively.

3.3.1. Discriminant Result of BeO Sample with Defects.

Theoretically, the ceramic sheet should have a good normal spectrum clustering effect by t-SNE, and the difference spectrum includes four sets of holes defect at the spectrum and the background spectrum scanning. It can be seen from Figures 3(c) and 3(d) that, under the condition of the same number of sampling points and iterations, the confusion levels of 30 and 50 have a similar clustering effect. As the number of sampling points increases to 10000, as shown in Figures 3(e) and 3(f), the degree of clustering of samples

improves significantly, while the number of discrete clusters tends to decrease, but the demarcation of each cluster does not improve significantly. When the number of iterations increases to 5000, as shown in Figure 3(g), the clustering boundary of samples is significantly improved, and the increased spatial distance between clusters can reflect the actual classification of sample points, indicating that increasing the number of iterations can improve the stability of the model under the condition of large samples.

Under the large number of iterations, if the degree of confusion is increased, as shown in Figures 3(h) and 3(i), the sample clustering degree will also be strengthened, and more sample points with similar spectra are clustered together. Although the space distance between clusters has decreased, there is still clear dividing line between the surface. Especially in Figure 3(i), the actual situation of the sample points

FIGURE 2: THz time-domain spectrum of sample imaging scanning: (a) transmission spectrum of Al_2O_3 ; (b) transmission spectrum of BeO ; (c) spectrum of monocrystalline silicon.

classification is better reflected, so this picture can be chosen as t-SNE representative clustering view of the model. In general, the data set of abnormal spectrum is significantly smaller than that of normal structure samples and is in free state, and the results are in line with the predicted analysis.

3.3.2. Discriminant Result of the BeO Sample with Zero Defects. It can be seen from Figure 4 that, through observation and analysis on spectral images of BeO sample with zero

defects, as shown in Figure 4(b), terahertz spectrum can be divided into two categories: background signal spectrum (air part) and normal BeO spectrum, and the number of reference signal spectra is smaller than that of BeO sample points. According to the cluster recognition result of difference spectrum, the experimental results are consistent with each other under different iteration times and sampling times when the perplexity is set to 100. The spectral data set was clearly clustered into two categories, and the classification boundaries were clear. In Figure 4(f), the classification boundaries of 5000

TABLE 2: Model parameters for four samples.

Sample name	Spectral dimension	PCA dimension reduction	t-SNE dimension	Maximum number of iterations	Perplexity	Number of spectra in data set	Number of sample points
BeO sheet with defects	800	50	2	1000, 5000	30–100	45240	2000–10000
BeO sheet with zero defects	800	50	2	1000, 5000	100	45240	3000–10000
Al ₂ O ₃ sheet with defects	800	50	2	1000, 5000	100	8400	8400
Monocrystalline silicon	800	50	2	1000, 5000	100	108800	5000–12000

FIGURE 3: Clustering results of the BeO sample with defects: (a) optical image; (b) THz image; (c) 5000/1000/30; (d) 5000/1000/100; (e) 10000/1000/30; (f) 10000/1000/100; (g) 5000/5000/30; (h) 5000/5000/100; (i) 10000/5000/100. Description format: sampled data/iterations/perplexity.

FIGURE 4: Clustering results of the BeO sample with zero defects: (a) optical image; (b) THz image; (c) 3000/1000/100; (d) 5000/1000/100; (e) 10000/1000/100; (f) 5000/5000/100. Description format: sampled data/iterations/perplexity.

FIGURE 5: Clustering results of the Al_2O_3 sample with defects: (a) optical image; (b) THz image; (c) 8400/1000/100; (d) 8400/2000/100; (e) 8400/3000/100; (f) 8400/5000/100. Description format: sampled data/iterations/perplexity.

FIGURE 6: Clustering results of monocrystalline silicon with zero defects: (a) optical image; (b) THz image; (c) 8000/1000/100; (d) 12000/1000/100; (e) 10000/5000/100; (f) 12000/5000/100. Description format: sampled data/iterations/perplexity.

iterations were not as obvious as those of other iterations. In general, the reference spectral data set was significantly smaller than that of the beryllium oxide samples, which was in line with the predicted analysis results.

3.3.3. Discriminant Result of the Al_2O_3 Sample with Defects.

As shown in Figure 5, cluster analysis is performed on the spectrum data set of the Al_2O_3 sample when the perplexity is set to 100. Under different iteration times, the experimental results are consistent, and the spectral data set is obviously clustered into four categories (the reference signal spectral set, the normal sample point spectral set, the background signal spectral set, and the sample point spectral set at the crack). Especially in Figure 5(d), the classification boundary of the spectral cluster is clear, and the scale of the differential spectral data set at the crack is much smaller than that of the normal spectral data set. The clustering results are consistent with the image features observed in the spectral image (Figure 5(b)). The results show that the t-SNE can be used for differential spectral clustering analysis to realize super-resolution identification of sample defect features.

3.3.4. Discriminant Result of the Monocrystalline Silicon Sample. The clustering results of monocrystalline silicon chip are shown in Figure 6. As can be seen from the figure,

the perplexity is set to 100, and consistent clustering results are obtained under different iteration times and sampling times. Spectral data sets are clearly clustered into two categories, and spectral clustering boundaries are quite clear. In Figures 6(c) and 6(f), small-scale discrete clusters can be clearly seen, but we are unable to observe obvious defects from the spectral image (Figure 6(b)), indicating that the sample may have a tiny defect structure lower than the optical resolution, which requires further imaging observation and analysis. This also reflects the effectiveness of t-SNE in superresolution identification of small defect structures.

4. Conclusion

For terahertz nondestructive testing, it is difficult to identify the structural features of samples with tiny defects due to the resolution limitation of the optical system. To solve this problem, we applied t-SNE to perform cluster analysis and identification of spectral data sets. Experiments on sample data set of scanned images indicate that t-SNE can precluster and identify the difference spectrum of the measured object. Furthermore, provide a priori predetection basis for the next step of pattern recognition classification and imaging analysis, thus improving the accuracy of spectral target recognition. In particular, the model clustering of the imaging spectral data set can overcome the inherent limitation

of optical resolution. From the perspective of spectral clustering, this method can provide a feasible method for realizing superresolution identification of samples which has an important research value for the rapid detection of large component samples in engineering applications.

The research method in this paper is different from the traditional method to identify target defects through images, and a new superresolution method to identify target defects through spectral clustering is created, which is an important auxiliary means to identify target defects through terahertz images. Since the method in this paper can only predict whether there is a defect in the detection target and the characteristics such as the type, shape, location, and size of the defect cannot be analyzed, the research focus in the next stage of this paper will aim to how to realize the judgment of defect characteristics through spectral recognition.

Data Availability

The data can be shared and used.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Support for Scientific Research Projects (Scientific Research Projects in Colleges and Universities) (nos. 440-99000332 and ZP2020038) and Foundation Funded Project of Doctoral (no. 440-99000617).

References

- [1] D. H. Auston, K. P. Cheung, and P. R. Smith, "Picosecond photoconducting hertzian dipoles," *Applied Physics Letters*, vol. 45, no. 3, pp. 284–286, 1984.
- [2] D. H. Auston, K. P. Cheung, J. A. Valdmanis, and D. A. Kleinman, "Cherenkov radiation from femtosecond optical pulses in electro-optic media," *Physical Review Letters*, vol. 53, no. 16, pp. 1555–1558, 1984.
- [3] P. Lopato, G. Psuj, and B. Szymanik, "Nondestructive inspection of thin basalt fiber reinforced composites using combined terahertz imaging and infrared thermography," *Advances in Materials Science and Engineering*, vol. 2016, Article ID 1249625, 13 pages, 2016.
- [4] L.-M. Xu, W.-H. Fan, and J. Liu, "High-resolution reconstruction for terahertz imaging," *Applied Optics*, vol. 53, no. 33, pp. 7891–7897, 2014.
- [5] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 1844–1868, 2014.
- [6] J. B. Tenenbaum, V. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [7] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [8] G. E. Hinton, S. Geoffrey, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [9] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [10] V. D. L. Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.
- [11] V. D. L. Maaten, "Accelerating t-SNE using tree-based algorithms," *Journal of Machine Learning Research*, vol. 15, pp. 3221–3245, 2014.
- [12] G. E. Hinton and S. T. Roweis, "Stochastic neighbor embedding," *Advances in Neural Information Processing Systems*, vol. 15, no. 4, pp. 833–840, 2003.
- [13] D. T. Montoro, F. Yuan, M. Biton et al., "A revised airway epithelial hierarchy includes CFTR-expressing ionocytes," *Nature*, vol. 560, no. 7718, pp. 319–324, 2018.
- [14] D. A. Cusanovich, J. P. Reddington, D. A. Garfield et al., "The cis-regulatory dynamics of embryonic development at single-cell resolution," *Nature*, vol. 555, no. 7697, pp. 538–542, 2018.
- [15] W. Li, J. E. Cerise, Y. Yang, and H. Han, "Application of t-SNE to human genetic data," *Journal of Bioinformatics and Computational Biology*, vol. 15, no. 4, Article ID 1750017, 2017.
- [16] G. C. Linderman, M. Rachh, J. G. Hoskins, S. Steinerberger, and Y. Kluger, "Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data," *Nature Methods*, vol. 16, no. 3, pp. 243–245, 2019.
- [17] W. Jiang, J. Zhou, H. Liu, and Y. Shan, "A multi-step progressive fault diagnosis method for rolling element bearing based on energy entropy theory and hybrid ensemble auto-encoder," *ISA Transactions*, vol. 87, pp. 235–250, 2019.
- [18] G. Traven, G. Matijević, T. Zwitter et al., "The galah survey: classification and diagnostics with t-SNE reduction of spectral information," *The Astrophysical Journal Supplement Series*, vol. 228, no. 2, p. 24, 2017.
- [19] E. Pouyet, N. Rohani, A. K. Katsaggelos, O. Cossairt, and M. Walton, "Innovative data reduction and visualization strategy for hyperspectral imaging datasets using t-SNE approach," *Pure and Applied Chemistry*, vol. 90, no. 3, pp. 493–506, 2018.
- [20] W. Liu, P. Zhao, C. Wu, C. Liu, J. Yang, and L. Zheng, "Rapid determination of aflatoxin B1 concentration in soybean oil using terahertz spectroscopy with chemometric methods," *Food Chemistry*, vol. 293, pp. 213–219, 2019.
- [21] G. Dimitriadis, J. P. Neto, and A. R. Kampff, "t-SNE visualization of large-scale neural recordings," *Neural Computation*, vol. 30, no. 7, pp. 1750–1774, 2018.
- [22] E. Roman-Rangel and S. Marchand-Maillet, "Inductive t-SNE via deep learning to visualize multi-label images," *Engineering Applications of Artificial Intelligence*, vol. 81, pp. 336–345, 2019.