Hindawi
Mathematical Problems in Engineering
Volume 2023, Article ID 5283437, 11 pages
https://doi.org/10.1155/2023/5283437

Research Article

Research on the Multilabel Data Stream Classification Method Based on Fuzzy Complex Set-Valued Measure Learning

Hai Wei Wu, Da Wei Yun, and Feng Run Gu

Hainan Vocational University of Science and Technology, Danzhou 571126, Hainan, China

Correspondence should be addressed to Hai Wei Wu; whaiwei2022@163.com

Received 25 August 2022; Revised 6 September 2022; Accepted 7 September 2022; Published 21 April 2023

Academic Editor: Baiyuan Ding

In the process of modern industrial production equipment developing towards the direction of structure, automation, and intelligence, motor is still the main power output equipment. If the data flow classification occurs during the operation of the motor, it will lead to problems such as the reduction of its operation efficiency and the increase in system energy consumption. In serious cases, it will even cause motor damage, and the overall system equipment will be shut down for maintenance for a long time, resulting in serious economic losses. Therefore, the research on intelligent multilabel data stream classification technology of motor is of great significance to ensure the stability and reliability of efficient operation of production equipment. In order to improve the recognition efficiency and accuracy of ms-1dcnn's multilabel data stream classification method in the environment of variable motor conditions and strong noise interference, a multiscale feature fusion framework is constructed based on the residual network structure. The implementation principles of two kinds of attention mechanism algorithms, squeeze and excitation module and convolution attention module, are studied, respectively. The attention module suitable for one-dimensional residual network is designed and embedded into the residual module to build a multiscale attention residual network model. Finally, the effectiveness and superiority of the proposed model are verified by using the experimental platform data.

1. Introduction

1.1. Research Background and Significance. As important power equipment in various production fields, the motor has the advantages of low price, relatively simple overall structure, relatively reliable, and so on, and undertakes more than 80% of the kinetic energy output in the process of modern industrial and agricultural production [1]. Especially in metallurgy, mining, machining, rail transit, and other industrial production fields, the installed capacity is huge and widely used. Motors and their related power equipment are important assets of enterprises, and their reliability and stability during operation are the key to ensure the safe and stable operation of mechanical equipment for a long time [2]. If the data flow classification of the motor occurs during the operation of the equipment, it will lead to unstable operation, sharp rise in energy consumption, and other conditions. In serious cases, it will even cause damage to the motor and equipment, and then affect the normal operation of the whole equipment. Sudden shutdown and maintenance must be carried out, resulting in problems such as slow production progress and economic losses. The regular maintenance of the motor is usually aimed at the inspection and maintenance of the motor, which has a certain positive significance in reducing the classification rate of the data flow of the motor, but there are problems such as insufficient maintenance, and blind repair [3]. At the same time, the traditional signal processing and threshold judgment methods using sensor signals are not enough in the process of industrial production in a complex environment, and because the early characteristics of data flow classification are weak or disturbed by noise, there are often cases of inadequate response, motor damage, and production pause when data flow classification is found, resulting in huge property losses and even casualties [4]. The development requirements of automation and intelligence of modern industrial equipment have led to the continuous innovation of multilabel data stream classification technology. At this stage, the intelligent multilabel data stream classification method based on signal processing, artificial intelligence, and other technologies has the advantages of high recognition accuracy and strong adaptability and is a hot issue in the field of multilabel data stream classification [5]. In the "made in China 2025" proposed in 2021, equipment intelligent diagnosis technology is listed as one of the important related technologies in the field of intelligent manufacturing, and it is an important technical means to realize the safe and stable operation of intelligent equipment [6, 7]. Therefore, it is of great practical significance to analyze the data flow classification mechanism of a motor and study the relevant intelligent multilabel data flow classification methods for maintaining the safe, stable, and efficient operation of industrial production.

At present, motor multilabel data stream classification technology can generally be divided into three categories: model-based, signal processing and machine learning, and deep learning [8]. At the beginning of the development of multilabel data stream classification technology, people have performed a lot of research on model-based multilabel data stream classification technology, but because it usually needs to build a mathematical model that can accurately reflect the actual running state of the motor, there are obvious defects and great difficulties. The traditional intelligent multi label data stream classification technology based on signal processing and machine learning has a good operability and recognition effect and has been widely used in the field of multilabel data stream classification. In recent years, with the continuous breakthroughs of computer, artificial intelligence, and other information technologies, as well as the requirements of modern industrial development, intelligent multilabel data stream classification technology based on deep learning algorithm has received extensive attention and research. It is a new research hotspot in the field of multilabel data stream classification and is in the process of rapid development [9, 10].

Based on the data of SIQ-MFS mechanical data flow classification test-bed, aiming at a variety of common types of motor data flow classification in the actual production process, this topic uses the motor vibration signal to realize the research on the motor data flow classification method. In the research of traditional intelligent multilabel data stream classification methods, it is usually necessary to analyze the collected data stream classification signals, extract the features that can reflect the classification status of the real data stream and further filter or reduce the dimension of the features through human intervention and finally get the diagnosis results through the pattern recognition algorithm. Based on the analysis of the generation mechanism of motor data stream classification, the EEMD method is used to analyze and extract the characteristics of motor vibration signals, and a feature selection method based on improved and adjusted mutual information is proposed for multilabel data stream classification. Aiming at the problems of a complex process and human intervention in traditional intelligent multilabel data stream classification methods, a one-dimensional convolution multiscale feature fusion framework is constructed, and an end-to-end intelligent motor multilabel data stream classification based on multiscale one-dimensional fuzzy complex value measure learning is proposed. A multiscale residual network model based on attention mechanism is proposed to solve the common problems of variable working conditions and strong noise interference in practical applications. The experimental results show that the proposed method can effectively realize the classification of multilabel data flow of motor, and has a good robustness and diagnostic effect under variable working conditions and noise interference, which is of practical significance to ensure the safe and stable operation of the motor.

1.2. Research Status

1.2.1. Research Status of Motor Multilabel Data Stream Classification Technology. Multilabel data stream classification is an important technical means to ensure the safe and smooth operation of equipment. It is a hot research direction of interdisciplinary and is in the process of continuous development. In the 1960s, through the monitoring of parameters such as current, voltage, vibration frequency, radial speed, and shaft radial flux, the classification characteristics of motor data flow were studied, and the purpose of multilabel data flow classification was achieved by threshold judgment [11]. Based on various measurement sensors, many researchers have adopted different monitoring methods, combined with signal processing, Internet of things and artificial intelligence technology designed a variety of multilabel data flow classification methods for the various parameters and possible operating conditions of mechanical equipment, which has led to the rapid development of equipment multilabel data flow classification technology, and the accuracy and speed of motor multilabel data flow classification have been improved. It plays a great role in promoting the safe operation of the motor [12, 13].

1.2.2. Multilabel Data Stream Classification Method Based on Signal Processing and Machine Learning. At present, in the process of studying the multilabel data stream classification method based on signal processing and machine learning, because the vibration signal of the motor will show complex nonstationary and nonlinear characteristics when it is in the data stream classification state, in order to fully reflect the classification characteristics of the data stream in the signal, it is necessary to use the time-frequency analysis method to analyze the changes of its vibration signal distribution parameters with time. In the process of multilabel data stream classification, time-frequency analysis methods such as short-time Fourier transform, wavelet transform, and Hilbert Huang transform are usually used to analyze vibration signals [14–16], obtain signal sequences that can reflect the time-frequency characteristics of data stream classification types and calculate their different types of statistical characteristics through a variety of methods, so as to build a data stream classification feature set for motor multilabel data stream classification. Huang et al. proposed to use iterative empirical wavelet packet (EWT) to analyze the gearbox data stream classification signal and used the multi iterative mutual information energy entropy to denoise and reconstruct the decomposed signal components. Finally, the sparse filter fuzzy complex value measure learning model is used to extract the effective data stream classification features and realize multilabel data stream classification [17]. Zhaohua and Huang proposed an adaptive frequency window empirical wavelet transform method to segment the signal spectrum through the frequency window and adaptively select the position of the window and then extract the early features of data flow classification by using the empirical wavelet packet [18]. Cheng-Chien et al.first performed EEMD decomposition on the bearing vibration signal, screened the IMF component based on the calculation of the correlation coefficient and combined it with the distance factor and reconstructed the signal, so as to eliminate the noise interference in the original signal and finally realized the multilabel data stream classification through the wavelet packet and the correlation coefficient [19]. Imaouchen et al. proposed to improve the EMD method by adding complementary noise. On the basis of overcoming the problem of modal aliasing, this method can effectively eliminate the noise residue in the algorithm process and improve the calculation efficiency of EEMD. Finally, IMF component reconstruction signals related to data stream classification features are selected, and amplitude and frequency modulation are extracted from the selected IMF by frequency weighted energy operator method for multilabel data stream classification [20, 21]. Among the above time-frequency signal analysis methods, the EMD method has many characteristics, such as intuitive, posterior, and adaptive. It can effectively decompose the nonstationary vibration signal into signal components of different frequencies in turn. It is one of the most commonly used time-frequency analysis methods. However, when the signal to be decomposed contains many discontinuous signals with different frequencies, the mode aliasing problem will appear in EMD decomposition. Therefore, in practical applications, EMD will be improved by noise-aided analysis and other methods [22].

1.2.3. Research Status of Fuzzy Complex Valued Measure Learning. In recent years, because the key problems of the deep learning algorithm in the training process have been solved, it has made outstanding achievements in image, computer vision, natural language processing, and many other fields by virtue of its incomparable advantages with other technologies, as well as a variety of algorithms that are applicable to various scenes and have a variety of different structure types.

Common deep learning algorithms include restricted Boltzmann machine, automatic encoder, and fuzzy complex value measure learning, on this basis, many researchers have proposed a variety of improved algorithms for different problems and scenarios, which greatly promotes the development and improvement of deep learning. Generally, deep learning algorithm is a kind of network structure with multiple hidden layers, which can realize the effective analysis of input data, learn the inherent characteristics and attributes contained therein, establish the mapping relationship between input data and characteristics, have excellent intelligent learning ability and have been widely used in many fields. Among them, fuzzy complex valued measure learning is one of the most widely used deep learning algorithms. Its excellent application potential has been widely valued and caused many scholars to conduct in-depth research [23].

Fuzzy complex value measure learning is convenient for processing a large amount of data and can automatically extract the deep features in the signal and realize the automation and intellectualization of multilabel data stream classification. As an end-to-end fuzzy complex value measure learning model, it overcomes the problems of low efficiency and effect of multilabel data stream classification caused by complex steps and too many human interventions of traditional intelligent multilabel data stream classification methods to a certain extent. Although fuzzy complex set-valued measure learning has many advantages and is a hot topic of current research, its research in the field of multilabel data stream classification is still in the process of continuous development. More and more in-depth research is needed on how to design a fuzzy complex set-valued measure learning structure suitable for multilabel data stream classification and how to improve the problems of variable working conditions and noise interference in the process of multilabel data stream classification [24].

1.3. Research Route.

(1) The whole research of this paper can be divided into two parts: multilabel data stream classification method based on signal processing and deep learning. In the multilabel data stream classification method based on signal processing, the time-frequency signal analysis and the corresponding spectral analysis methods are studied, respectively. Nine time-frequency and frequency-domain statistical features are selected to construct the feature set, and the sensitive features are screened. Finally, the dimension reduction and pattern classification algorithm are used to realize the construction of the motor diagnosis model. In the multilabel data stream classification method based on deep learning, the application of one-dimensional fuzzy complex value measure learning in the field of multilabel data stream classification is studied, and multiscale one-dimensional fuzzy complex value measure learning is constructed for motor multilabel data stream classification. Aiming at the problems of variable working conditions and noise interference of the motor, the attention mechanism and multiscale residual network are used to build a diagnostic model. Finally, the four common motor data flow classifications collected by the SQI-MFS data flow classification experimental platform and the vibration data under a normal motor state are fully verified by experiments [25].

(2) The time-frequency analysis method EMMD of vibration signals is studied, and the time-frequency features are extracted to construct the original

feature set. In view of the low correlation of its features, which cannot accurately reflect the classification state of data flow, and the redundancy between various features, a feature selection method AMISR based on improved and adjusted mutual information is proposed. The adjusted mutual information value obtained after feature clustering analysis is combined with the standard deviation as the feature evaluation index, quantitative analysis of feature sensitivity. On this basis, the LDA dimensionality reduction algorithm is used to map high-dimensional features to low-dimensional space, and the EMMD-AMISR-LDA-SVM model is constructed to realize the classification of motor multilabel data flow. The effectiveness of the proposed method is verified by two groups of comparative experiments.

(3) Aiming at the problems existing in the traditional intelligent multilabel data stream classification method based on signal processing and machine learning, such as complex process, too much intervention, weak generalization ability, and poor adaptability, this paper further studies the multilabel data stream classification method based on one-dimensional fuzzy complex value measure learning by referring to the current popular deep learning method and combining the one-dimensional attribute of vibration signal. The multiscale feature fusion framework is studied, and an end-to-end multiscale one-dimensional fuzzy complex value measure learning data flow classification diagnosis model is proposed. The stability, generalization performance, and anti-noise ability of the model are verified by motor off-duty and noise interference experiments [25].

2. Classification Principle and Signal Analysis of Multilabel Data Flow

2.1. Analysis of Data Flow Classification Mechanism

2.1.1. Classification Mechanism Analysis of Rotor Broken Bar Data Flow. Because the rotor is subjected to thermal, mechanical, and other stresses during operation, it is easy to classify the broken bar data flow of the rotor after long-time operation or frequent start-up and stop operations. Because the system is relatively complex, the vibration parts and corresponding vibration signals produced by different data stream classification types are different [19].

When operating in the state of no data flow classification, the stator current frequency only contains the main frequency components. When rotor bar breaking data flow classification occurs, electromagnetic and mechanical vibration will occur inside, resulting in harmonic components in the spectrum of stator current. Assuming that the power supply frequency is, the fundamental wave of the three-phase synthetic magnetomotive force of the stator current can be expressed as

$$F_1 = K_1 N_1 I_1 \sin(2\pi - p\theta). \quad (1)$$

The phase of the rotor winding can be obtained as $\varphi = \theta - 2\pi f_r t$, then according to the balance relationship between the stator and rotor magnetomotive force, we get

$$\begin{aligned} T_1(t,\theta) = \frac{K_2 N_2 I_2}{2}\{&\cos[2\pi(f_1 + 2f_r)t - 3\theta] \\ &-\cos[2\pi(f_1 - 2f_r)t + \theta]\}. \end{aligned} \quad (2)$$

When the number of pole pairs P is 1, the slip is

$$\begin{aligned} s &= \frac{(f_1 - f_r)}{f_1}, \\ T_1(t,\theta) &= \frac{K_2 N_2 I_2}{2}\{\cos[2\pi(3-2s)f_1 t - 3\theta] \\ &\quad -\cos[2\pi(1-2s)f_1 t - \theta]\}. \end{aligned} \quad (3)$$

It can be seen from the above analysis that when the rotor bar breaking data flow classification occurs, the frequency in the stator current spectrum will be $f_{bp} = (1-2s)f$. The harmonic component is specifically manifested in the periodic fluctuation of the stator current due to imbalance, which causes the periodic vibration of 2sf in rotor torque and speed, resulting in mechanical vibration.

2.1.2. Analysis of Data Flow Classification Mechanism of Stator Winding. Stator winding data flow classification is usually caused by the damage and falling off of the insulation paint on the surface of the conductor due to long-term high-temperature and corrosive operating environment. It is a common type in asynchronous data flow classification.

$$\begin{aligned} F_e &= \frac{f}{p} \times 2p = 2f, \\ f_s &= (1 \pm 2k(1-s))f, \\ f_s &= (n \pm 2k(1-s))f \quad n = 1,2,3\ldots, k = 0,1,2\ldots. \end{aligned} \quad (4)$$

It can be seen from the above analysis that when the stator winding data flow classification occurs, the harmonic component in the stator current is significantly enhanced, resulting in a certain change in the frequency of vibration. Therefore, the vibration signal can be used as an effective data source that can directly reflect the data flow classification state [26].

2.1.3. Bearing Data Flow Classification Mechanism Analysis. When the bearing has local surface type damage, the characteristic frequency that can reflect its own data flow classification state will appear in the data flow classification vibration signal and show the characteristics of periodicity and decreasing amplitude. Therefore, some signal time-frequency analysis methods can be used to obtain the characteristic information from the nonstationary data stream classification vibration signals with complex frequency components, so as to lay a foundation for constructing the data stream classification feature set and realizing the

Table 1: Four kinds of bearing fault characteristic frequency calculation formulas.

Data flow classification type	计算公式	公式编号
Inner circle data flow classification	$f_i = nr/120(1 + d/\mathrm{D}\sin\alpha)$	(8)
Outer ring data flow classification	$f_o = nr/120(1 - d/\mathrm{D}\cos\alpha)$	(9)
Rolling element data flow classification	$f_b = r/120(1 - d/\mathrm{D}\cos\alpha)$	(10)
Cage data flow classification	$f_c = Dr/120\ d[1 - (d/D)^2\cos^2\alpha]$	(11)

accurate multilabel data stream classification. The damage of different structural components of the bearing will lead to different types of bearing data flow classification, and the corresponding theoretical data flow classification characteristic frequencies are also different. The specific calculation formulas of four different bearing data flow classifications are shown in Table 1.

2.2. Vibration Signal Analysis Method. The frequency characteristics of data stream classification vibration signals are time-dependent, so it is necessary to analyze the change of signal frequency with time. Therefore, time-frequency analysis method is needed to obtain local and global information in the time domain and the frequency domain. In this section, the basic principles of empirical mode decomposition and its improved method are introduced in detail, and simple experimental verification is carried out through example signals [27].

2.2.1. Empirical Mode Decomposition. Empirical mode decomposition (EMD), proposed by Huang in 2021, is a part of Hilbert Huang transform. Because of its strong time-frequency analysis ability and adaptive characteristics, it has been widely used in many fields, including multilabel data stream classification. EMD decomposes the original vibration signal into a series of instantaneous frequency effective single component signals and a set of residual signals according to different frequency components. The single component signal is called the eigenmode function component, and each IMF component only contains the local characteristic signals of the same time scale in the original signal. In order to verify that the single component signal obtained is an IMF component, it is usually judged according to the following two conditions:

(1) Within the length of the signal sequence, the number of extreme points and the number of zero crossings are equal or at most the difference is no more than 1.

(2) For any time, the mean value of the local envelope of the signal is 0.

The specific steps of EMD decomposition of the original signal are as follows:

$$\begin{aligned} h_1(t) &= x(t) - m_1(t),\\ r_2(t) &= r_1(t) - c_2(t), \cdots, r_m(t)\\ x(t) &= \sum_{k=1}^{m} c_k(t) + r_m(t). \end{aligned} \tag{5}$$

From the decomposition process of the vibration signal, it can be seen that the sequence in which each IMF component is decomposed corresponds to the height of the component frequency component, that is, the IMF component decomposed first contains the high-frequency component in the signal, which contains obvious and important characteristic information.

2.2.2. Set Empirical Mode Decomposition. Verify the effect of EEMD on suppressing mode aliasing through the simulation signal and set the simulation signal as

$$x(t) = x_1(t) + x_2(t) + n(t). \tag{6}$$

Among them, $x_1(t) = \sin(100\pi t + \pi/2), x_2 2 = \sin(20\pi t)$, $n(t)$. It is intermittent noise signal.

Due to the superposition of intermittent noise signal components in the mixed signal, after EMD decomposition, it can be clearly seen that there are frequency components with greatly different time scales in imf1 to imf4 components, and there is modal aliasing, which fails to achieve the decomposition goal of the EMD algorithm.

By repeatedly using Gaussian white noise in the algorithm flow to smooth the intermittent signals in the signal to be decomposed, the accurate decomposition of signals with different frequency components can be achieved, and the problem of modal aliasing can be effectively suppressed.

3. Feature Selection and the Multilabel Data Stream Classification Method Based on Improved and Adjusted Mutual Information

3.1. Introduction. For the problems of high dimension and feature redundancy of a feature set, two representative dimensionality reduction algorithms, unsupervised principal component analysis (PCA) and supervised linear discriminant analysis (LDA), are studied to realize the mapping from high-dimensional features to low-dimensional feature space, which is convenient for the training and testing of the classification algorithm. Finally, combined with two classical machine learning classification algorithms, support vector machine and limit learning machine, a multilabel data stream classification model is constructed and verified by two groups of comparative experiments. The experimental results show that the model proposed in this chapter can achieve better recognition results.

3.2. Improved Feature Selection Method. Feature selection analyzes various features in the original feature set and uses some algorithms or indicators to select features that can reflect the classification status of data flow, so as to improve the diagnostic performance. Feature selection also reduces the dimension of feature set, avoids model over fitting and accelerates the operation speed of algorithm model. In this section, based on extracting the original feature set from the original signal sample by using the methods of set empirical mode decomposition (EEMD), spectral analysis, and statistical analysis, and clustering it by using the fuzzy C-means algorithm. A feature selection method that introduces the standard deviation into the cluster evaluation index adjusted mutual information (AMI) is proposed. The original features are evaluated quantitatively to select effective features for classification.

3.2.1. Fuzzy C-Means. Fuzzy C-means (FCM) realizes the soft interval division of all sample points in the data set by introducing membership weight, so that samples of the same category gather around the cluster center and has a higher tolerance for ambiguous samples in the sample points. After initializing the cluster center according to the sample data, the cluster center and the membership weight of each sample are iteratively updated by minimizing the objective function. When the position of the cluster center gradually tends to be stable during multiple iterative operations, and its change difference is within the set range, the algorithm update is stopped [28]. The objective function formula of FCM is shown in the following formula:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\| x_i - c_j \right\|^2, \quad 1 \le m < \infty. \tag{7}$$

Here, membership degree is U_{ij} and cluster center is C_j. The calculation formula of iterative update is expressed as

$$\begin{aligned} u_{ij} &= \frac{1}{\sum_{k=1}^{C}\left(\left\|x_i - c_j\right\| / \left\|x_i - c_k\right\|\right)^{2/m-1}}, \\ c_j &= \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}. \end{aligned} \tag{8}$$

The termination condition of iteration update is

$$\max_{ij}\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right|\right\} < \delta. \tag{9}$$

Because FCM is an unsupervised algorithm, there is a problem of the validity test. Therefore, it is necessary to evaluate the performance of the algorithm through some clustering evaluation indicators. The commonly used clustering evaluation indicators include entropy, Pearson correlation coefficient, mutual information, and Davies–Bouldin index.

3.2.2. Adjust Mutual Information. Mutual information is used to express the relationship between information. It is a measure of statistical correlation between random variables. Its definition is similar to cross entropy. It is used to measure the consistency of two data distributions and as an evaluation index of data clustering results. Assuming that and are the distributions of samples, the entropy of the two distributions are

$$H(U) = \sum_{i=1}^{|U|} P(i)\log(P(i)), H(V) = \sum_{j=1}^{|V|} P'(j)\log(P'(j)). \tag{10}$$

Then the mutual information between u and V can be expressed as

$$MI(U,V) = \sum_{i=1}^{|U|}\sum_{j=1}^{|V|} P(i,j)\log\left(\frac{P(i,j)}{P(i)P'(j)}\right). \tag{11}$$

The closer the AMI value is to 1, the more consistent the sample point is with the actual category.

$$\text{AMI} = \frac{MI - E(MI)}{\text{mean}(H(U),H(V)) - E(MI)}. \tag{12}$$

3.2.3. Feature Selection Method. The implementation steps of AMISR are as follows.

There are C types of data flow classification. The vibration signals of all states are analyzed by time-frequency analysis and spectral analysis. A total of N signal sample sequences corresponding to each data stream classification type are obtained, and a variety of statistical features are calculated to construct the original feature set $[\mathrm{OF}^1, \mathrm{OF}^2, \ldots, \mathrm{OF}^M]$, then the matrix formed by the mth feature of all n sample signals is shown in the following formula:

$$OF^m = \begin{bmatrix} F_{11}^m & F_{11}^m & \cdots & F_{1N}^m \\ F_{21}^m & F_{22}^m & \cdots & F_{2N}^m \\ \vdots & \vdots & \ddots & \vdots \\ F_{C1}^m & F_{C2}^m & \cdots & F_{CN}^m \end{bmatrix}, \tag{13}$$

$$SD_c^m = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(S_{ci}^m - \overline{S}_c^m\right)^2}, \tag{14}$$

$$\overline{S}_c^m = \frac{1}{N}\sum_{i=1}^{N}\left(S_{ci}^m\right). \tag{15}$$

Then, after calculation, the standard deviations of all C states are constructed in turn, and the standard deviation set of the mth feature is obtained as $[SD_1^m, SD_2^m, \ldots S_D^C]$, and sum it to get

$$\text{SSD}(m) = \sum_{j=1}^{M} SD_j^m. \tag{16}$$

After calculating the AMI value and standard deviation of all m statistical characteristics, the calculation formula of AMISR is defined to obtain the AMISR value sequence, and the calculation method is shown in the following formula:

$$\mathrm{AMISR}(m) = \frac{\mathrm{AMI}(m)}{SD(m)} \quad m = 1, 2, \ldots, M. \tag{17}$$

Through the above analysis process, it can be seen that the larger the AMISR value of the statistical feature, the higher its correlation with the current data flow classification type and has stronger feature expression ability, which indicates that the features screened by AMISR have the advantages of high degree of aggregation and low degree of deviation within the class. Therefore, using AMISR to evaluate the original feature set and screen out highly relevant features is conducive to improving the effect of multilabel data stream classification [29].

3.3. Extreme Learning Machine. Limit learning machine was proposed by Huang–Guangbin and others in 2004. It is a single hidden layer feedforward fuzzy complex set-valued measure learning model. In the training process, the connection weight between the input layer and the hidden layer of ELM and the threshold of hidden layer neurons are generated by a random algorithm, so it is only necessary to set the number of hidden layer nodes without iteratively updating the above parameters. Compared with the traditional feedforward fuzzy complex set-valued measure learning, ELM has a faster training speed and a better generalization performance while ensuring accuracy.

Figure 1 shows the typical structure of ELM, which is usually composed of input layer, hidden layer, and output layer. The neurons between different network layers are fully connected.

Suppose a single hidden layer fuzzy complex set-valued measure learning with L hidden nodes can be expressed as

$$\begin{aligned}
&\sum_{i=1}^{L} \beta_i g\left(W_i \cdot X_j + b_i\right) = o_j \quad j = 1, 2, \ldots, N,\\
&H\left(W_1, W_2, \ldots W_L; b_1, b_2, \ldots, b_L; X_1, X_2, \ldots X_L\right)\\
&\quad = \begin{bmatrix} g\left(W_1 \cdot X_1 + b_1\right) & \ldots & g\left(W_L \cdot X_1 + b_L\right) \\ \vdots & \ldots & \vdots \\ g\left(W_1 \cdot X_N + b_1\right) & \ldots & g\left(W_L \cdot X_N + b_L\right) \end{bmatrix},\\
&\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L\times m},\\
&T = \begin{bmatrix} T_1^T \\ \vdots \\ T_N^T \end{bmatrix}_{N\times m}.
\end{aligned} \tag{18}$$

3.4. Support Vector Machine. In the course of decades of development, support vector machine has been improved and perfected many times by hard margin, soft margin, kernel function, and other methods, gradually theorized as a part of statistical learning theory and has been successfully applied in pattern classification tasks in many fields, as shown in Figure 2.

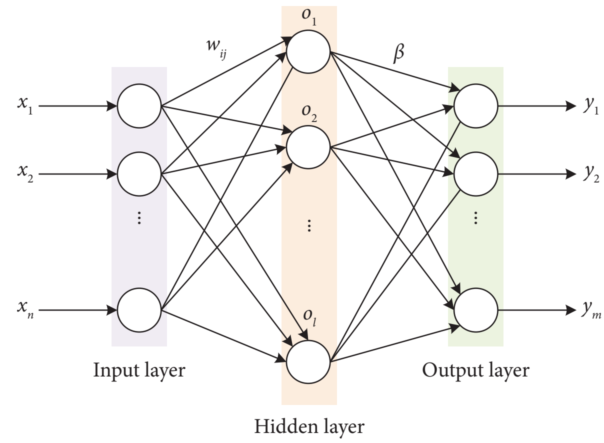

Figure 1: Structure of extreme learning machine.

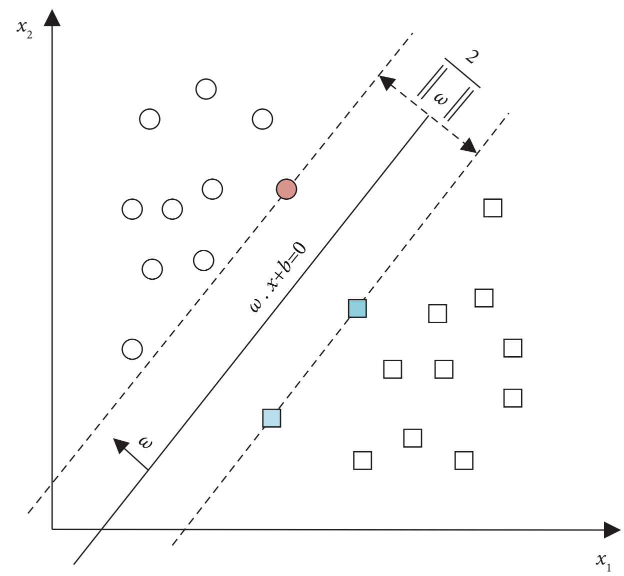

Figure 2: Linear support vector machine partitioning interface.

In order to divide as many sample data as possible correctly, SVM must first obtain the optimal separation hyperplane that maximizes the spacing between different categories of samples.

4. Multilabel Data Stream Classification Method Based on Multiscale One-Dimensional Fuzzy Complex Measure Learning

4.1. Introduction. Traditional intelligent multilabel data stream classification methods based on signal analysis and machine learning usually have the problems of a complex diagnosis process and too much human intervention. Many end-to-end intelligent multilabel data stream classification methods based on deep learning have the advantages of strong feature learning ability, simple diagnosis process, and high recognition accuracy, and gradually become the research hotspot in the field of multilabel data stream classification. As one of the widely used deep learning algorithms, fuzzy complex measure learning has not only made outstanding achievements in image related fields but also many researchers have realized its great research and application potential in the field of multilabel data stream classification.

4.2. Fuzzy Complex Valued Measure Learning Structure. Fuzzy complex valued measure learning (CNN) is a typical and representative deep learning algorithm, which was first proposed by Alexander Waibel et al. In 1987, CNN is a feedforward multilayer fuzzy complex set-valued measure learning model. Its main structure is usually composed of three parts: convolution layer, pooling layer, and full connection layer. The convolution layer and pooling layer mainly perform convolution, pooling, and other operations on the input signals of the network to realize feature learning. Then, the feature map obtained is spread and connected through the full connection layer, and the classification results are output in combination with the softmax function layer.

4.2.1. Convolution Layer. The convolution layer usually contains a set of convolution kernels. Each convolution kernel only convolutes the local area of the input signal or the feature map and then integrates the local features in the next level network to obtain the global feature map. Another main feature of the convolution layer is weight sharing, that is, each convolution kernel must traverse the output of the previous layer, and the weight parameters of any convolution kernel remain unchanged in this process. Weight sharing can effectively reduce the network parameters of the convolution layer, reduce the amount of calculation required for model training and avoid over fitting the model.

4.2.2. Activation Layer. The nonlinear mapping of the input features of the convolution layer is realized through the activation function, and the features in the original multidimensional space are mapped to another space to increase the linear separability of the data.

4.2.3. Pool Layer. Common pooling functions include Max pooling function and average pooling function. Maximum pooling refers to taking the maximum value of the activation value in the local receptive field as the output value, while average pooling refers to taking the average value of the activation value in the local receptive field as the output value. The specific pooling operation is shown in Figure 3, and the filter size is 2×2. The step size is 2.

4.2.4. Full Connection Layer. In the fuzzy complex set-valued measure learning model, the output of the previous layer of the network is spread and connected to realize the mapping from the feature to the sample tag space, forming the input of the full connection layer, and then through the hidden layer using the ReLU function as the activation function, combined with the softmax function to realize the output of the classification results. Its structure is shown in Figure 4.

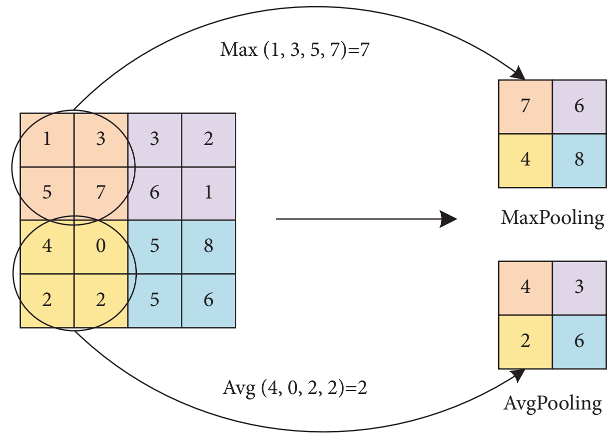

Figure 3: Pooling operation diagram.

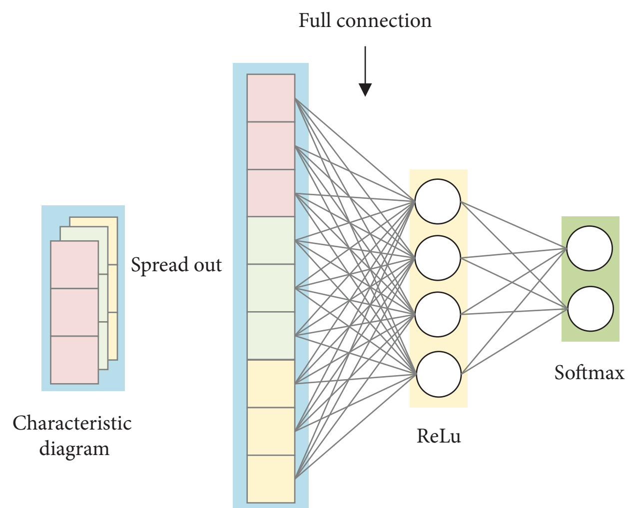

Figure 4: Schematic diagram of fully connected layer.

4.2.5. Loss Function. In the learning of fuzzy complex set-valued measure, peers use the loss function to measure the consistency between the output estimated value of the network model and the probability distribution of the target value. The smaller the loss function value is, the better the fitting degree of the network model to the training set samples is, as shown in Figure 5.

4.3. One-Dimensional Fuzzy Complex-Valued Measure Learning Structure. In the field of image, because the large convolution kernel will greatly increase the amount of calculation of network parameters, which is not conducive to the deepening of the model depth and will reduce the computational performance, the two-dimensional fuzzy complex-valued measure learning is usually constructed by stacking multiple small convolution kernels. For example, two 3's are usually used × The convolution layer stack of 3 replaces 1 with 5×5, the receptive field size remains unchanged, but the parameters are greatly reduced. However, for one-dimensional vibration signals, the amount of parameter calculation will increase after the stacking of smaller convolution cores

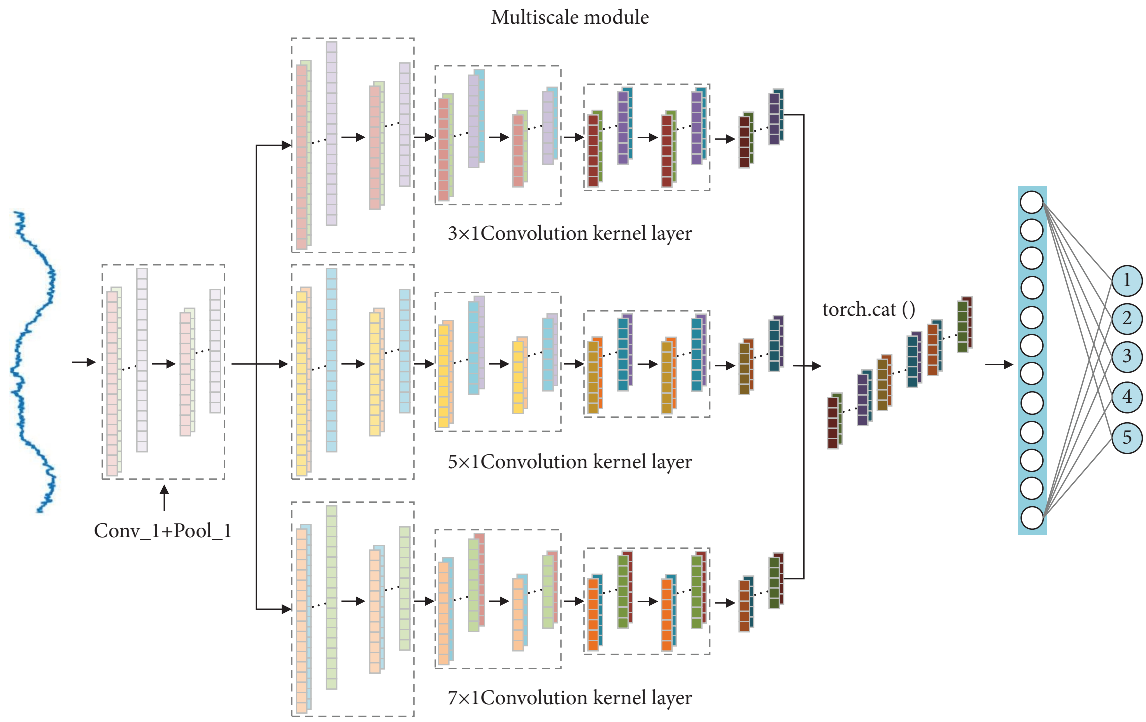

FIGURE 5: Multiscale one-dimensional convolutional neural network structure diagram.

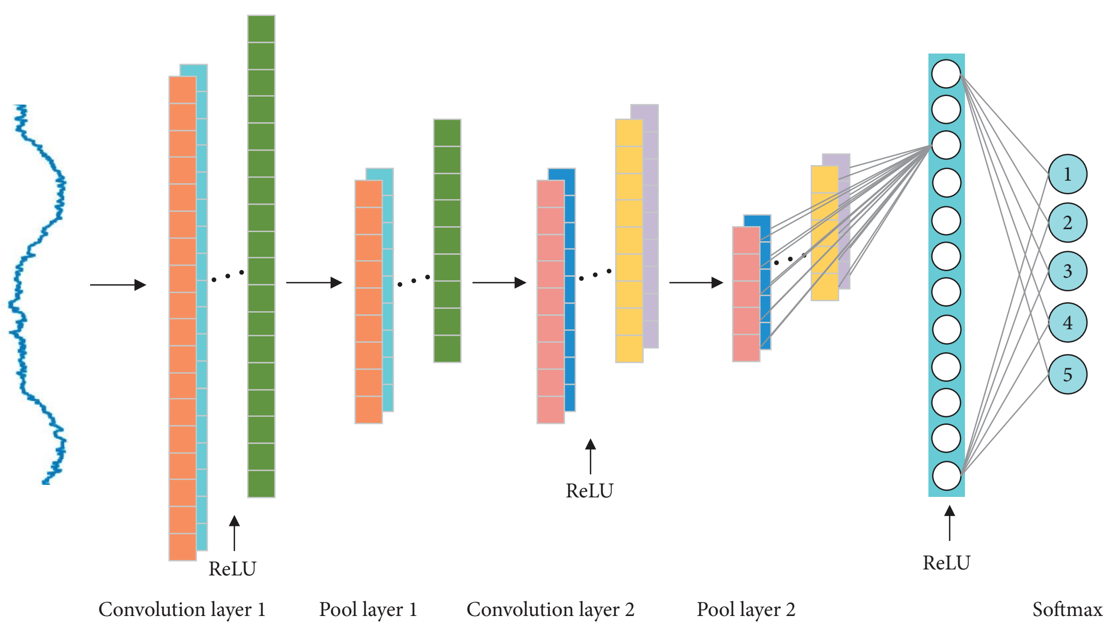

FIGURE 6: Structure of the one-dimensional convolutional neural network.

and convolution layers. Therefore, it is an appropriate method to use the one-dimensional convolution structure to build a multilabel data stream classification model for vibration signals. The common learning structure of one-dimensional fuzzy complex set-valued measures is shown in Figure 6.

4.4. Multiscale Feature Fusion. Fuzzy complex set-valued measure learning has strong feature information extraction and model fitting capabilities. The input data passes through CNN's convolution layer, activation layer, and pooling layer, respectively. The number of feature channels increases and the size of the feature map decreases. At the same time, the feature information extracted by convolution cores of different sizes is also different. Therefore, fusing the features extracted by convolution cores of different sizes can improve CNN's feature expression ability at various levels.

The multiscale fusion framework is constructed by using convolution kernels of different sizes to realize the complementarity of different scale information. Each layer of the multiscale framework is composed of convolution kernels of different sizes. From top to bottom, the size of convolution kernels is 3×1, 5×1, 7×1, and each layer has three groups of the same convolution layer and the pooling layer, which are used to extract the same scale features, and then through a 16×1, and finally through torch Cat() function performs multiscale feature fusion.

5. Summary

Using the vibration data of SQI-MFS mechanical data flow classification test-bed, this paper makes an in-depth study on the current popular intelligent multilabel data flow classification methods and constructs a variety of multilabel data flow classification models and verifies and analyzes them through experiments. The main research contents of the dissertation are summarized as follows.

Empirical mode decomposition (EMD), which is widely used in time-frequency signal analysis is studied, and its shortcomings are clarified. The mixed signal is constructed and tested to verify the decomposition effect of the improved algorithm set empirical mode decomposition (EEMD). After decomposing the vibration signal, it is proposed to use the IMF component and its marginal spectrum and envelope spectrum obtained by Hilbert transform to calculate 9 different statistical features to construct the original feature set.

Aiming at the adaptability of the model to changing conditions and strong noise interference, the ms-1dcnn model is further improved by the residual network and the attention mechanism. The multiscale feature fusion framework and the overall network structure are constructed by using the residual module to improve the efficiency and accuracy of network recognition. Se and CBAM attention modules are introduced and embedded into the residual module, respectively, and a multiscale residual network model based on attention mechanism is proposed. The algorithm not only has good robustness and recognition effect under variable working conditions but also can achieve more than 85% recognition accuracy under the strong noise interference of $\text{SNR} = -2$ db.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the present study.

Acknowledgments

The study was supported by the High level Talents Project of the Hainan Natural Science Foundation, 622RC727, 2022.

References

[1] C. Wang, M. Gan, and C. A. Zhu, “Fault feature extraction of rolling element bearings based on wavelet packet transform and sparse representation theory,” *Journal of Intelligent Manufacturing*, vol. 29, no. 4, pp. 1–15, 2021.

[2] H. Keskes, A. Braham, and Z. Lachiri, “Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet SVM,” *Electric Power Systems Research*, vol. 97, pp. 151–157, 2013.

[3] L. Dong Han, A. Jong Hyo, and K. Bong Hwan, “Fault detection of bearing systems through EEMD and optimization algorithm,” *Sensors*, vol. 17, no. 11, p. 2477, 2021.

[4] Y. Imaouchen, M. Kedadouche, R. Alkama, and M. A. Thomas, “A Frequency-Weighted Energy Operator and Complementary Ensemble Empirical Mode Decomposition for Bearing Fault detection,” *Mechanical Systems and Signal Processing*, vol. 82, 2021.

[5] D. Li, G. Hu, and C. J. Spanos, “A data-driven strategy for detection and diagnosis of building chiller faults using linear discriminant analysis,” *Energy and Buildings*, vol. 128, pp. 519–529, 2016.

[6] S. Juhamatti, D. Strömbergsson, J. Lundberg, and A. Thomson, “Detection and Identification of Windmill Bearing Faults Using a One-Class Support Vector Machine (SVM),” *Measurement*, vol. 137, 2021.

[7] S. S. Udmale and S. K. Singh, “Application of spectral kurtosis and improved extreme learning machine for bearing fault classification,” *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 11, pp. 1–12, 2021.

[8] E. Principi, D. Rossetti, S. Squartini, and F. Piazza, “Unsupervised electric motor fault detection by using deep autoencoders,” *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 02, pp. 108–118, 2021.

[9] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[10] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems*, vol. 25, no. 2, p. 116, 2021.

[11] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition*, IEEE Computer Society, Nashville, TN, USA, June 2021.

[12] L. Jing, M. Zhao, P. Li, and X. Xu, “A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring ofgearbox,” *Measurement*, vol. 111, pp. 1–10, 2017.

[13] R. Liu, G. Meng, B. Yang, C. Sun, and X. Chen, “Dislocated time series convolutional neural architecture: an intelligent fault diagnosis approach for electric machine,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1310–1320, 2017.

[14] J. Zhu, N. Chen, and W. Peng, “Estimation of bearing remaining useful life based on multiscale convolutional neural network,” *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3208–3216, 2019.

[15] L. Wen, X. Li, L. Gao, and Y. Zhang, “A new convolutional neural network-based data-driven fault diagnosis method,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2018.

[16] S. Guo, T. Yang, W. Gao, and C Zhang, “A novel fault diagnosis method for rotating machinery based on a convolutional neural network,” *Sensors*, vol. 18, no. 5, p. 1429, 2018.

[17] N. E. Huang, Z. Shen, S. R. Long et al., “The empirical mode decomposition and the Hilbert spectrum for nonlinear and

non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, vol. 454, pp. 903–995, 2021.

[18] W. U. Zhaohua and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 01, pp. 41–45, 2021.

[19] K. Cheng-Chien, L. Chien-Hsun, C. Hong-Chan, and K. J. Lin, "Implementation of a motor diagnosis system for rotor failure using genetic algorithm and fuzzy classification," *Applied Sciences*, vol. 7, no. 1, pp. 31–36, 2021.

[20] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *ProceeInternational Conference on International Conference on Machine Learning JMLR.org*, June 2021.

[21] L. Zeng, X. Xu, B. Cai, and S. Qiu, "Multi-scale Convolutional Neural Networks for Crowd Counting," in *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, China, September 2017.

[22] J. Hu, L. Shen, and S. Albanie, "Squeeze-and-Excitation Networks," in *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence*, Salt Lake City, UT, USA, June 2018.

[23] S. Woo, J. Park, and J. Y. Lee, *CBAM: Convolutional Block Attention Module*, Springer, Cham, Berlin, Germany, 2021.

[24] Q. Fu, B. Jing, P. He, S. Si, and Y. Wang, "Fault feature selection and diagnosis of rolling bearings based on EEMD and optimized Elman_AdaBoost algorithm," *IEEE Sensors Journal*, vol. 18, no. 12, pp. 5024–5034, 2018.

[25] G. Yangde, V. Francesco, L. Ming, and W. Song, "Multi-scale permutation entropy based on improved LMD and HMM for rolling bearing diagnosis," *Entropy*, vol. 19, no. 4, pp. 176–182, 2021.

[26] Z. Wei, Y. Wang, S. He, and J. Bao, "A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection," *Knowledge-Based Systems*, vol. 116, no. C, pp. 1–12, 2017.

[27] H. Liu, J. Zhou, Y. Zheng, W. Jiang, and Y. Zhang, "Fault Diagnosis of Rolling Bearings with Recurrent Neural Network-Based autoencoders," *ISA Transactions*, vol. 77, 2021.

[28] Y. Liu, H. Liu, S. Zhang, C. Liu, and W. Liu, "Application of computer big data and convolutional neural network technology in guiding the training standardization," in *Proceedings of the 2022 IEEE International Conference on Electrical Engineering Big Data and Algorithms (EEBDA)*, pp. 1184–1189, Changchun, China, February 2022.

[29] J. Zheng and C. Fan, "Sports Risk Analysis Based on Knowledge Discovery and Data Driven," *Security and Communication Networks*, vol. 2022, Article ID 8589200, 9 pages, 2022.