Enhancing Yarn Quality Wavelength Spectrogram Analysis: A Semi-Supervised Anomaly Detection Approach with Convolutional Autoencoder

Haoran Wang 1,2, Zhongze Han 2, Xiaoshuang Xiong 1,2,*, Xuewei Song 2 and Chen Shen 1,2,*

1 Hubei Key Laboratory of Digital Textile Equipment, Wuhan Textile University, Wuhan 430200, China; 2115053027@mail.wtu.edu.cn (H.W.)
2 School of Mechanical Engineering & Automation, Wuhan Textile University, Wuhan 430200, China; 2115373057@mail.wtu.edu.cn (Z.H.); 2215373024@mail.wtu.edu.cn (X.S.)
* Correspondence: cshen@wtu.edu.cn

Abstract: Abnormal detection plays a pivotal role in the routine maintenance of industrial equipment. Malfunctions or breakdowns in the drafting components of spinning equipment can lead to yarn defects, thereby compromising the overall quality of the production line. Fault diagnosis of spinning equipment entails the examination of component defects through Wavelet Spectrogram Analysis (WSA). Conventional detection techniques heavily rely on manual experience and lack generality. To address this limitation, this current study leverages machine learning technology to formulate a semi-supervised anomaly detection approach employing a convolutional autoencoder. This method trains deep neural networks with normal data and employs the reconstruction mode of a convolutional autoencoder in conjunction with Kernel Density Estimation (KDE) to determine the optimal threshold for anomaly detection. This facilitates the differentiation between normal and abnormal operational modes without the necessity for extensive labeled fault data. Experimental results from two sets of industrial data validate the robustness of the proposed methodology. In comparison to conventional Autoencoder and prevalent machine learning techniques, the proposed approach demonstrates superior performance across evaluation metrics such as Accuracy, Recall, Area Under the Curve (AUC), and F1-score, thereby affirming the feasibility of the suggested model.

Keywords: spinning machine; anomaly detection; deep learning; kernel density estimation; convolutional autoencoder

1. Introduction

In the yarn production process, malfunctions of spinning equipment components can significantly affect production efficiency and the quality of the resultant yarn. Therefore, the timely detection of abnormal conditions in spinning equipment holds paramount importance. The expeditious identification and resolution of equipment malfunctions are essential to guarantee the stable operation of spinning equipment and to elevate the consistency of yarn quality.

Existing anomaly detection techniques for production equipment predominantly incorporate vibration monitoring, temperature surveillance, oil analysis, sound detection, and image recognition [1–5]. In the specific context of spinning equipment, these conventional methods are complemented by data related to evenness quality, which serves as a crucial criterion for assessment. Throughout the continuous processing of fibrous materials, the quality detection system assesses various parameters, including weight deviation (A%), coefficient of variation (CV%) for different length segments, CV% histograms, weight deviation curves, and the wavelength spectrogram (WSG). The identification of anomalous values within this dataset empowers maintenance personnel to promptly discern equipment malfunctions, diagnose faults, and implement necessary maintenance measures.
Within the yarn quality data, the WSG encompasses two-dimensional data obtained from spectral analysis of yarn thickness uniformity, as captured by a sensor situated at the device’s output end. The current fault diagnosis system for drafting devices, which relies on the WSG, employs a knowledge base that contains methods for calculating and analyzing periodic unevenness. It diagnoses equipment faults by contrasting test results with statistical values or established thresholds. Consequently, the WSG serves as a critical instrument for evaluating and diagnosing defective components or underlying causes in drafting equipment.

Early classical fault diagnosis methods, such as Fuzzy Petri Net (FPN), were extensively utilized across various engineering fields [6]. A fault diagnosis and cause analysis (FDCA) model, which integrates Fuzzy Evidential Reasoning (FER) and Dynamic Adaptive Fuzzy Petri Nets (DAFPNs), has been proposed [7]. Despite its potential as a fault diagnosis tool, FPN is subject to certain drawbacks, including limited versatility and vulnerability to compromised diagnostic outcomes in situations with incomplete information.

Furthermore, these detection approaches hinge on manual expertise and are encumbered by limitations: the intricate nature of the spinning process introduces myriad potential factors contributing to periodic unevenness, many of which may not be comprehensively addressed by the analysis system. Additionally, the WSG does not manifest a consistent profile amidst fluctuating production conditions. Consequently, the existing diagnostic analysis system lacks the adaptability, versatility, and intelligence requisite for optimal performance.

The utilization of machine learning anomaly detection techniques offers a potential solution to the aforementioned challenges. Machine learning algorithms demonstrate the capacity to analyze sensor data and, with appropriate training, efficiently undertake equipment anomaly detection tasks [6]. Nevertheless, in manufacturing scene, the occurrence of abnormal events is relatively infrequent. Additionally, the gathering and labeling of anomalous data present notable challenges, and acquiring labeled datasets demands considerable time and specialized expertise, leading to elevated costs. Consequently, the application of supervised learning methods may not be feasible in such scenarios [7].

Unsupervised learning methods employ unlabeled data for training, explore the internal structure and patterns within data, offering an effective solution to these problems [8]. Such methods eliminate the need for labeling datasets, thereby reducing the cost of data annotation. Semi-supervised learning, positioned between supervised and unsupervised learning, utilizes a small amount of labeled training data and a large amount of unlabeled training data. A concept within semi-supervised anomaly detection involves training and developing anomaly detectors using normal data, utilizing only normal datasets for training to learn the characteristics of normal data [9]. By establishing a specific decision threshold, the model can discern whether the test data are abnormal.

This study presents a method for detecting anomalies in WSG data utilizing a convolutional autoencoder (CAE) with a semi-supervised training process. The CAE consists of two primary components: the encoder, and the decoder. These components are concurrently trained to compress the original data (encoder) and then reconstruct the input data (decoder). Following training, the reconstruction error serves as an indicator for identifying abnormal behavior. It effectively navigates the challenge of acquiring and annotating anomalous data in real-world contexts. By capitalizing on the reconstructive data feature of CAE, anomalies within WSGs can be identified, exploiting the characteristic that the reconstruction error of anomalous data surpasses that of normal data.

This paper conducts a comparative analysis of the efficacy of conventional anomaly detection methods in the context of industrial equipment anomaly detection based on two real engineering data experiments. Furthermore, it discusses the capacity of semi-supervised methods in detecting anomalies in WSG data.

The primary contributions of this paper can be summarized as follows:

1. Introduction of a semi-supervised anomaly detection method for WSG data utilizing a CAE. In contrast to traditional WSG anomaly detection techniques, this approach
eliminates the need for extensive manual expertise and is not constrained by data labeling requirements.
2. The method incorporates Kernel Density Estimation (KDE) in conjunction with grid search to establish a threshold value for identifying anomalies based on reconstruction errors. The adaptive nature of the threshold selection ensures effective anomaly detection across diverse datasets under varying process conditions.
3. Comparative analysis reveals that, when compared to commonly utilized machine learning techniques, the proposed anomaly detection method exhibits superior performance in accurately identifying abnormal data within an unsupervised setting.

This paper is organized as follows: The second Section provides an overview of related research. Following that, the third Section elucidates the principles of CAE and outlines the methodologies for threshold determination. The efficacy of the proposed anomaly detection method is substantiated using industrial datasets in the fourth Section. Finally, the fifth Section provides a summary of the work conducted in this study and discusses future avenues for research.

2. Related Works

2.1. Fault Diagnosis by Wavelet Spectrogram Analysis (WSA)

Due to imperfections in the spinning machinery or process, the yarn produced may exhibit periodic variations in quality during spinning. Consequently, the final fabric product is susceptible to notable quality degradation. The identification of these periodic defects can be efficiently achieved through spectral analysis.

The spectrum effectively portrays the diverse periodic uneven wavelengths. The abscissa of the WSG represents the logarithm of the wavelength, while the ordinate denotes the relative amplitude value, as depicted in Figure 1. Mechanical defects within the spinning equipment give rise to periodic uneven mechanical waves. Meanwhile, improper settings of process parameters lead to hill-like drafting waves, indicated by the red shaded area in Figure 1. Analyzing the shape of the spectrum allows to discern the characteristics of the yarn’s uneven defects and calculate the mechanical components and factors causing these defects.

![Figure 1. Example of anomalous WSG.](image)

The advantage of spectrum detection lies in its ability to estimate the location of equipment failure through analysis of the amplitude, waveform, and position of the spectrum. The common method for spectrum analysis is calculation, as demonstrated in Equation (1). In this equation, $\lambda_1$ represents the wavelength of the mechanical wave, $D_1$ denotes the diameter of the rotating part generating the mechanical wave, and $E_1$ signifies the draft ratio between the rotating part and the output part of the product.

$$\lambda_1 = \pi \cdot D_1 \cdot E_1$$

(1)
The mechanical wave, often due to spinning machine failure, can lead to periodic unevenness in the yarn, which is reflected in the 'chimney' shape on the spectrum. For instance, Figure 2 illustrates the spectrum obtained from the evenness experiment of the FA506 spinning frame, showing a 'chimney' shape at 5 cm and 10 cm. The CV value exceeds the normal level by 0.9 percentage points. The diameter range of the front rubber roller in the spinning frame is between 2.8 cm and 3.0 cm, with a corresponding draft ratio of 1. Hence, $\lambda_1$ is defined as the product of the front rubber roller diameter and the draft multiple, ranging approximately from 8.8 cm to 9.4 cm, signifying the mechanical wave induced by the rubber roller.

$$\lambda_1 = \text{front rubber roller diameter} \times \text{draft ratio}$$

Figure 2. Example of WSG with mechanical fault.

The challenge associated with spectrum detection lies in the requirement for fault assessment, where one needs to utilize the wavelength and amplitude distribution of mechanical periodic waves in the spectrum to deduce uneven characteristics and subsequently pinpoint the origin of mechanical defects. However, this process demands spectrum calculation results, knowledge of equipment characteristics, and production experience. The integration of machine learning methods for anomaly detection in spectral data enables the identification of device anomalies without prior information and diminishes the dependence on manual experience.

2.2. Anomaly Detection Methods for Industrial Data

The industrial data anomaly detection process utilizing machine learning comprises several pivotal steps: data preprocessing, feature extraction, detection classification, and fault diagnosis. To enable the automated identification of faulty components, machine learning algorithms leverage data to formulate a detection model. The data employed for training this model essentially consist of features that have been constructed and extracted from the original dataset.

Sensor data features are generally categorized into time-domain and frequency-domain features [10]. Time-domain features encompass statistical measures such as the maximum value, minimum value, peak-to-peak value, mean value, variance, standard deviation, mean square error, kurtosis, etc. On the other hand, frequency-domain features include centroid frequency, average frequency, root mean square frequency, and frequency standard deviation [11]. It is important to note that the rules established in feature engineering may be specifically tailored to specific datasets. In the textile industry, the WSG serves as a distinctive frequency-domain feature.

Anomaly detection techniques for industrial sensor data are commonly categorized into statistical, clustering, classification, and nearest-neighbor methods [12]. The majority of machine learning algorithms require optimization through supervised training, which entails a substantial amount of labeled data for the training process. Nevertheless, in industrial production settings, preventing fault states is crucial, and obtaining sufficient and accurate labeled data for abnormal datasets can be particularly challenging. Moreover, numerous factors contribute to the periodic irregularities observed in the WSG dataset. The most frequent mechanical defects include eccentricity in rollers or rubber rollers,
concave deformation in rubber rollers, flaws in the draft transmission gear, or suboptimal meshing. Acquiring labels for abnormal WSG data demands extensive expert knowledge and a significant amount of time, presenting significant challenges for supervised anomaly detection techniques.

Unsupervised anomaly detection identifies anomalies based on the intrinsic characteristics of the data. The training dataset remains unlabeled, and the objective is to either cluster or differentiate the observed values. Predominantly employed unsupervised classification techniques are grounded in clustering algorithms and density-based approaches [13]. Clustering, an unsupervised learning technique, segregates data into distinct groups or clusters by computing the similarity or distance between samples. During the clustering process, outliers typically represent points that fail to be allocated to any cluster or lie at a considerable distance from all cluster centroids. K-means clustering stands as one of the most frequently utilized clustering algorithms. On the other hand, the density-based method detects anomalies by evaluating the local density of the data. For instance, the Local Outlier Factor (LOF) algorithm, a density-based unsupervised anomaly detection technique, pinpoints outliers by comparing the local density of data points with that of their neighbors.

Deep neural networks have been extensively applied in this domain [14–17], proving to be adaptable to various semi-supervised anomaly detection scenarios. Notable examples include the use of CAE for anomaly detection in concrete structures [18], deep autoencoders for water level monitoring [19], and LSTM autoencoders for sensor signal anomaly detection [20]. These studies offer valuable tools and methodologies for semi-supervised anomaly detection, contributing to improved accuracy and efficiency in the detection of anomalies in industrial equipment.

An autoencoder (AE) serves as a semi-supervised deep learning model utilized for anomaly detection based on the concept of reconstruction error [19]. The primary training goal is to minimize the reconstruction error on standard data, typically quantified using mean square error (MSE). Anomalies are then detected by setting a predefined threshold for the reconstruction error. In the context of anomaly detection in industrial time series data, anomalies represent a minority compared to normal data. When the dissimilarity between the input and the reconstructed output from the AE exceeds the predetermined threshold, the data are considered anomalous.

Therefore, after training the AE with normal data, the reconstruction error significantly increases for anomalous inputs, whereas it remains minimal for normal data inputs. The traditional AE comprises a fully connected layer with an extensive number of parameters, which in certain instances induces overfitting. When confronted with two-dimensional or high-dimensional data, it is typically requisite to convert the original data into one-dimensional vectors. This transformation disrupts the structural characteristics within the data and generates a multitude of superfluous parameters. Conversely, the CAE can circumvent the aforementioned issues. CAE replace these with convolutional and pooling layers, which are more effective in preserving the feature information of the input data. In the context of the WSG data, CAE demonstrate superior performance in feature extraction compared to conventional autoencoders. Consequently, this study aims to employ CAE for anomaly detection. The training dataset consists of WSG data acquired from the device under standard operating conditions. This approach straddles the boundary between supervised and unsupervised learning and can be categorized as a semi-supervised learning method.

3. Proposed Approach

3.1. Anomaly Detection Process

The process of WSG anomaly detection proposed in this paper is depicted in Figure 3. The WSG dataset is standardized, and the AE model is trained using the normal dataset to obtain optimal parameters. The reconstruction error is derived from this training process. Subsequently, the test dataset is input into the trained AE model for further evaluation. The
optimal threshold, denoted as $Z_{\text{threshold}}$, is determined using the grid search method. If the reconstruction error surpasses the optimal threshold, it is classified as abnormal; otherwise, it is considered normal.

![Figure 3. Anomaly detection framework utilizing CAE.](image)

3.2. CAE

AE consists of an encoder and a decoder. Each encoder comprises a feedforward neural network, which contains multiple hidden fully connected layer [21]. The original data is compressed by the encoder to generate the actual features [20], and the decoder reconstitutes the sample from the coding features as input. The AE function $f_w(\cdot)$ maps the input value $x$ to the latent space to generate the latent value $f_w(x)$. The latent value $f_w(x)$ is a nonlinear representation of the input data, which retains the underlying characteristics of the input data. The decoder uses the latent value $f_w(x)$ to reconstruct the output. Therefore, the latent value is a feature extraction of the input data [8,22]. The decoder function $g_a(\cdot)$ passes and outputs the potential value $f_w(x)$ to obtain the output value $\hat{x}$, as shown in Equation (2).

$$\hat{x} = g_a(f_w(x))$$ (2)

The primary objective of AE is to train the functions $f_w(\cdot)$ and $g_a(\cdot)$ to minimize the discrepancy between the original data $x$ and the reconstructed data $\hat{x}$. Figure 4 illustrates the architecture of the CAE. Unlike traditional AE, CAE, on the other hand, combines the semi-supervised learning approach of AE with the unique structural advantages of CNNs, the CAE replaces fully connected layers with convolutional and pooling layers while it adheres to the same underlying principle. This variant efficiently eliminates superfluous information and extracts salient features from the data. The features of WSG are abstract, which enables both automatic extraction of WSG features and effective retention of structural information in the data. The encoder effectively encapsulates the essential characteristics of the input data, while the decoder reconstitutes the input from the encoded features. The convolution operation is depicted in Equation (3). Here, $H^k$ denotes the k-th convolution surface, $\sigma$ symbolizes the activation function, $W^k$ and $b^k$ signify the k-th weight and bias, respectively, and * indicates the convolution operation. Equation (3) elucidates that the convolution operation encompasses the traits of local connectivity and weight sharing. This mitigates the parameter count in a conventional AE and shares the weight
across the local receptive field, consequently extracting the structural attributes of the original data proficiently and diminishing the overfitting risk.

\[ H^k = \sigma(x * W^k + b^k) \quad (3) \]

Figure 4. Structure of convolutional autoencoder.

3.3. Prediction Error Calculation

The determination of anomaly detection relies on the notable distinction between abnormal and normal data [23]. The reconstruction error resulting from the AE’s reconstruction of abnormal data, which has been trained using normal data, exhibits a substantial difference from that of normal data [24].

Reconstruction error, also referred to as prediction error, can be calculated by statistical estimation. Commonly used reconstruction errors include mean square error (MSE) [24], mean absolute error (MAE) [25], square prediction error (SPE) [26], and absolute error [27]. In the context of anomaly detection, AE typically utilize MSE as the reconstruction error [19,20,28]. MSE of the reconstruction error is shown in Equation (4), where \( X = [x_1, x_2, \ldots, x_N] \) is the input segment and \( \hat{X} = [\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_N] \) is its corresponding reconstruction output.

\[ \text{MSE} = \frac{1}{N} \sum_{i=1}^{n} (\hat{x}_i - x_i)^2 \quad (4) \]

3.4. Probability Density Estimation of Prediction Error

To determine the optimal classification threshold, this investigation employed KDE to model the error distributions of both abnormal and typical data. KDE is a nonparametric estimation technique designed to assess unknown distribution density functions. Given data points \( x_1, x_2, \ldots, x_n \), derived from a continuous probability density function \( f(x) \), with bandwidth \( h \) and kernel function \( K \), the kernel density estimate at any specified point is articulated by Equation (5).

\[ \hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left( \frac{x - x_i}{h} \right) \quad (5) \]

3.5. Anomaly Decision-Making

The final step in constructing a classification model involves determining an appropriate threshold, which directly impacts the model’s performance. This study employs the grid search method to identify the optimal threshold, denoted as \( Z_{\text{threshold}} \). As a hyperparameter optimization technique, grid search utilizes exhaustive optimization to pinpoint
the most suitable parameters within a predetermined set. Specifically, the range of $Z_{\text{threshold}}$ and the step size for the search are derived from the distribution characteristics of the kernel density estimation of the reconstructed samples. Subsequently, the grid search method is applied to obtain the value of $Z_{\text{threshold}}$ that maximizes the $F_1$ score in the test set, as referenced in [29]. The decision-making process is articulated in formula (6), where a value of ‘0’ denotes a normal data sample, while a value of ‘1’ indicates an abnormal data sample.

$$\text{Result} = \begin{cases} 0 , & \text{threshold} < Z_{\text{threshold}} \\ 1 , & \text{threshold} > Z_{\text{threshold}} \end{cases}$$

4. Experimental Procedure

To validate the effectiveness of the proposed method, this study utilizes two distinct industrial datasets for empirical analysis. Due to the lack of comparable research on spectral data detection, Case 1 employs the publicly available rolling bearing dataset to benchmark the method presented in this paper against current detection techniques. In Case 2, authentic WSG data are utilized to substantiate the detection accuracy.

4.1. Evaluation Metrics

In order to assess the performance of the methods, Accuracy, Recall, F1-score, and Area Under the Curve (AUC) were chosen as evaluation indicators for testing data [30]. The calculation methods are outlined below:

$$\text{TPR} = \frac{TP}{TP + FN}$$

(7)

$$\text{FPR} = \frac{FP}{FP + TN}$$

(8)

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

(9)

$$\text{Recall} = \frac{TP}{TP + FN}$$

(10)

$$F1 \ - \ score = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

(11)

Among these, the positive class and the negative class denote normal data and abnormal data in the detection process, respectively. The definitions for True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN) are as follows:

True Positive (TP): The actual category of the sample is positive, and the model correctly recognizes it as positive.

False Negative (FN): The actual category of the sample is positive, but the model incorrectly identifies it as a negative class.

False Positive (FP): The actual category of the sample is negative, but the model incorrectly identifies it as a positive class.

True Negative (TN): The actual category of the sample is negative, and the model correctly identifies it as negative.

4.2. Case 1

4.2.1. Bearing Dataset

The dataset utilized in this experiment is sourced from the study referenced as [31], containing vibration signal data. It includes vibration signals recorded by two PCB352C33 unidirectional accelerometers with a frequency setting of 25.6 kHz, an interval of 1 min, and a duration of 1.28 s each time.

Figure 5a presents the time-domain signal of the experimental data, where the horizontal axis represents time and the vertical axis shows the amplitude of the bearing vibration.
signal. Prior to the $0.851 \times 106$th sample, the data are marked as normal, while after the change point, the rolling bearing experiences inner ring wear, leading to increased vibration. Although time-domain signals can reveal faults through temporal trends, early faults are challenging to detect. It is widely acknowledged that frequency-domain signals are more sensitive to faults than time-domain signals [32], as depicted in Figure 5b,c, which represent the frequency-domain plots under faulty and normal conditions, respectively. In Figure 5b, a significant amplitude of periodic vibration is observed at a frequency of 1000 Hz, which is absent in Figure 5c, indicating a frequency characteristic of the fault occurrence. Therefore, this experiment uses frequency-domain values obtained from the fast Fourier transform to extract data features for anomaly detection.

![Figure 5. Vibration signal generated by bearing with its frequency spectrum. (a) Time domain signal. (b) Frequency spectrum of abnormal signal. (c) Frequency spectrum of normal signal.](image)

For this research, the Bearing2_1 dataset has been chosen, encompassing 491 samples, comprising 461 normal samples and 30 fault samples, with the inner ring identified as the bearing failure component. The training set constitutes 65% of the data, while the test set represents 35%.

### 4.2.2. Experimental Setup

Before the model training, a linear transformation is applied to the original data using the min-max normalization method, mapping the results to the interval (0,1). The normalized data are represented by $x$, where $x_{\text{new}}$ denotes the normalized data, $x_{\text{min}}$ and $x_{\text{max}}$ represent the minimum and maximum values in the current data. The calculation method is as follows:

$$x_{\text{new}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$$  \hspace{1cm} (12)
In the CAE model, the encoder comprises an input layer, followed by three one-dimensional convolutional layers, three normalization layers, and an activation layer. The number of convolution kernels is set to 1, with a convolution step size of 1, and the activation function employed is ReLU. The decoder is constructed from three one-dimensional deconvolution layers, three normalization layers, and activation functions. MSE serves as the loss function during training, with Adam being the optimization algorithm. The learning rate is fixed at 0.001, the batch size is 16, and the training process is executed over 50 epochs.

Figure 6 illustrates the KDE curves for the reconstruction samples. The x-axis represents the reconstruction error, while the y-axis indicates the density. It delineates the distribution of data density for both normal and abnormal data of the bearings, with the blue section illustrating the distribution of normal data, whereas the orange section depicts the distribution of anomalous data. KDE curve facilitates the determination of a threshold range, enabling the identification of the optimal threshold within this range. As indicated by the x-axis in Figure 6, the maximum reconstruction errors for AE and CAE are 0.15 and 0.5, respectively. Consequently, the threshold search intervals for AE and CAE are designated as (0, 0.15) and (0, 0.5).

![Figure 6. KDE curve for AE and CAE.](image)

Subsequently, the grid search method was employed to determine the value of \( Z \) that maximizes the performance index F1-score and accuracy in the test set, as illustrated in Figure 7a,b. The optimal thresholds were determined to be 0.006 for the AE model and 0.004 for the CAE model.

![Figure 7. Grid search results.](image)
4.2.3. Experimental Results

The experiments compare the classification efficacy of the proposed method with commonly used anomaly detection algorithms. These include supervised algorithms such as Random Forest, K-Nearest Neighbors (KNN), Logistic Regression, and Support Vector Machines (SVM), as well as unsupervised algorithms like Isolation Forest and AE. These algorithms have been utilized by researchers for anomaly detection in various applications [33]. The model employs the default parameters set forth in the Scikit-learn library.

Figure 8 demonstrates the ability of CAE and AE to distinguish between normal and anomalous data. In the figure, dots represent the reconstruction error of normal data, while crosses signify anomalous data. The blue line denotes the optimal threshold, used to separate the data in the test set. It is evident that after processing with the autoencoder method, the reconstruction error values of normal data are lower and more concentrated, while the reconstruction error of anomalous data is higher and dispersed. Moreover, compared to AE, the reconstruction error obtained from CAE can more accurately differentiate between normal and anomalous data.

![Visualization of anomaly detection for AE](image1)

![Visualization of anomaly detection for CAE](image2)

**Figure 8.** Visualization of anomaly detection.
The experimental results are summarized in Table 1 and depicted in Figure 9. Figure 9 shows a visual comparison of the results on the four metrics. With the vertical axis representing the metric values, and the distribution of the histograms on the horizontal axis divided into four groups, each of which displays four metrics, namely, Accuracy, Recall, AUC, and F1-score, for each of the seven methods. Within the supervised algorithm category, SVM stands out with an accuracy rate of 0.992, a recall rate of 0.878, an AUC of 0.938, and an F1-score of 0.934, surpassing other supervised algorithms. The unsupervised algorithm Isolation Forest exhibits lower performance across all four indicators compared to supervised methods.

Table 1. Bearing data experiment results.

<table>
<thead>
<tr>
<th>Models</th>
<th>Accuracy</th>
<th>Recall</th>
<th>AUC</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest</td>
<td>0.980</td>
<td>0.863</td>
<td>0.925</td>
<td>0.832</td>
</tr>
<tr>
<td>KNN</td>
<td>0.984</td>
<td>0.777</td>
<td>0.888</td>
<td>0.872</td>
</tr>
<tr>
<td>Logistic</td>
<td>0.987</td>
<td>0.808</td>
<td>0.903</td>
<td>0.892</td>
</tr>
<tr>
<td>SVM</td>
<td>0.992</td>
<td>0.878</td>
<td>0.938</td>
<td>0.934</td>
</tr>
<tr>
<td>Isolation Forest</td>
<td>0.932</td>
<td>0.623</td>
<td>0.811</td>
<td>0.766</td>
</tr>
<tr>
<td>CAE</td>
<td>0.989</td>
<td>1.000</td>
<td>0.990</td>
<td>0.968</td>
</tr>
<tr>
<td>AE</td>
<td>0.973</td>
<td>0.900</td>
<td>0.943</td>
<td>0.915</td>
</tr>
</tbody>
</table>

Figure 9. Bearing experiment comparison of each index.

Regarding semi-supervised methods, CAE achieves higher accuracy than AE and is marginally inferior to SVM. With a Recall of 1, an AUC of 0.99, and an F1-score of 0.968, it attains the highest scores among all methods. In conclusion, the CAE-based anomaly detection method exhibits commendable performance in detecting anomalies within bearing datasets, outperforming other methods.

4.3. Case 2

4.3.1. WSG Dataset

WSG data are collected from a single axis detection spinning machine used in the test phase, and the amplitude spectrum corresponding to 140 wavelength segments is obtained by filtering the sensor current signal through a multi-channel bandwidth filter. The dataset comprises 323 spectral samples, with 303 being normal and 20 indicating fault conditions. The training set is composed of 200 normal samples, while the test set includes the remaining 103 normal samples and the 20 fault samples.
4.3.2. Experimental Setup

According to the reconstructed data from the test set, KDE curves derived from AE and CAE models are illustrated in Figure 10. The curves in Figure 10 are KDE curves, and the area enclosed by the curves and the horizontal coordinates indicates the distribution of normal and abnormal data, blue sections indicate normal data, and orange sections indicate abnormalities. As demonstrated on the abscissa of Figure 10, the maximum reconstruction errors for AE and CAE are approximately 0.08 and 0.12, respectively. Consequently, the threshold search intervals for AE and CAE are defined as (0, 0.08) and (0, 0.12), respectively. Utilizing the grid search method, the value of Z that maximizes the F1-score of the performance index within the test set is ascertained. As demonstrated in Figure 11, the optimal thresholds for AE and CAE models are determined to be 0.041 and 0.078, respectively, with the corresponding maximum F1-scores being 0.927 for AE and 0.955 for CAE.

![KDE curve](image1.png)

**Figure 10.** KDE curve.

![Grid search results](image2.png)

**Figure 11.** Grid search results.

4.3.3. Experimental Results

The visual representation of AE and CAE detection results are shown in Figure 12. For WSG data, CAE still shows a high degree of differentiation between normal and abnormal data. The experimental outcomes are summarized in Table 2 and illustrated in Figure 13. The AUC of the CAE model is 0.99, which surpasses the 0.94 AUC of the AE model. Additionally, the CAE model exhibits superior performance across other evaluation metrics. This indicates that the CAE model possesses a stronger anomaly detection capability compared to other models when analyzing WSG data anomalies. The recall of the CAE model reaches 1, while the AE model's recall is 0.9, suggesting a slight superiority in detecting WSG data abnormalities. As demonstrated on the abscissa of Figure 10, the maximum reconstruction errors for AE and CAE are approximately 0.08 and 0.12, respectively. Consequently, the threshold search intervals for AE and CAE are defined as (0, 0.08) and (0, 0.12), respectively. Utilizing the grid search method, the value of Z that maximizes the F1-score of the performance index within the test set is ascertained. As demonstrated in Figure 11, the optimal thresholds for AE and CAE models are determined to be 0.041 and 0.078, respectively, with the corresponding maximum F1-scores being 0.927 for AE and 0.955 for CAE.
detection capability compared to other models when analyzing WSG data anomalies. The recall of the CAE model reaches 1, while the AE model’s recall is 0.9, suggesting a slight superiority of the CAE model. Figure 12 illustrates that the CAE model outperforms the AE model in differentiating abnormal data and can accurately identify all WSG anomalies. Among traditional machine learning algorithms, SVM achieves the highest performance with a recall of 0.88 and an AUC of 0.94, yet it still falls short of the CAE model. This demonstrates that conventional machine learning algorithms have a lower efficacy in detecting WSG data anomalies compared to deep learning algorithms. Consequently, the CAE model holds significant advantages over alternative algorithms for the task of WSG data anomaly detection.

Figure 12. Visualization of anomaly detection.
Table 2. WSG data experiment results.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Recall</th>
<th>AUC</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAE</td>
<td>0.984</td>
<td>1.000</td>
<td>0.990</td>
<td>0.955</td>
</tr>
<tr>
<td>AE</td>
<td>0.970</td>
<td>0.900</td>
<td>0.940</td>
<td>0.927</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.947</td>
<td>0.764</td>
<td>0.864</td>
<td>0.880</td>
</tr>
<tr>
<td>KNN</td>
<td>0.896</td>
<td>0.520</td>
<td>0.760</td>
<td>0.684</td>
</tr>
<tr>
<td>Logistic</td>
<td>0.931</td>
<td>0.680</td>
<td>0.840</td>
<td>0.800</td>
</tr>
<tr>
<td>SVM</td>
<td>0.974</td>
<td>0.880</td>
<td>0.940</td>
<td>0.936</td>
</tr>
<tr>
<td>Isolation Forest</td>
<td>0.784</td>
<td>0.240</td>
<td>0.587</td>
<td>0.324</td>
</tr>
</tbody>
</table>

Figure 13. WSG experiment comparison of each index.

5. Conclusions

In this study, we present a semi-supervised anomaly detection approach customized for yarn quality data within the spinning industry. The primary objective is to overcome challenges arising from the limited availability of labeled data in industrial settings and to diminish the reliance on manual fault detection for spinning equipment components. By leveraging normal data for model training, our method tackles the problem of insufficient abnormal data in practical scenarios. We have implemented this technique on two distinct industrial datasets and compared its performance against conventional anomaly detection methodologies. The ensuing key findings are summarized as follows:

1. The semi-supervised anomaly detection method, grounded in CAE, outperforms conventional anomaly detection techniques. This model exhibits commendable performance across various evaluation metrics, including Accuracy, Recall, AUC, and F1-score.
2. The proposed anomaly detection approach adeptly distinguishes between anomalous samples and normal samples within WSG data. However, due to the relatively intricate nature of anomaly detection applications, further analysis and validation will be conducted in subsequent work, particularly regarding scenarios involving the simultaneous occurrence of multiple faults or the linear superposition of normal data resulting in anomalous data.
3. Owing to the fact that WSG data imply fault types, this method demonstrates superior interpretability compared to other machine learning techniques. In subsequent research, we shall delve into the machine learning fault diagnosis method predicated on WSG, building upon the detection outcomes of this study.
Considering the imperative of ensuring consistency and performance in real-world industrial applications through the utilization of CAE for anomaly detection, several factors are pivotal. Firstly, data quality and preprocessing are paramount, necessitating clean, representative data prepared via techniques like normalization and feature engineering. Secondly, continuous monitoring and adaptation of the model, supported by appropriate evaluation metrics like precision, recall, and F1-score, ensure effectiveness as industrial environments evolve. Additionally, integration of domain expertise, inherent in WSG data, can further refines anomaly detection processes. These aspects will receive increased attention in subsequent work.

**Author Contributions:** Conceptualization, H.W. and C.S.; methodology, H.W.; software, H.W.; validation, H.W. and C.S.; formal analysis, Z.H.; investigation, H.W.; resources, C.S.; data curation, H.W.; writing—original draft preparation, H.W. and C.S.; writing—review and editing, C.S.; visualization, X.X.; supervision, X.S.; project administration, C.S.; funding acquisition, C.S. All authors have read and agreed to the published version of the manuscript.

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

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

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

**Nomenclature**

- **AE** Autoencoder
- **AUC** Area Under the Curve
- **CAE** Convolutional Autoencoder
- **CV%** Coefficient of Variation
- **KDE** Kernel Density Estimation
- **MSE** Mean Square Error
- **WSA** Wavelet Spectrogram Analysis
- **WSG** Wavelength Spectrogram

**References**


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