GEORGE BROWN PRESS
JOURNALS, BOOKS & PROCEEDINGS

Journal of Computer, Signal, and System Research

Article

Fault Diagnosis of Wind Turbine Bearings Using SwinT-CBAM-BiGRU

Fei Li [1,*] and Xueming Zhai [1]

[1] College of Control and Computer Engineering, North China Electric Power University, Baoding, Hebei, China
* Correspondence: Fei Li, College of Control and Computer Engineering, North China Electric Power University, Baoding, Hebei, China

Abstract: To address the challenge that traditional diagnostic models struggle to simultaneously capture local impacts and global features due to background noise interference under the complex operating conditions of wind turbine bearings, this study proposes a novel multimodal fusion fault diagnosis model based on SwinT-CBAM-BiGRU. Specifically, the Gramian Angular Difference Field (GADF) encoding technique is employed to transform 1D vibration signals into 2D feature images. A Swin Transformer integrated with the Convolutional Block Attention Module (CBAM) is utilized to extract deep spatial features and precisely pinpoint core fault regions. Concurrently, a Bidirectional Gated Recurrent Unit (BiGRU) is combined to mine the long-range temporal evolution patterns of the signals. Through comparative and ablation experiments conducted on the Case Western Reserve University (CWRU) bearing dataset, the results demonstrate that the proposed fusion model exhibits superior diagnostic capabilities. Ultimately, this method effectively breaks the constraints of single-dimensional features, demonstrating stronger discriminative stability and robustness in multi-class bearing fault diagnosis tasks.

Keywords: wind turbine; bearing fault diagnosis; multimodal fusion

1. Introduction

During their lifecycle, wind turbines often suffer from major component failures, leading to substantial quality costs. Statistical data indicate that among various mechanical faults in wind turbine drivetrain systems, downtime incidents caused by early bearing failures account for the highest proportion. Consequently, bearings have become a critical bottleneck restricting the reliable operation of wind turbines [1]. Unlike general-purpose industrial bearings, the loads and rotational speeds of wind turbine bearings fluctuate significantly due to dynamic wind flows. Under certain extreme operating conditions, these large bearings endure excessive loads and are prone to damage [2]. Therefore, developing reliable and cost-effective fault diagnosis methods for wind turbine bearings is of paramount importance.

The primary mission of bearing fault identification is to precisely determine the health status of mechanical components by interpreting gathered operational datasets. Over the past decades, a variety of sensing technologies have been engineered to harvest fault-related information. Methods such as acoustic emission, wear debris monitoring, power quality checks, and thermal analysis have all been validated as effective tools for supervising large-scale wind turbine bearings [3]. Notably, vibration-based monitoring has emerged as the dominant and most practical strategy. This preference stems from its extensive frequency coverage (typically spanning 1–30 kHz) and the simplicity of its industrial deployment [4].

Diagnostic strategies centered on vibration signals are broadly bifurcated into two paradigms: conventional signal processing and modern deep learning architectures [5]. The former typically follows a two-stage workflow involving feature engineering and

Received: 31 January 2026
Revised: 21 March 2026
Accepted: 02 April 2026
Published: 05 April 2026

subsequent pattern recognition [6]. Within this framework, signal processing covers time, frequency, and joint time-frequency domains. Time-domain indicators focus on statistical metrics like kurtosis and peak factors, while frequency-domain studies utilize power and envelope spectra to detect anomalies [7]. Furthermore, advanced time-frequency tools, including STFT [8], WT [9], and EMD [10], are widely employed. In terms of classification, statistical learning models such as SVM and Random Forests (RF) are frequent choices [11]. For instance, Meng et al. successfully separated rolling bearing features using EMD combined with kurtosis-based filtering [10]. In a similar vein, an EEMD-SVM hybrid scheme was introduced by Hu et al., utilizing energy ratios and approximate entropy for feature refinement [12].

Despite their utility, conventional diagnostic schemes are often constrained by a heavy reliance on human-crafted features and specialized domain expertise [13]. This dependence introduces subjective biases that may undermine the final diagnostic precision. To mitigate these drawbacks, data-driven deep learning models have transitioned into a phase of rapid advancement [14]. These architectures excel at autonomous feature representation, facilitating highly accurate fault detection. As an illustration, Wang et al. pioneered a multi-scale CNN architecture to integrate 1D and 2D signal correlations, significantly enhancing classification accuracy [15]. Additionally, the MDFT-GAN framework was proposed by Guo et al. to resolve data scarcity and class imbalance in wind farms [16]. Keshun et al. also contributed by merging Quadratic Neural Networks with Bi-LSTM, creating a high-speed hybrid model for rolling bearing analysis [17].

Current academic consensus suggests that isolated deep learning architectures often struggle to process intricate, long-range temporal sequences. Consequently, a major bottleneck in intelligent diagnosis is the simultaneous suppression of background noise and the capture of global topological structures. To fill this research gap, this paper develops Swin-CBAM-BiGRU, an innovative diagnostic framework rooted in multimodal feature fusion. Through rigorous validation on the Case Western Reserve University (CWRU) benchmark, the proposed architecture exhibits superior robustness and precision, proving its effectiveness in identifying multi-class faults amidst complex noise interference.

2. Bearing Fault Diagnosis Based on Swin Transformer and BiGRU

2.1. Overall Framework of the Proposed Model

The overall framework of the proposed Swin-CBAM-BiGRU fault diagnosis model, based on multimodal fusion, primarily consists of a dual-branch feature extraction network. In the spatial feature extraction branch, the original 1D vibration signals are first transformed into 2D feature images using the Gramian Angular Difference Field (GADF) technique. These images are subsequently fed into the Swin Transformer network to extract deep spatial textures. Furthermore, the Convolutional Block Attention Module (CBAM) is integrated to precisely localize the core regions associated with fault patterns. Meanwhile, the temporal feature extraction branch focuses on the introduction of BiGRU. By reorganizing the 1D long sequences into a matrix structure as input, the network overcomes the limitations of a single time axis. This allows the model to fully leverage both past and future contextual information, thereby deeply mining and extracting the instantaneous fluctuations and temporal evolutionary patterns of the signals.

In the feature aggregation and classification stage, the model performs a linear concatenation of the spatial feature vectors from the SwinT-CBAM branch and the temporal feature vectors from the BiGRU branch. This process constructs a joint feature representation that integrates multi-dimensional information across the spatial, temporal, and frequency domains. Subsequently, the fused high-dimensional features are fed into a fully connected (FC) layer. By employing the Softmax function to calculate the probability distribution of each sample, the model achieves efficient and accurate classification of various operating conditions for wind turbine bearings (Figure 1).

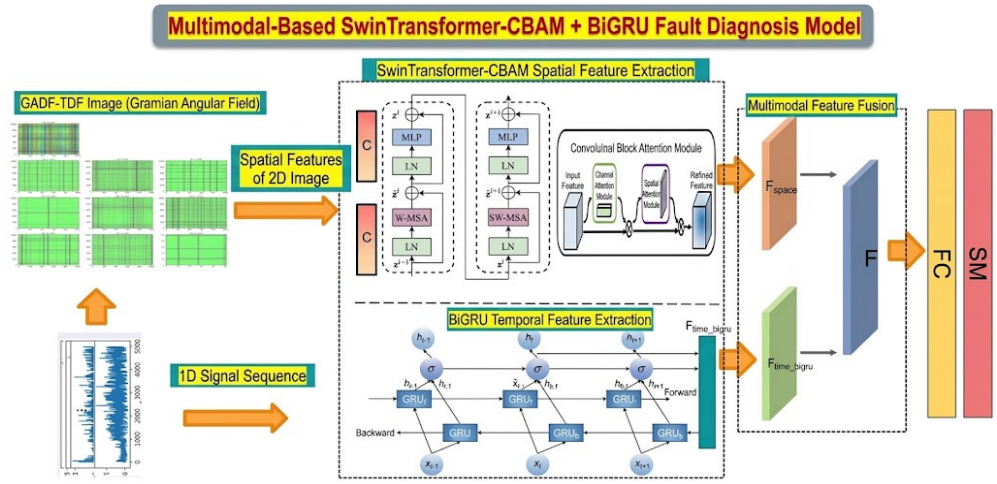

Figure 1. Schematic diagram of the overall framework.

2.2. Image Encoding of Vibration Signals Based on Gramian Angular Field

To convert 1D time series into 2D images, this study employs the Gramian Angular Field (GAF) encoding technique. By applying a polar coordinate transformation, GAF establishes interdependencies among different time points while preserving the original temporal characteristics. Consequently, this method effectively captures the temporal dynamics of the signals. The conversion process consists of the following steps:

Data Normalization: The raw 1D time series data is normalized and scaled to the interval [0, 1]. Given a raw input time series $X = \{x_1, x_2, \ldots, x_n\}$, where n represents the sequence length, the minimum and maximum values of the sequence are calculated as $X_{\min} = \min(X)$ and $X_{\max} = \max(X)$, respectively. Each data point is then linearly normalized as follows:

$$\tilde{x}_i = \frac{x_i - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

The normalized data preserves the relative temporal dynamics of the original signal while eliminating the impact of varying data scales.

Polar Coordinate Transformation: The 1D time series is mapped into a polar coordinate system, encoding the temporal information into geometric angles. By treating the scaled value $\tilde{x}_i$ as the cosine of an angle and mapping the timestamp t_i to the radius r_i, the 1D signal is converted into a polar representation:

$$\begin{cases} \phi_i = \arccos(\tilde{x}_i), & 0 \le \tilde{x}_i \le 1 \\ r_i = \frac{t_i}{N}, & t_i \in \mathbb{N} \end{cases} \tag{2}$$

where N is a constant factor regulating the span of the polar radius. The advantage of this transformation is that it generates distinct trajectories in the polar coordinate system over time. Moreover, the monotonic property of φ within the interval $[0,\pi]$ ensures a bijective mapping (i.e., uniqueness of the mapping).

Gramian Matrix Generation: An $n \times n$ Gramian matrix is constructed by computing the angular relationships between any two time steps, i and j, in the polar space. Depending on the method of angular operation, GAF is categorized into the Gramian Angular Summation Field (GASF) and the Gramian Angular Difference Field (GADF).

GASF defines its matrix elements by calculating the cosine of the sum of angles at two different time steps:

$$GASF_{i,j} = \cos(\phi_i + \phi_j) = \tilde{x}_i \cdot \tilde{x}_j - \sqrt{1 - \tilde{x}_i^2}\sqrt{1 - \tilde{x}_j^2} \tag{3}$$

GASF tends to capture the overall envelope characteristics of the signal, resulting in visually smoother generated images.

GADF constructs the matrix by calculating the sine of the angular difference. The structure of its generated matrix G is formulated as follows:

$$G = \begin{bmatrix} \sin(\phi_1 - \phi_1) & \sin(\phi_1 - \phi_2) & \cdots & \sin(\phi_1 - \phi_n) \\ \sin(\phi_2 - \phi_1) & \sin(\phi_2 - \phi_2) & \cdots & \sin(\phi_2 - \phi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \sin(\phi_n - \phi_1) & \sin(\phi_n - \phi_2) & \cdots & \sin(\phi_n - \phi_n) \end{bmatrix} \quad (4)$$

Expanded via trigonometric identities, each element $G_{i,j}$ in the matrix can be further expressed as:

$$G_{i,j} = \sin(\phi_i - \phi_j) = \sqrt{1 - \tilde{x}_i^2} \cdot \tilde{x}_j - \tilde{x}_i \cdot \sqrt{1 - \tilde{x}_j^2} \quad (5)$$

Due to the differencing operation, GADF is more sensitive to instantaneous fluctuations and phase shifts in the signal. Its main diagonal (i.e., when $i = j$) is identically zero, visually presenting a finer and more complex texture. Therefore, GADF is ultimately selected as the input for the image branch in the proposed model design (Figure 2).

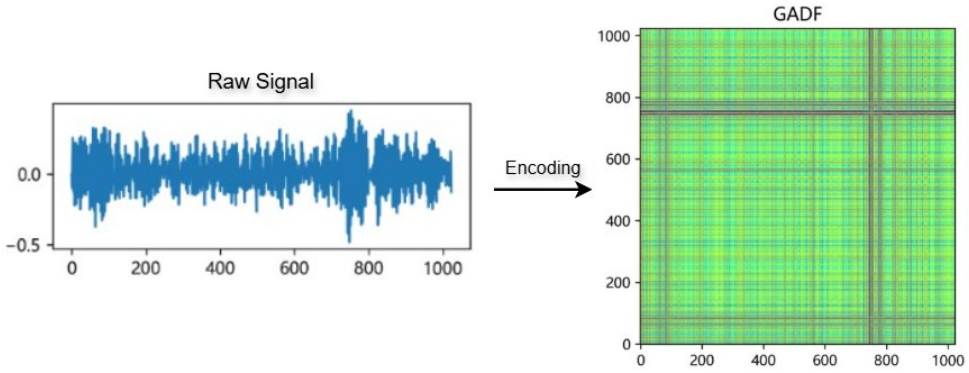

Figure 2. Image encoding of vibration signals.

2.3. Spatial Feature Extraction Based on Swin Transformer

The spatial feature extraction branch of this study is centered on the Swin Transformer, aiming to deeply excavate the complex frequency-domain textures and phase correlations embedded in the generated GADF images. To address the high-dimensional characteristics of time-frequency images of wind turbine bearings, the Swin Transformer architecture resolves the computational bottleneck of traditional Transformers, where the computational cost grows quadratically with the image size. The core of this model lies in constructing a Window-based Multi-head Self-Attention (W-MSA) mechanism. By partitioning the image into $M \times M$ non-overlapping windows, the computational complexity Ω achieves a linear relationship with the image size hw : $\Omega = 4hwC^2 + 2M^2hwC$. Furthermore, the model introduces a relative position bias $B \in \mathbb{R}^{M^2 \times M^2}$ to enhance the spatial perception capability for periodic impact features in the time-frequency images, formulated as follows: where Q , K , and V represent the query, key, and value matrices, and d is the feature dimension,

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}} + B\right)V \quad (6)$$

To compensate for the lack of communication between isolated local windows, the model employs the Shifted Window technique to alternate window shifts across consecutive blocks. Coupled with the Patch Merging layer for hierarchical downsampling, this enables the model to concurrently capture high-frequency detailed impacts and the overall trend of fault evolution. In two consecutive blocks, the evolution of the feature vector z^l follows a residual structure comprising W-MSA/SW-MSA and MLP:

$$\begin{cases} \hat{z}^l = \text{W-MSA}\big(\text{LN}(z^{l-1})\big) + z^{l-1} \\ z^l = \text{MLP}\big(\text{LN}(\hat{z}^l)\big) + \hat{z}^l \\ \hat{z}^{l+1} = \text{SW-MSA}\big(\text{LN}(z^l)\big) + z^l \\ z^{l+1} = \text{MLP}\big(\text{LN}(\hat{z}^{l+1})\big) + \hat{z}^{l+1} \end{cases} \quad (7)$$

where LN denotes Layer Normalization. Through this cyclic alternation, the model achieves cross-window global modeling without increasing the computational burden.

Following the Swin Transformer, the Convolutional Block Attention Module (CBAM) is embedded. CBAM refines the feature matrix F output by the Swin Transformer by integrating channel attention and spatial attention. The channel attention mechanism employs global max pooling and average pooling to automatically filter and emphasize the feature channels that contribute most significantly to fault category discrimination. Its calculation is formulated as Equation (8):

$$M_c(F) = \sigma\left(\mathrm{MLP}\left(\mathrm{AvgPool}(F)\right) + \mathrm{MLP}\left(\mathrm{MaxPool}(F)\right)\right) \tag{8}$$

The spatial attention mechanism pinpoints the regions with the most pronounced fault impacts at the pixel level.

$$M_s(F') = \sigma\left(f^{7\times7}\left([\mathrm{AvgPool}(F'); \mathrm{MaxPool}(F')]\right)\right) \tag{9}$$

where σ denotes the Sigmoid activation function, and $f^{7\times7}$ represents a convolution operation with a filter size of 7×7 .After multi-level hierarchical extraction and attention weighting, the spatial feature map is ultimately converted into a 1D spatial feature vector via global average pooling.

2.4. Temporal Feature Extraction Based on BiGRU

When processing bearing vibration signals, single-dimensional spatial features are insufficient to describe the dynamic trend of fault evolution over time. This branch employs the Bidirectional Gated Recurrent Unit (BiGRU) as the core of temporal feature extraction, which consists of two Gated Recurrent Units (GRUs) operating in opposite directions. Compared to traditional RNNs, GRU can effectively mitigate the vanishing gradient problem during long sequence training. Furthermore, it possesses fewer parameters, making it more suitable for the real-time inference requirements of wind power monitoring equipment. Compared to a unidirectional GRU, the advantage of BiGRU lies in its ability to break the constraints of the time axis, simultaneously utilizing past and future contextual information to enhance the representation of the current state.

The basic unit of BiGRU is composed of a forward hidden layer $\vec{h}_t$ and a backward hidden layer $\overleftarrow{h}_t$. Its internal gating mechanism operates collaboratively through an update gate z_t and a reset gate r_t , with the specific mathematical expressions formulated as follows:

$$\begin{aligned} z_t &= \sigma(W_z \cdot [h_{t-1}, x_t]) \\ r_t &= \sigma(W_r \cdot [h_{t-1}, x_t]) \\ \tilde{h}_t &= tan\,h(W \cdot [r_t \odot h_{t-1}, x_t]) \\ h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \end{aligned} \tag{10}$$

where σ denotes the Sigmoid activation function, and $\odot$ represents element-wise multiplication. At each time step, x_t is simultaneously fed into the forward and backward GRU cells. Finally, the hidden states from both directions are concatenated to obtain the complete feature representation at the current time step: $H_t = [\vec{h}_t; \overleftarrow{h}_t]$.

2.5. Model Feature Fusion and Classification

During the feature aggregation stage, the spatial branch outputs a 100-dimensional dense spatial feature vector, denoted as $S_{features}$. Simultaneously, the temporal branch generates a 128-dimensional high-order temporal feature vector, $T_{features}$ (formed by concatenating bidirectional 64-dimensional hidden states). Because these two vectors characterize the bearing conditions from two distinct dimensions---frequency-domain spatial textures and temporal dynamic evolution---they exhibit significant heterogeneous complementarity. The model constructs a 228-dimensional joint feature representation, F_{fusion} , via linear concatenation along the feature dimension:

$$F_{fusion} = \left[S_{features}, T_{features}\right] \tag{11}$$

The fused 228-dimensional high-level feature vector is subsequently fed into a global classification layer. Composed of a linear fully connected network, this layer is responsible for mapping the fused features into the predefined fault category space. Ultimately, the model employs the Softmax activation function to normalize the output of the

classification layer, computing the probability distribution that indicates the likelihood of the sample belonging to each respective fault category:

$$\hat{y} = \text{Softmax}\left(W_f \cdot F_{fusion} + b_f\right) \quad (12)$$

In this expression, W_f and b_f denote the weight matrix and bias term for the output classification stage, respectively. The final diagnostic output is determined by identifying the fault category that corresponds to the maximum probability score.

3. Experiments

3.1. Dataset Description

The widely utilized Case Western Reserve University (CWRU) bearing benchmark was selected in this research for model verification. The underlying experimental setup involves a 2-horsepower motor, a dynamometer, and associated power electronics. Vibration information was captured via magnetic-base sensors positioned at the 12 o'clock location on both the drive-end (DE) and fan-end (FE) housings. In our study, we specifically focused on the DE signals acquired at a 12 kHz sampling rate.

To assess the model's diagnostic accuracy in a multi-class context, ten distinct health states were defined, as summarized in Table 1. During the data preparation phase, the original records were segmented into 10,000 individual samples. Each sample, comprising 1,024 data points, was transformed into a 32×32 matrix to serve as the network input. Furthermore, the total sample pool was distributed into training, validation, and testing subsets according to a 7:2:1 proportion.

Table 1. Bearing fault types and descriptions.

Fault Location	Fault Diameter (mils)	Class Label	Sampling Frequency (kHz)
Normal Condition	0	0	12
Inner Race Fault	7 / 14 / 21	1 / 4 / 7	12
Outer Race Fault	7 / 14 / 21	2 / 5 / 8	12
Ball Fault	7 / 14 / 21	3 / 6 / 9	12

3.2. Analysis of Experimental Results

The model is trained using the Adam optimizer with a learning rate of 0.0003, a batch size of 32, and a total of 50 epochs. Cross-entropy is utilized as the loss function. As observed from the training and validation loss curves, the model exhibits excellent convergence performance. During the first 10 epochs, the loss values decrease sharply. After 20 epochs, both curves tend to plateau, eventually stabilizing around 0.05. Although the validation loss experiences minor fluctuations in the later stages, no overfitting occurs. Furthermore, the accuracy curves indicate a rapid increase in precision for both the training and validation sets, surpassing 98% around the 15th epoch and ultimately reaching 99.02% upon the completion of training. The achievement of such high accuracy is directly attributed to the multimodal fusion strategy (Figure 3).

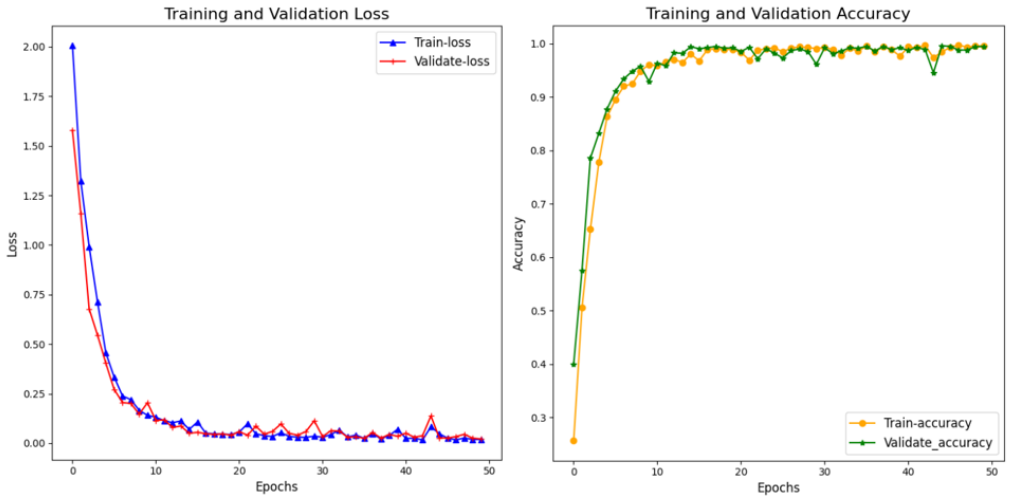

Figure 3. Training loss and accuracy curves.

To further illustrate the fault classification performance of the proposed model on the test set, the classification results were visualized. The t-SNE algorithm was employed to map high-dimensional features onto a 2D plane. By observing the clustering effect of the sample points, the feature extraction and classification capabilities of the model were intuitively demonstrated. Figure 4 illustrates the t-SNE visualization of the 2D feature distribution for both the raw test set data and the fusion feature vectors processed by the proposed model. Compared to the raw data, the visualization results have changed significantly, with feature clusters of the 10 states achieving clear separation in space.

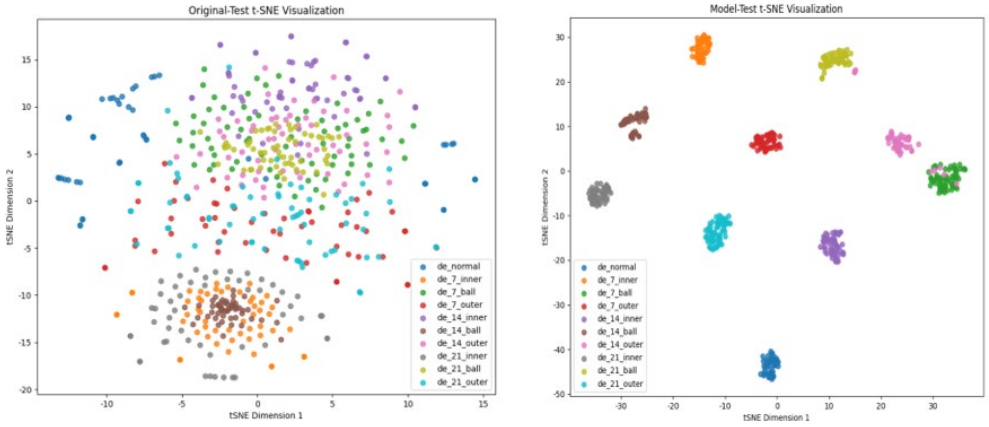

Figure 4. t-SNE visualization.

3.3. Ablation Experiments

An ablation framework with four alternative models was designed to quantify the role played by each primary element of the proposed system in diagnosing wind turbine bearing failures. All models were trained for 50 epochs under identical hardware environments using the CWRU dataset (0 HP load, 10-class classification), with all experimental parameters kept strictly consistent.

The ablation control groups were established as follows:

1. A temporal-only branch directly processing 1D raw signals;
2. A spatial-only branch processing GADF images without the CBAM module;
3. A dual-branch fusion architecture lacking the CBAM module in its spatial branch;
4. The fully intact multimodal fusion model proposed in this study.

Following the 50 training epochs, the average accuracy of each variant on the test set is detailed in Table 2:

Table 2. Results of the ablation experiments.

Model Configuration	Average Accuracy (%)
Single-branch: 1D Temporal (BiGRU)	91.24
Single-branch: 2D Spatial (Swin)	94.67
Dual-branch Fusion: No CBAM	97.58
Proposed Model (Full Model)	99.02

These ablation results sufficiently prove that the proposed fusion model not only boosts classification accuracy but also fortifies diagnostic robustness under complex working conditions via the synergistic interplay between its modules.

3.4. Comparative Experiments

To comprehensively validate the superiority of the proposed model in wind turbine bearing fault detection, this study designed comparative experiments involving five models (including the proposed one). The baseline models are detailed as follows:

WDCNN: A classic 1D convolution-based bearing fault diagnosis model, adept at extracting spatial features directly from raw vibration signals.

LSTM: A standard Recurrent Neural Network (RNN) architecture, commonly utilized for processing 1D bearing sequence data with long-range temporal dependencies.

ViT: An image classification model based on the standard Transformer architecture. Unlike the Swin Transformer, it employs fixed-size image patches and lacks hierarchical feature extraction capabilities.

ResNet-BiLSTM (Residual Fusion Model): Adopts the widely used ResNet-18 as the spatial branch and combines it with a BiLSTM for multimodal feature concatenation.

After 50 epochs of training under identical experimental conditions, the various classification evaluation metrics for each model on the test set are presented in Table 3.

Table 3. Results of the comparative experiments.

Diagnostic Model	Input Modality	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
LSTM	1D Raw Signal	82.35	81.92	82.10	82.01
WDCNN	1D Raw Signal	91.18	90.85	91.00	90.92
ViT	2D GADF Image	94.67	94.38	94.55	94.46
ResNet-BiLSTM	Multimodal Fusion	97.25	97.05	97.18	97.11
Proposed Model	Multimodal Fusion	99.02	98.95	99.00	98.97

Empirical findings indicate that the introduced framework maintains remarkable stability across key evaluation indices, including Precision, Recall, and F1-score. These metrics, which collectively reflect the overall efficacy of the diagnostic system, consistently converge at a level near 99%. This proves that the model avoids obvious false negative or false positive biases when handling complex faults, possessing extremely strong classification robustness. Ultimately, the proposed SwinT-CBAM-BiGRU model outperforms all other models in the experimental groups.

4. Conclusion

This paper proposes a multimodal fusion diagnostic model based on SwinT-CBAM-BiGRU. Compared with traditional single-branch models, such as WDCNN and LSTM, which exclusively process 1D signals, or single-image representation methods, the constructed dual-branch fusion architecture breaks through the feature limitations of a

single dimension. It achieves the synchronous extraction of instantaneous signal impacts and global topological features. Experimental results indicate that the fusion model reaches an accuracy of 99% on the CWRU dataset, proving that the multimodal architecture possesses stronger discriminative stability and robustness in handling complex background noise and multi-class fault recognition tasks.

Current diagnostic methods often rely on balanced samples from specific datasets. However, sample class imbalance and variable rotational speed interference, which are common in actual operating conditions, remain critical issues that have yet to be fully resolved. Future research should focus on enhancing the generalization capability of the model under extreme operating conditions. Furthermore, it is recommended to explore lightweight architectures to meet the real-time requirements of on-site monitoring for wind turbines, thereby providing more practical decision support for reducing quality costs and improving the operational reliability of wind turbines.

References

1. E. Hart, A. Turnbull, J. Feuchtwang, et al., "Wind turbine main-bearing loading and wind field characteristics," Wind Energy, vol. 22, no. 11, pp. 1534-1547, 2019.
2. Z. Liu and L. Zhang, "A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings," Measurement, vol. 149, p. 107002, 2020.
3. X. Chen, R. Yang, Y. Xue, et al., "Deep transfer learning for bearing fault diagnosis: A systematic review since 2016," IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-21, 2023.
4. W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis---Part I: Components and subsystems," IEEE Transactions on Industrial Electronics, vol. 62, no. 10, pp. 6536-6545, 2015.
5. T. Sun, G. Yu, M. Gao, et al., "Fault diagnosis methods based on machine learning and its applications for wind turbines: A review," IEEE Access, vol. 9, pp. 147481-147511, 2021.
6. Z. Chen, A. Mauricio, W. Li, et al., "A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks," Mechanical Systems and Signal Processing, vol. 140, p. 106683, 2020.
7. W. Lu and F. Chu, "Condition monitoring and fault diagnostics of wind turbines," in Proc. 2010 Prognostics and System Health Management Conference, IEEE, 2010, pp. 1-11.
8. D. Liu, W. Cheng, and W. Wen, "Rolling bearing fault diagnosis via STFT and improved instantaneous frequency estimation method," Procedia Manufacturing, vol. 49, pp. 166-172, 2020.
9. P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Rolling element bearing fault diagnosis using wavelet transform," Neurocomputing, vol. 74, no. 10, pp. 1638-1645, 2011.
10. D. Meng, H. Wang, S. Yang, et al., "Fault analysis of wind power rolling bearing based on EMD feature extraction," Computer Modeling in Engineering & Sciences, vol. 130, no. 1, p. 543, 2022.
11. S. Maurya, V. Singh, and N. K. Verma, "Condition monitoring of machines using fused features from EMD-based local energy with DNN," IEEE Sensors Journal, vol. 20, no. 15, pp. 8316-8327, 2019.
12. C. Hu, M. Huang, Q. Yang, et al., "On the use of EEMD and SVM based approach for bearing fault diagnosis of wind turbine gearbox," in Proc. 2016 Chinese Control and Decision Conference (CCDC), IEEE, 2016, pp. 3472-3477.
13. D. T. Hoang and H. J. Kang, "A survey on deep learning based bearing fault diagnosis," Neurocomputing, vol. 335, pp. 327-335, 2019.
14. A. U. Rehman, W. Jiao, Y. Jiang, et al., "Deep learning in industrial machinery: A critical review of bearing fault classification methods," Applied Soft Computing, vol. 171, p. 112785, 2025.
15. D. Wang, Q. Guo, Y. Song, et al., "Application of multiscale learning neural network based on CNN in bearing fault diagnosis," Journal of Signal Processing Systems, vol. 91, no. 10, pp. 1205-1217, 2019.
16. C. Guo, V. V. Potekhin, P. Li, et al., "MDFT-GAN: A Multi-Domain Feature Transformer GAN for Bearing Fault Diagnosis Under Limited and Imbalanced Data Conditions," Applied Sciences, vol. 15, no. 11, p. 6225, 2025.
17. K. Yeshun, W. Puzhou, and G. Yingkui, "Toward efficient and interpretative rolling bearing fault diagnosis via quadratic neural network with Bi-LSTM," IEEE Internet of Things Journal, vol. 11, no. 13, pp. 23002-23019, 2024.