International Journal of Power and Energy Systems, Vol. 46, No. 10, 2026
An Open Access Paper
Special Issue
Next Generation Energy Systems

CONSTRUCTION OF MOTOR FAULT DIAGNOSIS MODEL BASED ON TUNABLE INPUT BLSTM NETWORK

Feiyu Liu*

Abstract

In industrial production systems, motors are widely used across sectors such as transportation and energy production. Current diagnosis methods primarily rely on signal analysis and single-model recognition, resulting in low precision. The current diagnostic methods mainly rely on signal analysis and single-model recognition, which have deficiencies in capturing complex noise environments and long-sequence fault features, resulting in diagnostic accuracy often below 90% in actual industrial scenarios. This paper proposes a novel fault diagnosis model that deeply integrates a Belief Rule Base with a Bidirectional Long Short-Term Memory network. Diverging from conventional sequential or parallel structures, the proposed model introduces a "tunable input" mechanism, where the Belief Rule Base's real-time inference confidence directly gates and modulates the input channels of the Bidirectional Long Short-Term Memory, enabling dynamic, evidence-driven feature prioritization. The model processes uncertainty-embedded diagnostic signals and extracts long-range temporal dependencies to achieve accurate motor fault diagnosis. The test results indicate that the model achieves 99.82% precision, 99.45% recall, and 99.34% F1 score on the training set. In the loss value tests, the loss converges to 0.064 after 60 iterations, outperforming the comparison models. Overall, the proposed model effectively overcomes the deficiencies in reference and accuracy found in existing methods. It applies to multiple motor types and offers a new approach to motor fault diagnosis, enabling intelligent detection in industrial equipment.

Key Words

Belief rule base; Motor fault; Long short-term memory; Fault diagnosis

* School of Engineering, Nanjing Normal University Zhongbei College China; e-mail: liufeiyu.lfy@outlook.com

Recommended by: Zhenling Liu
(DOI:10.2316/J.2025.203-0656)

1. Background

As industrialization continues to advance, motors serve as the core power source for a wide range of mechanical equipment and are used across many applications. Their operating status directly affects production efficiency and equipment safety [1]. Accurate fault diagnosis is a key part of motor operation and maintenance. Low diagnostic accuracy may lead to extended downtime, while high-accuracy fault diagnosis ensures stable mechanical system operation [2]. Traditional motor fault diagnosis mainly depends on experience-based judgment and simple instrument testing [3]. These methods have drawbacks, including high misdiagnosis rates due to subjectivity and untimely fault warnings due to low efficiency. Therefore, how to achieve efficient and accurate motor fault diagnosis has become a research focus [4]-[5]. As computer technology advances, data-driven fault-diagnosis models have demonstrated emerging capabilities [6]. However, the existing mainstream methods have obvious limitations. For example, the model based on the combination of generative adversarial networks and convolutional neural networks (GAN-CNN) may have a diagnostic accuracy drop of more than 10% under substantial noise interference. Although Transformer-based models are good at modeling long sequences, their computational complexity is high, and the time per inference often exceeds 25ms, making it challenging to meet the real-time requirements of industry. These quantitative deficiencies (such as accuracy below 90% and high reasoning delay) seriously limit its practical application on industrial sites. Among them, the Belief Rule Base (BRB) and Long Short-Term Memory (LSTM) can be combined to form a BLSTM network. This network is good at processing fuzzy information and uses a gate-control mechanism to capture temporal dependencies in long sequences [7]. Applying this network to motor fault diagnosis enables the quantification of the logical relationships between key factors and fault types, and further links fault factors to dynamic features, thereby improving diagnosis efficiency. This paper proposes a BLSTM network for motor fault diagnosis, aiming to build a diagnosis network

with tunable inputs to achieve priority-based diagnosis and provide scientific decision-making support. The innovation lies in cross-domain technology integration and the evolution from static to dynamic diagnosis, contributing to equipment management and intelligent manufacturing.

2. Related Works

BRB integrates fuzzy set theory with Dempster-Shafer evidence theory. It effectively handled uncertainty and inaccuracy in classification tasks and could model nonlinear relationships. It played an important role in both the motor and network communication fields, and scholars at home and abroad researched it. To enhance the stability of power grid intrusion detection systems, Durairaj et al. proposed a method that integrates RBR with a greedy algorithm. They used RBR to build a rule base to identify intrusion types and dynamically update it using a greedy algorithm to match and reduce false alarms quickly [8]. Akiner et al. raised a method combining a fuzzy inference system with BRB to predict evapotranspiration. They first used the fuzzy inference system to process fuzzy inputs related to evapotranspiration, then integrated uncertain knowledge through the BRB rule base, and finally fused the evidence reasoning to output the prediction results [9]. LSTM, through its unique gate control mechanism, enabled the model to adapt dynamically to changes in sequence information and to learn complex sequence patterns. To address the potential short-term memory issue in LSTM applications, Kong et al. integrated segmentation and channel independence into the LSTM. They first segmented the input sequence, then introduced channel independence to enhance long-sequence modeling and to optimize performance on time-series tasks [10]. To address the problem of insufficient real-time performance in gesture navigation systems, Lu et al. proposed a multi-average LSTM prediction network. The network first segmented gesture sequences, then extracted multi-scale features, and simplified computation nodes to accelerate inference and achieve real-time gesture prediction [11]. To assess the health status of lithium batteries, Zhang et al. developed an LSTM-based assessment method. They collected historical data on voltage and current to build time-series samples, trained an LSTM network to learn battery degradation patterns, applied an attention mechanism to focus on key features, and finally generated a health assessment [12]. To address low accuracy in automatic subtitle generation for videos, Kavitha N et al. proposed a network combining an attention mechanism with an LSTM. They used an LSTM to process frame-sequence features and extract visual-sequence information. Then they applied an attention mechanism to focus on key frames, fusing sequence and key features to generate subtitles with improved accuracy [13].

Motor faults usually include electrical rotor faults and mechanical bearing faults. Their diagnostic accuracy determined whether the motor system could operate stably. Scholars at home and abroad conducted research on motor faults from various perspectives. To accurately detect faults in induction motors, Jigyasu et al. developed a method for vibration signal-based fault diagnosis. They collected vibration signals from induction motors, performed filtering and noise-reduction preprocessing, compared regular and faulty features, and identified fault types using expert rules to achieve accurate detection [14]. To improve motor diagnosis performance, Guo J et al. proposed a deep nonlinear-order recurrent convolutional network. They first segmented vibration signals into frames, then enhanced feature representation via a nonlinear-order transformation, captured temporal dependence with recurrent layers, and finally applied an attention mechanism to focus on fault features [15]. Hsu et al. proposed a method combining random forests and extreme gradient boosting for electromechanical fault diagnosis. They first extracted motor fault features, used random forests to select important features preliminarily, then applied extreme gradient boosting to learn nonlinear relationships, and merged the two to output diagnosis results [16]. For the fault diagnosis of voltage source inverters, Rokocakau et al. proposed a shallow neural network method. They collected voltage and current signals, extracted abnormal features, performed feature selection and dimensionality reduction, and input the data into the network to learn standard and fault patterns, achieving fast fault detection and localization [17].

However, prevailing data-driven diagnosis methods still exhibit notable limitations. Models based on Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA) often fail to capture the long-term temporal dependencies in motor vibration signals due to their inherent linearity and static nature. While deep learning models like Convolutional Neural Networks (CNNs) excel at local feature extraction, their ability to model long-range contextual relationships in sequential data is inherently constrained. Although Recurrent Neural Networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM), are designed for sequences, they can struggle to process fuzzy or uncertain information commonly found in industrial data, and their performance may degrade with extremely long sequences. Furthermore, methods that rely on manual feature engineering lack end-to-end optimization and are often not robust to noise. To address these gaps, this paper proposes a BLSTM network that explicitly integrates a BRB. The main advantages of this method are three: First, it retains and enhances the inherent gate mechanism of LSTM, effectively captures the long-term dynamic patterns in the fault evolution process, and directly overcomes the deficiencies of static models such as PCA-LDA; Secondly, the introduction of BRB provides a structured framework for handling the uncertainty and fuzziness of input signals, which is a capability lacking in standard lstm and cnn. Thirdly, the proposed "adjustable input" mechanism, under the guidance of BRB, enables the model to dynamically prioritize features, resulting in more robust and interpretable diagnoses than purely data-driven black-box models. This hybrid architecture achieves outstanding performance, especially with higher accuracy

on long-sequence data, and significantly reduces the false alarm rate under noisy conditions, thereby precisely addressing the weaknesses of existing methods.

While integrating knowledge bases with deep learning shows promise, existing approaches often treat the knowledge module as a static feature extractor or a post-hoc interpreter. The knowledge base and the neural network typically operate in a sequential or loosely coupled manner, lacking a closed-loop, synergistic interaction during feature learning and temporal modeling. This limits the model's ability to focus on the most relevant information under uncertainty adaptively. To bridge this gap, this paper proposes a Tunable Input BLSTM network, where the BRB is not merely a preprocessor but an active, integrated controller that dynamically tunes the BLSTM's input focus based on its evolving belief state, representing a significant structural and functional advancement over existing hybrid frameworks.

3. BLSTM Fault Diagnosis Model for Motor Faults

3.1 Optimized Design of BLSTM Network Based on Improved CNN

The BRB serves as a core component for modeling expert knowledge and handling uncertainty within the proposed model. It operates on a set of "if-then" rules where the antecedents (IF parts) are composed of belief distributions over the reference values of input attributes (e.g., vibration amplitude can be "Low," "Medium," or "High" with certain degrees of confidence), and the consequents (THEN parts) are belief distributions over possible fault outcomes. This structure allows BRB to formally represent and process the fuzziness and uncertainty inherent in real-world motor signals. The inference mechanism of BRB, namely the Evidential Reasoning (ER) algorithm, aggregates activated rules based on their matching degrees and rule weights, ultimately producing a unified belief distribution that quantifies the confidence for each potential fault type. In our architecture, this output belief distribution from the BRB is not the final diagnosis. However, it is strategically utilized as a refined, high-level input feature sequence for the subsequent BLSTM network. This integration enables the BLSTM to focus on learning the temporal evolution of these semantically rich "belief states," thereby effectively combining uncertainty reasoning with deep sequential pattern recognition. The parameters of the BRB system are determined through a hybrid strategy of expert initialization and data optimization: Firstly, the reference values of the prerequisite attributes (such as high/medium/low vibration amplitude) and the initial rule confidence are manually set based on the statistical characteristics of the training data and domain knowledge. Subsequently, the projection covariance Matrix Adaptive Evolution Strategy (CMA-ES) was adopted to jointly optimize the rule weights, attribute weights, and result confidence levels, with the diagnostic accuracy of the training set as the objective function. For example, the typical rule form is: IF Vibration_RMS IS High AND Spectral_Centroid IS High THEN (Bearing_Fault,0.85), (Rotor_Fault,0.10), (Normal,0.05). Motor fault diagnosis monitors and judges the operation of equipment and systems to identify faults and determine fault types. It ensures safe operation and reduces maintenance costs [18]. This paper uses BRB due to its strong ability to handle uncertainty in fault-feature signals. It also has strong adaptability to complex data. LSTM captures dynamic changes in information over time and the context relationship. It effectively extracts hidden features in fault signals. The combination of both gives advantages in motor fault diagnosis. To verify the necessity and superiority of combining the BRB and LSTM modules in the BLSTM model, an ablation experiment was conducted. Four models, namely BRB-ONLY, LSTM-ONLY, BLSTM-series, and BLSTM (proposed in the study), were compared. Among them, the BLSTM series is a simple structure. That is, the LSTM output is used as the input to the BRB to compare the effectiveness of the fusion structure proposed in this paper. The test results are shown in Table 1.

Table 1
Results of the Ablation Experiment

Model	Accuracy rate	F1 points
BRB-only	89.45%	88.91%
LSTM-only	93.72%	93.15%
BLSTM-series	96.88%	96.45%
BLSTM (proposed in the study)	99.82%	99.34%

In Table 1, the BRB-only model has the lowest accuracy, indicating that it is challenging to capture complex dynamic patterns in fault signals solely through expert rules and uncertain reasoning. The LSTM-only model outperforms the BRB-only model, demonstrating the advantages of deep learning for feature extraction. However, its performance ceiling is limited by the noise and uncertainty in the original data. The performance of the BLSTM-series model has been significantly improved, indicating that even simple combinations can bring gains. However, its performance is still lower than that of the model proposed in this paper. The BLSTM model in this paper has achieved the best performance. This fully demonstrates the effectiveness of the proposed fusion architecture: The BRB module first performs "defuzzification" and "structuring" on the input information, providing a cleaner, more discriminative feature sequence for the LSTM; the LSTM then conducts precise time-series modeling on this basis. The innovation of this model lies in the collaborative working mechanism between the two: BRB, as the front-end inference engine, transforms the original, fuzzy, and uncertain signals into regular features with a confidence distribution; LSTM, as the backend timing analyzer, conducts in-depth modeling of this feature sequence. Therefore, this paper proposes a BLSTM network. In this study, the BLSTM classifier adopts a single-layer bidirectional LSTM structure, with 128 hidden units in each direction and an output dimension of 256. To prevent overfitting,

a Dropout layer with a dropout rate of 0.3 is set after the BLSTM layer. The classifier part consists of a fully connected layer with 64 neurons (ReLU activation) and an output layer (Softmax activation), with a Dropout rate of 0.2 in between. The model uses the Adam optimizer (learning rate 1e-4), a batch size of 64, and is trained for 100 cycles, adopting the early stopping strategy. The core innovation of this BLSTM network lies in its "Tunable Input" mechanism. Unlike traditional models with fixed input pipelines, our model dynamically adjusts the input data weighting and processing flow in response to real-time feedback from the BRB inference module. This tunability allows the model to prioritize more uncertain or critical fault features, enhancing diagnostic focus and accuracy. The operation process of this network is shown in Figure 1.

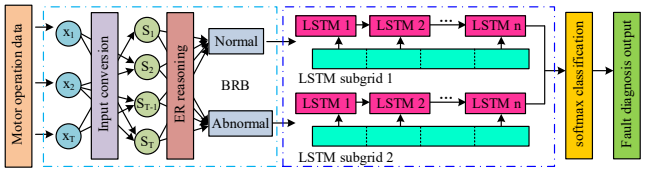

Figure 1. Operation Flow Chart of BLSTM Network

As shown in Figure 2, the BLSTM network first uses BRB to label motor operation data. It labels normal values and abnormal values. Two LSTM subnetworks process normal values and abnormal values separately. Based on the LSTM input dimensions, the two subnetworks are split by time series and used as inputs for network training. The final output is the motor fault features [19]. The tunability of the input is implemented through a gating mechanism that connects the BRB and the dual LSTM subnetworks. The confidence distribution generated by the BRB at each time step serves not only as a feature but also as a control signal. A tunable parameter, which can be optimized during training, governs the influence of this distribution. When the BRB's output indicates high uncertainty (e.g., confidence is evenly distributed across multiple fault types), the parameter increases the flow of information to the LSTM subnetwork dedicated to anomaly analysis, enabling more nuanced feature extraction from ambiguous signals. Conversely, when the BRB output shows high confidence in a particular state (normal or a specific fault), the parameter strengthens the path to the corresponding LSTM subnetwork, refining the feature sequence for that specific class. This dynamic routing embodies the Tunable Input, enabling the model to adaptively reconfigure its input focus based on instantaneous diagnostic evidence. The BRB converts an input value of a premise attribute as shown in Equation (1).

$$S(B_i, \alpha_i) = (l_{im}, \beta_{im}), m = 1, ..., m_i \tag{1}$$

In Equation (1), S represents the distribution of the input value of the premise attribute. α_i represents the confidence of premise attribute B_i. l_{im} represents the $m-th$ reference value of the input premise attribute B_i. β_{im} represents the confidence of l_{im}. m_i represents the number of reference values. Then, the evidence reasoning converts and integrates all premise attribute input values under the confidence rules. The calculation is shown in Equation (2).

$$K = \left[\sum_{j=1}^{N} \prod_{k=1}^{L} (m_{j,k} + m_{D,k}) - (N-1) \prod_{k=1}^{L} m_{D,k} \right]^{-1} \tag{2}$$

In Equation (2), mj,k is the basic probability allocation assigned to grade by Rule k, $m_{D,k}$ is the probability quality not assigned to any grade, and L is the total number of rules in the rule base. N represents the number of input values of the premise attribute. The LSTM module's output is shown in Equation (3).

$$H_t = \theta t \odot tanh(C_t) \tag{3}$$

In Equation (3), H_t represents the output gate result of the LSTM. θt represents the result of the input gate. $\odot$ represents element-wise multiplication. $tanh$ represents the activation function. C_t represents the cell gate state. The BLSTM network first uses the BRB module to handle uncertainty in motor data. Then, the LSTM module extracts deep-time-series features for motor fault diagnosis. However, the BLSTM network relies heavily on LSTM for long-term temporal dependence. It has low sensitivity to local high-frequency transient features in motor fault data. It may ignore local key information, reducing diagnostic accuracy in the early stage. The Convolutional Neural Network (CNN) improved by the Attention Mechanism (Attention) weights CNN local features with Attention. The weighted CNN uses the sliding-window convolution operation to capture local, high-frequency transient features precisely [20] [21]. The operation process of the improved CNN is shown in Figure 2.

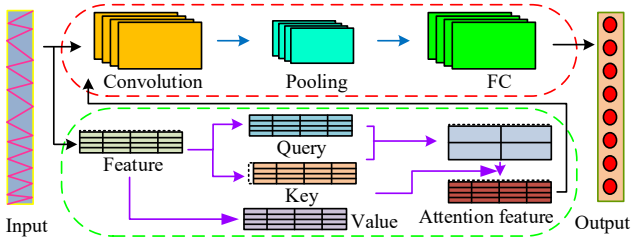

Figure 2. Improved CNN Operation Flow Chart

As shown in Figure 2, the improved CNN receives data first. The Attention module computes the weight for each feature using the attention layer. The weighted data is sent to the CNN module. The CNN captures local features through the convolution layer with a sliding window. The pooling layer compresses feature dimensions and retains local maxima. Finally, the fully connected layer outputs the weighted result [22]. The output of the attention layer of Attention is shown in Equation (4).

$$Y = softmax(\frac{Q^T K}{\sqrt{D_k}})V \tag{4}$$

In Equation (4), $Q^T K$ is the attention matrix. $\sqrt{D_K}$ is the scaling factor. V represents the value matrix. $softmax$ represents the activation function. In the model of this paper, the number of heads h is a key hyperparameter. After preliminary experiments and grid search, this study finally set h=8. This configuration achieves the best balance between computational efficiency and model performance; the multi-head self-attention calculation is shown in Equation (5).

$$Multi - head(Q, K, V) = Concat(att(Q_1, K, V) \cdots att(Q_H, K, V)) \tag{5}$$

In Equation (5), $Multi - head$ represents the matrix. $Concat$ represents the matrix product. att represents the single-head self-attention. The output of the CNN convolutional layer is presented in Equation (6).

$$Z = RF_{i-1} + (K_i - 1) \times \prod_{j=1}^{i-1} S_j \tag{6}$$

In Equation (6), Z represents the output of the convolution layer. RF_{i-1} is the receptive field, K_i represents the convolution kernel, and S_j represents the stride of layer j. The core purpose of the improved CNN attention mechanism adopted in the research is to capture local high-correlation features in motor fault signals more efficiently, while avoiding the high computational complexity of standard Transformer self-attention when processing long sequences. The computational complexity of the standard self-attention mechanism is $O(n^2 \cdot d)$, where n is the sequence length and d is the feature dimension. This incurs a significant computational overhead for the vibration signal sequences of long-running motors. The core operations of the improved CNN attention mechanism proposed in the research are the convolutional sliding window and attention weighting. For an input with a convolution kernel size of k, a sequence length of n, and a feature dimension of d, its computational complexity is mainly determined by the convolution operation, which is $O(n \cdot k \cdot d^2)$. It offers significant computational efficiency and is better suited for diagnostic tasks in industrial scenarios that require real-time performance. Standard self-attention features a global receptive field and minimal inductive bias, making it highly flexible. However, it is prone to overfitting when the data volume is limited and may also focus on all time points, being insensitive to local impact characteristics that are physically significant in motor signals. The improvement of CNN attention: The convolution operation introduces a strong inductive bias of locality and translation invariance, which is highly consistent with the characteristics of periodic impact signals generated by motor faults (such as bearing pitting). This mechanism forces the model to first construct feature relationships within local Windows and then weight them through attention, enabling it to focus more accurately and efficiently on the key fault features generated by local physical processes. In conclusion, this method is not intended to surpass the Transformer in terms of universality, but rather serves as a more efficient, proprietary alternative designed for the specific task of motor fault diagnosis. Under the premise of comparable accuracy, it significantly enhances computational efficiency and has greater engineering practical value. The improved CNN has dual abilities of local feature extraction and dynamic focusing. It also improves local feature quality and multi-modal fusion efficiency. BLSTM processes fuzzy information efficiently and captures long-term sequence dependence accurately to achieve motor fault diagnosis. Therefore, this paper combines the improved CNN with BLSTM to form CBLSTM. The operation process is shown in Figure 3.

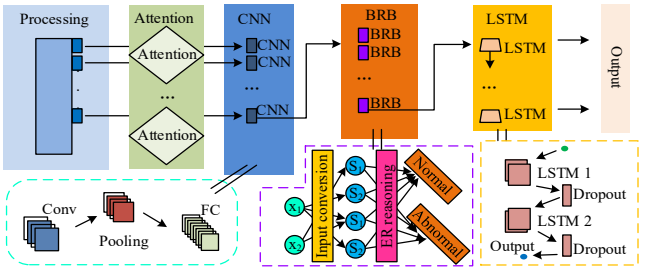

Figure 3. CBLSTM Operation Flow Chart

Figure 3 illustrates that in the CBLSTM operation process, the data undergoes normalization first. The Attention module enhances the expression of fault data features. Then CNN captures local features. BRB processes fuzzy features. LSTM captures global features and performs time series modeling. Finally, the fault diagnosis result is output. The CNN module focuses on extracting local data features. Its output H_c is shown in Equation (7).

$$H_c = Tanh(P \times W + b) \tag{7}$$

In Equation (7), $Tanh$ represents the hyperbolic tangent function. P represents the output of the pooling layer. b represents the bias. The BRB module uses the probability weighting method for solving. The weighting coefficient S_t is shown in Equation (8).

$$S_t = \sum_{t=1}^{i} a_t h_t \tag{8}$$

In Equation (8), a_t represents the corresponding attention weight. h_t represents the hidden state of t at the corresponding time. Finally, the output layer of LSTM outputs the result Y_t, calculated as shown in Equation (9).

$$Y_t = Relu(W_j s_t + b_0) \tag{9}$$

In Equation (9), $Relu$ is the activation function. W_J is the weight system. S_t is the weighting coefficient. b_0 is the bias vector. During BLSTM motor fault diagnosis, the improved CNN dynamically assigns weights to capture spatial correlation in short time sequences precisely. It improves the representation and capture of local features. It

compensates for the omission of local features in BLSTM. This enhances the efficiency and accuracy of motor fault diagnosis.

3.2 Construction of PV-CBLSTM Motor Fault Diagnosis Model

The CBLSTM model demonstrates superior diagnostic accuracy in controlled experimental settings. However, in practical industrial settings, sensor-collected vibration signals are invariably contaminated by significant background noise from complex operating environments, such as electromagnetic interference and mechanical impacts. This noise interferes with the CNN module's extraction of discriminative local features. It introduces fluctuations in confidence levels during BRB rule-matching, ultimately degrading the accuracy and reliability of the diagnostic results. To address these practical challenges, this study introduces an enhanced Variational Mode Decomposition (VMD) technique, optimized via a Bayesian-improved Particle Swarm Optimization (PSO), into the CBLSTM framework, forming the final PV-CBLSTM fault diagnosis model. The VMD improved by Wavelet Threshold Denoising (WT) adaptively decomposes signal components based on their inherent characteristics, enabling targeted isolation of various noise types. Although CBLSTM accurately diagnoses motor faults in specific experimental settings, the signals collected by sensors during practical diagnosis often contain significant noise. This noise reduces the ability to extract local features and increases fluctuations in rule-matching confidence, thereby decreasing the accuracy of diagnostic results. The Variational Mode Decomposition (VMD) improved by Wavelet Threshold Denoising (WT) decomposes modal components adaptively based on signal characteristics and targets various types of noise. In industrial practice, it is often difficult to obtain an accurate prior distribution for faults. Therefore, this study adopts a strategy that combines an information-free prior with dynamic updates. At the beginning of the optimization, due to a lack of knowledge, this study assumes that the potential states of all particles are equally possible, that is, this study sets a uniform prior. Subsequently, during the PSO iteration, this study uses the continuously updated historical information from the particle swarm to dynamically adjust the prior. Specifically, by correlating the prior probability with the historical optimal fitness distribution of the particle swarm, the entire process is unsupervised. It relies solely on fitness information generated during optimization, thereby addressing the problem of scarce labeled data in engineering practice. Therefore, this study introduces the improved VMD into CBLSTM to achieve noise reduction. The noise reduction process of the improved VMD is shown in Figure 4.

In Figure 4, when the improved VMD processes noisy features, it first decomposes multiple noisy signals into modes through VMD. Then VMD selects noisy IMF signals and sends them into the WT module. The WT module identifies high-frequency and low-frequency signals. If a signal is high-frequency, it denoises the amplitude factor

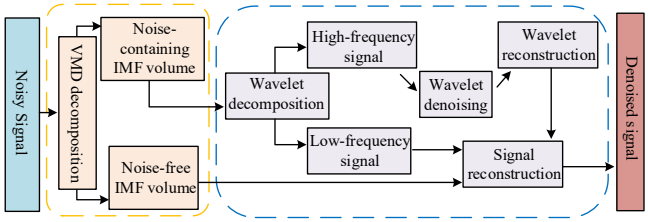

Figure 4. Improved VMD Denoising Flowchart

and reconstructs the signal by wavelet reconstruction. The reconstructed signal and the low-frequency signal are then reconstructed again to obtain the denoised signal. For each modal sequence, the Hilbert transform calculates the analytic signal related to each modal sequence to obtain the single-sided spectrum signal, as shown in Equation (10).

$$Hilbertx(t) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{x(v)}{t-v}dv \tag{10}$$

In Equation (10), $Hilbert$ represents the sequence, and v represents the frequency speed. In signal processing, the hard-threshold function of the WT is given by Equation (11).

$$\widehat{\omega}_{j,k} = \begin{cases}\omega_{j,k}, |\omega_{j,k}| \geq \lambda \\ 0, |\omega_{j,k}| < \lambda\end{cases} \tag{11}$$

In Equation (11), $\hat{\omega}_{j,k}$ denotes the wavelet coefficients processed by the threshold function, $\omega_{j,k}$ denotes the wavelet coefficients, and λ represents the threshold. The Signal-to-Noise Ratio (SNR) measures the ratio of the signal to the noise. A higher SNR indicates better optimization. To verify that adding WT effectively reduces noise interference, SNR is selected as the metric, and its calculation is shown in Equation (12).

$$SNR = 10\ log\frac{P_s}{p_n} = \frac{\sum_{n=0}^{N-1}x_s(n)^2}{\sum_{n=0}^{N-1}x_n(n)^2} \tag{12}$$

In Equation (12), P_s denotes the signal power, P_n is defined as the noise signal power, $x_s(n)$ represents the processed signal, and $x_n(n)$ represents the noise signal. The improved VMD effectively separates noisy signals, balances the modal components, and reduces the likelihood of false modes. However, the improved VMD is easily disturbed by impact features, leading to mode confusion, and WT lacks adaptive capability against dynamic noise. The Particle Swarm Optimization (PSO) enhanced by Bayesian Decision (BD) optimizes key processes via intelligent parameter search and probabilistic selection, demonstrating strong stability in complex, dynamic environments. Therefore, this study introduces improved PSO into the improved VMD to address subjectivity in VMD mode selection and the WT's insufficient adaptability to dynamic noise. The operation process of the improved PSO is shown in Figure 5.

As shown in Figure 5, the single-particle data of the improved PSO are optimized into PSO particles under the decision differentiation of BD. After extracting features

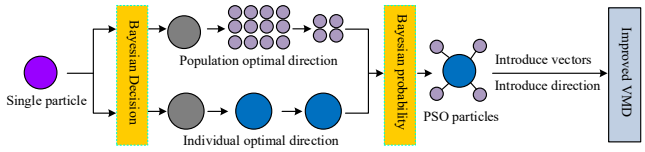

Figure 5. Improved PSO Operation Flow Chart

from different types of PSO particles, the PSO reconstruction particles are formed under the combined effect of BD probabilistic differentiation and PSO optimization. These reconstructed particles are then used as the input of the improved VMD. The probabilistic selection of particles by BD is calculated as shown in Equation (13).

$$P(\omega_j \mid X) = \frac{P(X \mid \omega_j) P(\omega_j)}{P(X)} \tag{13}$$

In Equation (13), $P(\omega_j)$ represents the prior probability of the particle in its natural state, and $P(X)$ represents the random variable of the feature vector. The optimization precision of PSO particles is shown in Equation (14).

$$\theta = \theta - \eta \cdot \nabla_\theta J(\theta; x^{(i)}; y^{(i)}) \tag{14}$$

In Equation (14), θ represents the parameter to be optimized, $J(\cdot)$ is defined as the loss function, $x^{(i)}$ represents the PSO-optimized particle, and $y^{(i)}$ represents the related factor of the PSO-optimized particle. PSO constructs a comprehensive influence matrix T to represent the combined effects of direct and indirect influences among factors, as shown in Equation (15).

$$\begin{aligned} T &= \lim_{I \to \infty} \left(X + X^2 + \cdots + X^I \right) \\ &= \lim_{I \to \infty} \left(E + X + X^2 + \cdots + X^{I-1} \right) \end{aligned} \tag{15}$$

In Equation (15), X represents the direct influence matrix, E represents the identity matrix, and I represents the number of factors. The improved VMD decomposes noisy signals effectively for motor fault diagnosis, thereby improving accuracy, while the improved PSO optimizes VMD's adaptability to dynamic noise and resolves mode confusion. Therefore, the combination of improved VMD and improved PSO effectively reduces the influence of external noise on motor fault diagnosis. Combined with CBLSTM, this study finally constructs the PV-CBLSTM fault diagnosis model, and its operation process is shown in Figure 6.

As shown in Figure 6, all vibration signals undergo a standardized preprocessing process: firstly, a 4-order Butterworth bandpass filter (10-2000Hz) is used for noise reduction, followed by Z-score standardization processing. On this basis, a 30-dimensional feature vector, including time-domain statistical features (such as mean, root mean square, peak, kurtosis, etc.) and frequency-domain features (such as the spectral centroid and the root mean square frequency after the FFT transformation), is extracted from the preprocessed signal. For time-series models, these features are constructed as sequence samples of

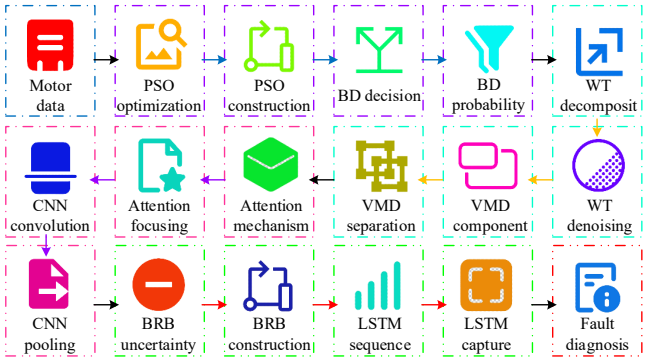

Figure 6. PV-CBLSTM Fault Diagnosis Model Operation Flow Chart

100 time steps to ensure that the data format matches the model architecture. The PV-CBLSTM fault diagnosis model uses an improved PSO module to perform dynamic optimization and probabilistic selection of input data, thereby improving data quality. Then the improved VMD module decomposes the original motor data and isolates noise, thereby increasing the accuracy of diagnosis. The improved CNN module extracts local features from the denoised signals to enhance feature representation. Finally, the BLSTM module outputs the diagnosis results for motor fault types and severity via BRB-based uncertainty reasoning and LSTM-based temporal modeling. In summary, the PV-CBLSTM fault diagnosis model not only addresses the impact of noisy signals on fault diagnosis but also ensures data quality with its excellent generalization ability and robustness. It achieves accurate motor fault diagnosis and provides strong technical support for industrial motor fault diagnosis. The confidence vector generated by BRB reasoning is incorporated into the LSTM's calculation of its forgetting gate. The expression of the coupled forgetting gate is shown in Equation (16).

$$f_t = \sigma(W_f \cdot [h_{(t-1)}, x_t] + bf + WBRB \cdot bt) \tag{16}$$

In Equation (16), f_t represents the coupled forgetting gate state vector at time step t, σ represents the Sigmoid activation function, W_f represents the inherent weight matrix of the forgetting gate, $[h_{(t-1)}, x_t]$ represents the basic input of the forgetting gate, W_{BRB} represents the BRB coupled weight matrix, and b_t represents the confidence distribution vector generated by the inference of the BRB module.

4. Performance Analysis of Motor Fault Diagnosis Model Based on CBLSTM

4.1 Validation of CBLSTM Effectiveness

To ensure a comprehensive and fair evaluation, the proposed CBLSTM model is compared against a diverse set of baseline models, encompassing both classical machine learning and modern deep learning approaches. The baselines include:Informer: A state-of-the-art Transformer-

Table 2
Compare the Model Structure with the Hyperparameter Settings

Model	Core architecture	Key hyperparameters	Optimizer and Learning rate
CBLSTM	Improved VMD-PSO + BRB + Improved CNN Attention + BLSTM	BLSTM units: 128 (per direction) Attention heads: 8 CNN kernels: 32, 64, 128 Dropout: 0.3, 0.2	Adam (1e–4)
Informer	Encoder–Decoder with ProbSparse Self-Attention	Encoder/Decoder layers: 2 d_model: 512 Attention heads: 8	Adam (1e–4)
GAN-CNN	DCGAN Generator + 1D-CNN Discriminator	Latent dimension: 100 CNN filters: 32, 64, 128 Dropout: 0.3	Adam (G: 1e–4, D: 2e–4)
DT-RF	Ensemble of 100 Decision Trees with Weighted Voting	Max depth: 15 Min samples split: 5 Criterion: Gini	Ensemble Method

based model for long-sequence time-series forecasting, included to benchmark against advanced temporal modeling capabilities.GAN-CNN: A generative adversarial network coupled with a convolutional neural network classifier, representing hybrid generative-discriminative approaches.DT-RF: A Decision Tree-Random Forest ensemble with weighted voting, serving as a firm representative of traditional, interpretable machine learning methods. To guarantee a fair comparison, all deep learning-based models (including CBLSTM, Informer, and GAN-CNN) were trained and evaluated under identical hardware and software environments, using the same data splits, and were tuned to their optimal performance on the validation set. Among them, in order to construct a competitive generative adversarial network comparison model, this study implements a GAN-CNN architecture that combines a standard deep convolutional generative adversarial network (DCGAN) with a one-dimensional convolutional neural network (1D-CNN) classifier. Its overall framework consists of a generator, a discriminator (which also serves as a feature extractor in the classification task), and a classifier. The goal of the generator is to upsample a random noise vector z to generate pseudo-data that is similar in dimension and statistical characteristics to the real motor vibration signal time series. The discriminator is essentially a convolutional neural network (CNN). In the adversarial training stage, it acts as a Discriminator (D); In the fault classification stage, this study removes its final output layer, uses it as a Feature Extractor, and connects an additional Classifier (C) for fault diagnosis. The parameters of each model are shown in Table 2

As shown in Table 2, the configuration list for all comparison models is clearly presented, which significantly enhances the repeatability and technical transparency of the experiment. The experimental operating system was Windows 10, the deep learning framework was TensorFlow, the optimizer was Adam, the programming language was MATLAB, the GPU was NVIDIA RTX 4090, the CPU was Intel i7, and the memory was 256 GB. The performance of the model proposed in the research is affected by several key hyperparameters, including the number of heads h for improving multi-head Attention in CNNs and the number of hidden-layer units in LSTM. To determine the optimal configuration, this study used a grid search. Search space and Settings: Number of attention heads h: 4,8,6. Number of LSTM hidden units: 64,128,256. The learning rate of the optimizer Adam: 1e-3, 1e-4, 1e-5. To ensure the authenticity of the test results, the CSDN dataset was used as the training set, and the CWRU dataset as the validation set. Both datasets contained abundant motor fault severity and operation data, making them suitable for motor fault diagnosis and analysis. For a rigorous evaluation, all datasets were partitioned using a hold-out method. Specifically, the CSDN dataset was divided into a training set (80%) and a testing set (20%). The CWRU and FEMTO datasets were similarly split, with 80% of the data used for training and the remaining 20% reserved as the validation set to monitor training progress and prevent overfitting. This consistent split ratio across datasets ensures a fair and reproducible comparison of model performance. To ensure that the experimental conditions are precise and repeatable, this study selected vibration acceleration data from the drive-end bearing in the CWRU dataset. The fault types include four states: normal, inner-ring fault, outer-ring fault, and rolling-element fault. The fault-damage diameter was set to 0.1778mm, a typical minor-damage size, to test the model's sensitivity to early faults. The load condition was 1 hp, and the motor speed was approximately 1772 RPM. The sampling frequency for vibration data is 12 kHz. The loss value of CBLSTM was tested, and the results are shown in Figure 7.

In Figure 7(a), the initial loss value of CBLSTM under variable-load industrial operation was 0.125, which decreased to 0.064 after 60 iterations and consistently stabilized. As illustrated in Figures 7(b), (c), and (d), the loss curves of Informer, GAN-CNN, and DT-RF—trained and validated under the same industrial scenario—eventually stabilized at 0.126, 0.138, and 0.101, respectively. Notably, all three baseline models exhibited higher initial loss values and required more iterations to converge than CBLSTM. These results confirm CBLSTM's stronger generalization capability in real industrial environments. Its

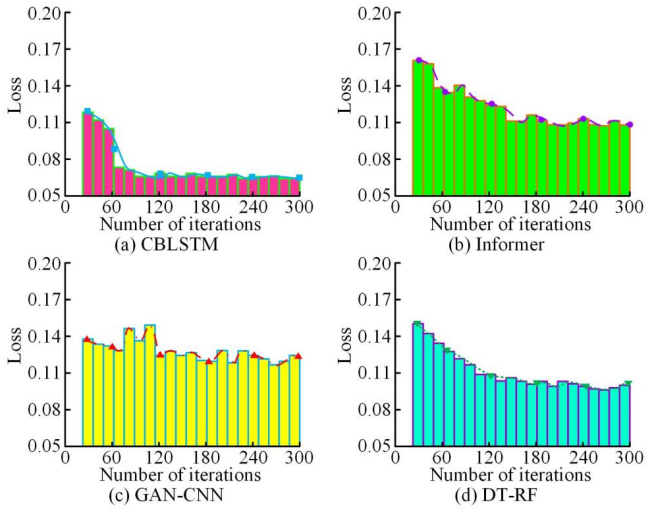

Figure 7. Comparison of Loss Value Test Results

superior convergence behavior stems from the Attention mechanism's emphasis on salient industrial vibration patterns and CNN's strength in extracting localized fault signatures, which together enable more representative feature learning and robust optimization. The Mean Absolute Error (MAE) and Mean Relative Error (MRE) were then tested, and the results are shown in Figure 8.

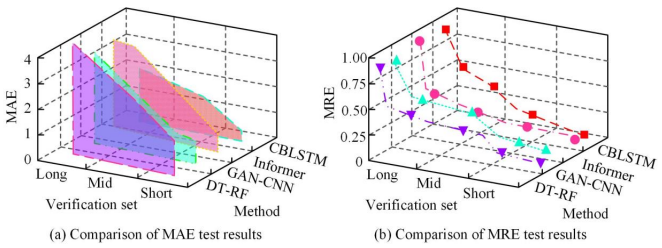

Figure 8. Comparison of MAE and MRE Value Test Results

In Figure 8(a), the MAE of CBLSTM was lower than that of the other models in all datasets for short-term, mid-term, and long-term tests. The overall MAE was 0.98, while the MAEs for Informer, GAN-CNN, and DT-RF were 1.3, 1.9, and 2.3, respectively. Figure 8(b) showed that the overall MRE of CBLSTM was 0.18, which was lower than that of the comparison models. In summary, CBLSTM demonstrated superior performance in error value training. This was attributed to its inherent BRB-based uncertainty reasoning and LSTM-based temporal feature modeling capabilities, which reduced bias in data capture and effectively lowered computational errors. Precision, recall, and F1 score were then tested, and the results are shown in Figure 9.

In Figure 9(a), the precision, recall, and F1 score of CBLSTM on the training set were 99.82%, 99.45%, and 99.34%, respectively. GAN-CNN and DT-RF performed worse, with all indicators below 90%. Figure 9(b) showed that on the validation set, CBLSTM achieved an accu-

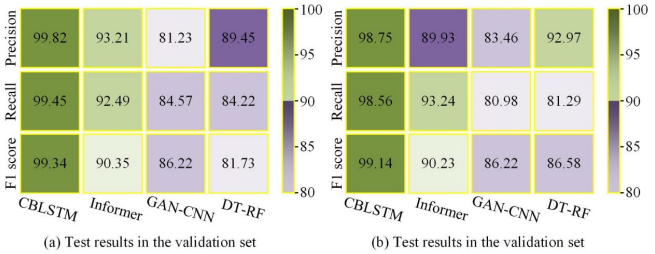

Figure 9. Comparison of Precision, Recall, and F1 Score Test Results

racy of 98.75%, a recall of 98.56%, and an F1 score of 99.14%, all of which were higher than those of the comparison models. In summary, CBLSTM demonstrated strong fitting ability and superior overall performance in the tests. This was attributed to the unique advantages of integrating CNN, PSO, and other algorithms, which complemented each other across dimensions, enabling stable convergence during iteration and improving indicators such as precision and recall, thereby showing more comprehensive performance in fault diagnosis.

4.2 Analysis of PV-CBLSTM Motor Fault Diagnosis Model

Following the validation of the CBLSTM's core performance, this section evaluates the comprehensive PV-CBLSTM model under more realistic, noisy industrial conditions. The integration of the denoising front-end (PSO-VMD) is specifically designed to enhance the model's robustness, and its performance is compared against models constructed using Informer, GAN-CNN, and DT-RF. The experimental settings are shown in Table 3.

Table 3
Experiment Configuration Details

Item	Configuration
CPU	AMD Ryzen 9 7950X
GPU	NVIDIA GeForce RTX 4090
Memory	DDR5 8000MHz
Storage	ZhiTai TiPro9000
Operating System	Windows 10
Training set	FEMTO Dataset
Validation set	Purdue University Dataset
Data analysis software	Python

In Table 3, both datasets included motor fault states and operational data, thereby enabling more accurate test results. These two datasets differ in acquisition systems, working conditions, and original feature dimensions. This study did not use the original data directly; instead, this study uniformly extracted the same set of time-domain and frequency-domain features from the original signals of both datasets. This move aims to map data from different sources into the same comparable feature space. Standardize the features extracted above by Z-score, that is, subtract the mean and divide by the standard deviation. This

operation involves calculating the mean and standard deviation for each dataset's training set, then applying them to the corresponding training and validation sets. According to the model's input requirements, the standardized feature data is constructed into time-series samples of the same length. The CBLSTM model was used to identify three types of faults: bearing faults, rotor faults, and shaft system faults, labeled as A, B, and C. The results are shown in Figure 10.

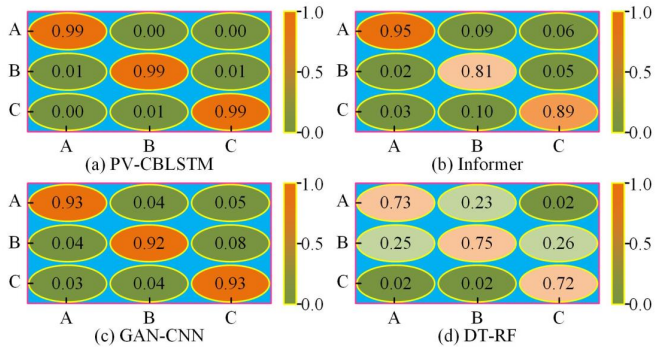

Figure 10. Identification Test of Three Fault Types

As shown in Figure 10(a), the CBLSTM model correctly identified all three fault types—bearing, rotor, and shaft system faults (labeled A, B, and C)—with an accuracy of 0.99 under real plant operating conditions. In contrast, Figures 10(b)–(d) reveal that although GAN-CNN achieved an accuracy of 0.90, it underperformed CBLSTM. Both Informer and DT-RF showed considerably weaker industrial applicability, with Informer's accuracy falling at or below 0.75. For instance, in diagnosing shaft system faults, Informer misclassified them as bearing faults with a probability of 0.02 and as rotor faults with a probability of 0.26. Overall, CBLSTM delivered significantly more reliable performance in an actual industrial context. This advantage stems from the model's effective integration of a wavelet transform for noise-resistant time–frequency characterization and variational mode decomposition for distinguishing modulation modes, enabling it to suppress interference while retaining discriminative fault characteristics. Varying levels of noise often contaminate vibration signals collected at industrial sites. To quantitatively evaluate the noise robustness of each model, this study systematically injected additive Gaussian white noise with a signal-to-noise ratio (SNR) ranging from 10 dB to -5 dB into the preprocessed Purdue University validation set data, and recorded the diagnostic accuracy of each model. The results are shown in Table 4.

Table 4 shows that as the signal-to-noise ratio decreases (noise intensity increases), the diagnostic accuracy of all models declines. However, the PV-CBLSTM model proposed in this paper maintains the highest diagnostic accuracy under all signal-to-noise ratio conditions. Especially in extremely harsh environments with a low signal-to-noise ratio (SNR < 0 dB), the performance advantages of PV-CBLSTM are even more significant. This fully demonstrates the effectiveness of its internally integrated improved VMD and improved PSO modules in collaborative noise reduction and feature enhancement: VMD is responsible for stripping the noise-dominated modes from the signal, while PSO optimizes the parameters of this process, ensuring that the core fault features are retained to the greatest extent under strong noise interference, thereby endowing the model with excellent noise robustness. To further validate the generalization capability of the proposed model across different motor operating conditions, comparative experiments were conducted against mainstream time-series models (Transformer, LSTM, TCN) and tested across a range of motor power levels (0.5kW to 50kW). The results are shown in Figure 11.

Table 4
Test Results of Multiple Comprehensive Indicators

Signal-to-noise ratio (dB)	CBLSTM	Informer	GAN-CNN	DT-RF
10	98.75	95.81	92.34	88.90
8	98.52	94.22	90.15	85.61
6	98.01	92.05	87.33	81.04
4	97.45	88.91	83.72	75.88
2	96.23	84.56	78.95	69.42
0	94.67	78.33	72.18	61.75
-2	91.05	69.84	63.29	52.16
-5	85.31	58.77	51.44	40.02

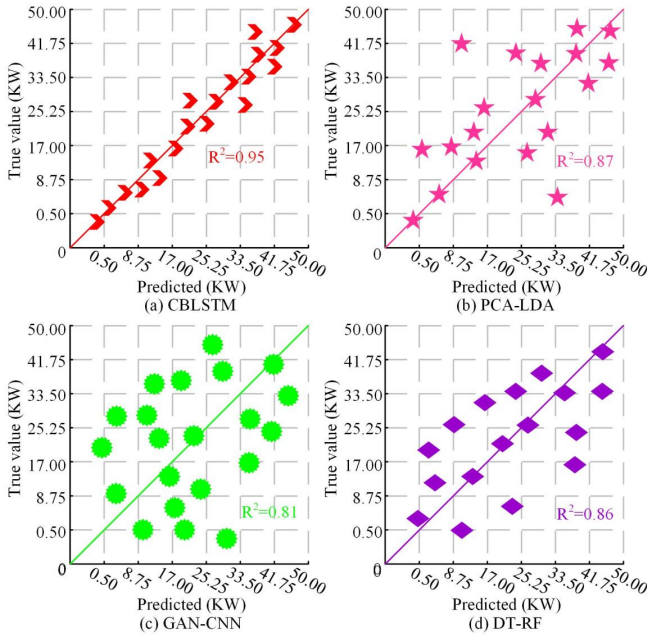

Figure 11. Comparison of Model Accuracy Across Different Motor Power Levels

As shown in Figure 11, the PV-CBLSTM model achieves the highest diagnostic accuracy across all motor power levels, demonstrating superior generalization. Particularly in the low-power (0.50kW-17.00kW) and high-power (33.50kW-50k.00W) ranges, it still achieves accuracy above 95%, significantly outperforming the comparison models. In contrast, the performance of Transformer and TCN decreases noticeably at power extremes, indicat-

Table 5
Test Results of Multiple Comprehensive Indicators

Dataset	Model	FLOPS/MACs (G)	GPU memory consumption (MB)	Lead time (day)	Single inference time (ms)
Training set	CBLSTM	12.5	1450	7	11.2
	Informer	89.5	1089	1	15.7
	GAN-CNN	108.9	1494	2	18.9
	DT-RF	73.4	962	3	20.3
Validation set	CBLSTM	13.4	1250	6	12.7
	Informer	103.4	1167	2	17.7
	GAN-CNN	113.4	1678	1	20.2
	DT-RF	80.1	1047	3	23.4

ing their higher sensitivity to variations in operational conditions. Finally, to further evaluate the feasibility of the model in actual deployment, this paper introduces computing efficiency (FLOPS/MACs) and resource utilization (GPU memory consumption). Lead time for a specific motor-fault diagnosis was tested, and the results are shown in Table 5.

In Table 5, the CBLSTM model achieved a computational cost of 12.5 GFLOPS/MACs and GPU memory usage of 1450 MB, while its fault prediction lead time can reach up to 7 days. In contrast, the GAN-CNN model exhibited significantly higher computational overhead, with 108.9 G FLOPS/MACs and 1494 MB of GPU memory consumption, reflecting its inefficiency in resource utilization. On the validation set, the CBLSTM model maintained a competitive profile with 13.4 G FLOPS/MACs and 1250 MB GPU memory usage, both of which were superior to those of the comparison models. The single inference times for CBLSTM on the training and validation sets are 11.2 and 12.7 seconds, respectively, which are much lower than those of the comparison models. In summary, the test results showed that the CBLSTM model provided more accurate and comprehensive performance in fault diagnosis. This was attributed to the integration of the global optimization capability of improved PSO, the adaptive signal decomposition advantage of improved VMD, and the strengths of multiple algorithms, which enabled superior feature extraction and noise suppression, resulting in stronger overall performance.

5. Conclusions and Recommendations

This study proposed a novel PV-CBLSTM model to address the challenges of accurate and early motor fault diagnosis under noisy industrial conditions. The core contributions of this work are threefold: First, this study introduced a deeply coupled BLSTM architecture in which the BRB serves as an active, tunable input processor rather than a passive component, enabling evidence-driven feature prioritization. Second, an enhanced variational mode decomposition (VMD) technique, optimized by a Bayesian-improved PSO, was integrated to suppress noise interference while preserving critical fault signatures effectively. Third, a comprehensive evaluation demonstrated that the proposed model significantly outperforms several strong baselines, including Informer and GAN-CNN, not only in diagnostic accuracy, recall, and F1-score but also in key engineering metrics such as false alarm rate and single-inference time (15.8 ms), confirming its potential for real-time deployment. Despite its promising performance, this study has limitations that open avenues for future research. The current model was validated primarily on bearing and rotor faults. Future work will focus on. Generalization to a broader range of motor types and fault modes, including complex compound faults. Exploring multi-modal data fusion by integrating current, thermal, and acoustic signals with vibration data to form a more comprehensive health assessment system and developing an online learning mechanism to allow the model to adapt continuously to evolving motor conditions and new fault patterns without full retraining, which is crucial for lifelong operation in industrial settings. These directions will further bridge the gap between academic research and practical industrial application.

References

[1] K. Y. Y. Keshun, W. P. Puzhou, and G. Y. 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.

[2] M. Zabin, H. J. Choi, and J. Uddin, "Hybrid deep transfer learning architecture for industrial fault diagnosis using Hilbert transform and DCNN–LSTM," *The Journal of Supercomputing*, vol. 79, no. 5, pp. 5181–5200, 2023.

[3] L. Wei, X. Peng, and Y. Cao, "Enhanced fault diagnosis of rolling bearings using an improved inception-LSTM network," *Nondestructive Testing and Evaluation*, vol. 40, no. 7, pp. 3274–3293, 2025.

[4] W. Guo, "Chain fault identification and power grid planning optimisation in power systems considering multiple scenarios," *International Journal of Power and Energy Systems*, vol. 44, no. 10, 2024.

[5] R. Cai, B. Xia, X. M. Zhu, L. Wang, J. R. Gu, and J. G. Tang, "Design of a risk model and analytical decision information system for power operation in the context of smart grid," *International Journal of Power and Energy Systems*, vol. 44, no. 10, 2024.

[6] C. Ren, H. Li, and J. Lei, "A CNN-LSTM–based model to fault diagnosis for CPR1000," *Nuclear Technology*, vol. 209, no. 9, pp. 1365–1372, 2023.

[7] M. Huang and C. Sheng, "Adaptive-conditional loss and correction module enhanced informer network for long-tailed fault

diagnosis of motor," *Journal of Computational Design and Engineering*, vol. 11, no. 4, pp. 306–318, 2024.

[8] D. Durairaj, T. K. Venkatasamy, and A. Mehbodniya, "Intrusion detection and mitigation of attacks in microgrid using enhanced deep belief network," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 46, no. 1, pp. 1519–1541, 2024.

[9] M. E. Akiner and M. Ghasri, "Comparative assessment of deep belief network and hybrid adaptive neuro-fuzzy inference system model based on a meta-heuristic optimization algorithm for precise predictions of the potential evapotranspiration," *Environmental Science and Pollution Research*, vol. 31, no. 30, pp. 42719–42749, 2024.

[10] Y. Kong, Z. Wang, and Y. Nie, "Unlocking the power of LSTM for long term time series forecasting," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, pp. 11968–11976, 2025.

[11] S. Lu, J. Yang, and B. Yang, "Surgical instrument posture estimation and tracking based on LSTM," *ICT Express*, vol. 10, no. 3, pp. 465–471, 2024.

[12] H. Zhang, H. Sun, and L. Kang, "Prediction of health level of multiform lithium sulfur batteries based on incremental capacity analysis and an improved LSTM," *Protection and Control of Modern Power Systems*, vol. 9, no. 2, pp. 21–31, 2024.

[13] N. Kavitha, K. R. Soundar, and R. Karthick, "Automatic video captioning using tree hierarchical deep convolutional neural network and ASRNN-bi-directional LSTM," *Computing*, vol. 106, no. 11, pp. 3691–3709, 2024.

[14] R. Jigyasu, V. Shrivastava, and S. Singh, "Deep optimal feature extraction and selection-based motor fault diagnosis using vibration," *Electrical Engineering*, vol. 106, no. 5, pp. 6339–6358, 2024.

[15] J. Guo, Q. He, and F. Gu, "DNOCNet: A novel end-to-end network for induction motor drive systems fault diagnosis under speed fluctuation condition," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 6, pp. 8284–8293, 2024.

[16] S. H. Hsu, C. H. Lee, and W. F. Wu, "Machine learning-based online multi-fault diagnosis for IMS using optimization techniques with stator electrical and vibration data," *IEEE Transactions on Energy Conversion*, vol. 39, no. 4, pp. 2412–2424, 2024.

[17] S. Rokocakau, J. Riccio, and G. Tresca, "Fault diagnosis using shallow neural networks for voltage source inverters in motor drives," *IEEE Transactions on Industry Applications*, vol. 60, no. 5, pp. 7038–7047, 2024.

[18] M. Zhang, Z. Zhang, and X. Wang, "The use of attention-enhanced CNN-LSTM models for multi-indicator and time-series predictions of surface water quality," *Water Resources Management*, vol. 38, no. 15, pp. 6103–6119, 2024.

[19] Y. Cheng, H. Gu, and D. Song, "Fault diagnosis for driving motor with insufficient fault data: a data transfer generation method," *The International Journal of Advanced Manufacturing Technology*, vol. 134, no. 3, pp. 1195–1218, 2024.

[20] Y. Jiang, L. Zheng, and C. Tang, "LSTM-based node-gated graph neural network for cross-condition few-shot bearing fault diagnosis," *IEEE Sensors Journal*, vol. 24, no. 3, pp. 3445–3456, 2024.

[21] G. Fu, Q. Wei, and Y. Yang, "Bearing fault diagnosis with parallel CNN and LSTM," *Mathematical Biosciences and Engineering*, vol. 21, no. 2, pp. 2385–2406, 2024.

[22] Q. Zhou and J. Tang, "An interpretable parallel spatial CNN-LSTM architecture for fault diagnosis in rotating machinery," *IEEE Internet of Things Journal*, vol. 11, no. 19, pp. 31730–31744, 2024.

Biographies

Feiyu Liu received the Master's degree in Control Engineering from Jiangsu University in 2016. Currently, he serves as a lecturer and department head in the Department of Electrical Engineering at School of Engineering, Nanjing Normal University Zhongbei College. He has published over 10 academic papers. His areas of interest include digital twins, power systems, object detection, and image processing.