

Article

IMU-Based Time-Domain Fault Diagnosis of BLDC Motors Using an End-to-End 1D-CNN

Ke Hao Wang, Hwi Gyu Lee, Seon Min Yoo  and In Soo Lee *

School of Electronic and Electrical Engineering, Kyungpook National University, Daegu 41566, Republic of Korea; ws20000726@knu.ac.kr (K.H.W.); bbi05179@gmail.com (H.G.L.); seon01@knu.ac.kr (S.M.Y.)

* Correspondence: insoolee@knu.ac.kr; Tel.: +82-10-5312-5324

Abstract

Reliable fault detection in brushless DC motors is challenging owing to environmental complexity and high equipment costs. To address these challenges, we propose an effective and cost-effective approach using an optimized end-to-end one-dimensional convolutional neural network. Specifically, a real experimental platform simulating bearing and eccentricity faults was developed. Statistical *t*-tests indicated that three-axis accelerometer signals from a low-cost inertial measurement unit provided sufficient fault information for the present diagnosis task. Unlike traditional methods such as support vector machines, multilayer neural networks, and random forests, which rely on manual feature extraction, our model learns directly from raw waveforms and can handle signal drift. Under the present controlled experimental setting and the leave-one-day-out evaluation protocol, the model achieved 100.00% average window-level classification accuracy, considerably outperforming traditional methods, the performances of which declined to 67.95–71.37% under environmental shifts. Moreover, with an inference time of only 0.96 ms, 32 times faster than that of random forests, this approach is well suited for real-time embedded monitoring. The proposed method demonstrates strong potential for cost-efficient and robust fault diagnosis under the present experimental setting.

Keywords: Brushless DC motor; fault diagnosis; 1D convolutional neural network; inertial measurement unit sensor; time-domain analysis; vibration monitoring

1. Introduction

Brushless DC (BLDC) motors, owing to their high power density, efficiency, and low maintenance requirements, are widely used in the fields of industrial automation, aerospace, and electric vehicles [1,2]. Although these motors are generally reliable, prolonged operation in harsh environments inevitably causes component aging and wear. Studies indicate that bearing failures and rotor eccentricity account for over 50% of all known motor fault types [3,4]. If such faults are not detected promptly, they can progress and cause severe system damage or even a complete shutdown, resulting in significant economic losses. Therefore, developing intelligent systems for motor fault diagnosis is essential.

Traditionally, vibration analysis has been the primary approach for motor condition monitoring. Common techniques include fast Fourier transform [5], wavelet transform (WT) [6], and empirical mode decomposition [7]. These methods typically extract frequency-domain features before conducting fault diagnosis. However, these methods are highly dependent on expert knowledge and require manual feature design and extraction. Furthermore, detecting high-frequency vibration signals caused by faults requires considerably



Academic Editors: Margarita Terziyska and Miroslava Ivanova

Received: 19 March 2026

Revised: 28 April 2026

Accepted: 30 April 2026

Published: 2 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

high sampling rates, often exceeding 10 kHz, to satisfy the Nyquist–Shannon theorem for capturing fault harmonics [8]. Although the affordability of sensors has increased, acquiring, processing, and transmitting such high-frequency data imposes substantial demands on communication bandwidth, storage capacity, and computational resources. This significantly increases system complexity and cost and limits the practicality of employing these methods in low-power, resource-constrained monitoring systems.

In recent years, deep learning has considerably advanced fault diagnosis by automatically learning relevant features. Convolutional neural networks (CNNs), in particular, have demonstrated strong diagnostic performance [9,10]. Recent studies published in high-impact international journals have further advanced intelligent techniques for fault diagnosis. These studies can be broadly categorized into two main types.

First, lightweight one-dimensional deep learning models have garnered significant attention owing to their efficiency and suitability in real-time applications. These methods directly process raw time-series signals and have demonstrated promising performance in terms of reducing model size while maintaining accuracy. For instance, Tao et al. [11] developed a lightweight one-dimensional CNN (1D-CNN) that considerably reduced the number of network parameters while preserving diagnostic performance, and similar approaches were explored in [12,13].

Second, time–frequency representation-based methods have been widely explored, in which raw vibration signals are transformed into two-dimensional representations, such as short-time Fourier-transform spectrograms or wavelet scalograms, and subsequently processed using two-dimensional CNNs [14,15]. Although these approaches afford high diagnostic accuracy, they typically involve complex preprocessing steps and incur high computational and memory costs, limiting the applicability of these approaches in real-time embedded systems.

Despite the promising performance of existing methods, significant challenges remain in terms of real-world industrial applications. In particular, several studies have assumed that the training and testing data follow the same distribution (i.i.d. assumption) [16]. In practice, however, factors such as sensor reinstallation errors, environmental temperature variations, and gradual loosening of mechanical structures can generate baseline drift in vibration signals over time [17]. This phenomenon, known as domain shift, can cause models that perform well on historical data to suffer considerable accuracy degradation when applied to unseen conditions [18,19]. In addition, several existing approaches rely on high-fidelity and high-cost sensors, which limit their scalability in practical industrial scenarios. Recent studies have further broadened the application scope of intelligent diagnosis of motor faults from conventional single-condition classification to practical settings involving multisensor fusion, domain shift, and complex drive systems. For example, recent representative studies have highlighted fault diagnosis and fault-tolerant operations as critical challenges in advanced motor-drive systems [20]. In parallel, multi-sensor fusion approaches that integrate complementary signal sources have been explored to improve diagnostic performance [21]. In addition, domain-robust and domain-adaptive diagnosis methods have been proposed to address performance degradation under distribution mismatch between training and test conditions [22]. These studies offer an important additional context to the present study.

In contrast to multi-sensor fusion frameworks that rely on rich sensing configurations, the present study focuses on a low-cost inertial measurement unit (IMU)-based setup and investigates whether reliable fault diagnosis can be achieved using only three-axis acceleration signals. Moreover, rather than introducing an explicit domain-adaptation module, this study evaluates robustness through a strict cross-day leave-one-day-out

(LODO) protocol, which directly determines the stability of the learned representation under practical conditions of baseline drift.

New feature extraction strategies are urgently required to further improve the feasibility of deploying deep learning models on resource-constrained industrial edge devices and to enhance robustness against environmental baseline drift. This study shifts focus to widely available, low-cost microelectromechanical systems-based IMU sensors, emphasizing waveform morphology, rather than traditional spectral peaks [23–25]. Specifically, this study employs an end-to-end wide-convolution-kernel 1D-CNN model to directly extract robust features from IMU acceleration signals. The main contributions of this study are as follows:

- Experimental platform and validation protocol: A custom test platform for the BLDC motor was developed, and the LODO cross-day validation method was used to simulate sensor baseline drift commonly encountered in industrial environments.
- Sensor optimization: Statistical *t*-test analysis [26] demonstrated that gyroscope data is redundant. Using only three-axis acceleration signals was sufficient for fault detection, thereby reducing hardware costs.
- Robustness to environmental changes: Experiments revealed that the proposed wide-convolution-kernel 1D-CNN effectively learns waveform shape features from raw signals, achieving 100% accuracy. In contrast, the performances of traditional support vector machine (SVM) and multilayer neural network (MLNN) models substantially degraded owing to feature drift [27].
- Additional cross-speed validation: Additional cross-speed validation was conducted to examine the robustness of the proposed method under variable rotational-speed conditions.

2. Experimental System Description

2.1. Electromechanical Testbed Setup

Experimental data were collected using a custom-built motor fault simulation rig. Figure 1 depicts that the system includes a PG42-BL4261 BLDC motor (Twirl Motor, Changzhou, China; 24 V, 1/61 reduction ratio) as the target machine.

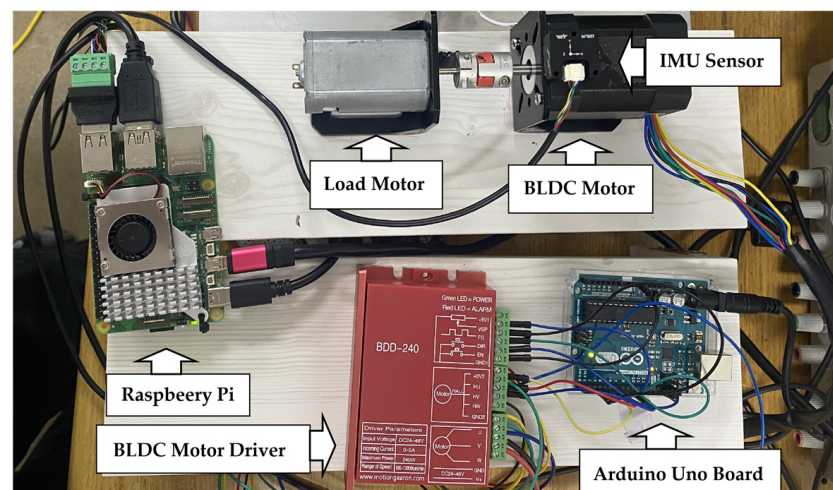


Figure 1. Experimental equipment.

- Control: An Arduino microcontroller generates pulse-width modulation signals to maintain a constant rotational speed of 1200 RPM.
- Load: The target motor is coupled to a secondary permanent magnet motor via a rigid coupling to provide a consistent mechanical load.

- Data acquisition: A low-cost six-axis IMU (MPU-605) is mounted on the motor housing near the output shaft. Vibration data are collected by a Raspberry Pi via the I2C interface at a sampling rate of 126 Hz.

2.2. Fault Injection Methodology

To construct a reliable dataset, three motor states are artificially generated, as illustrated in Figure 2.

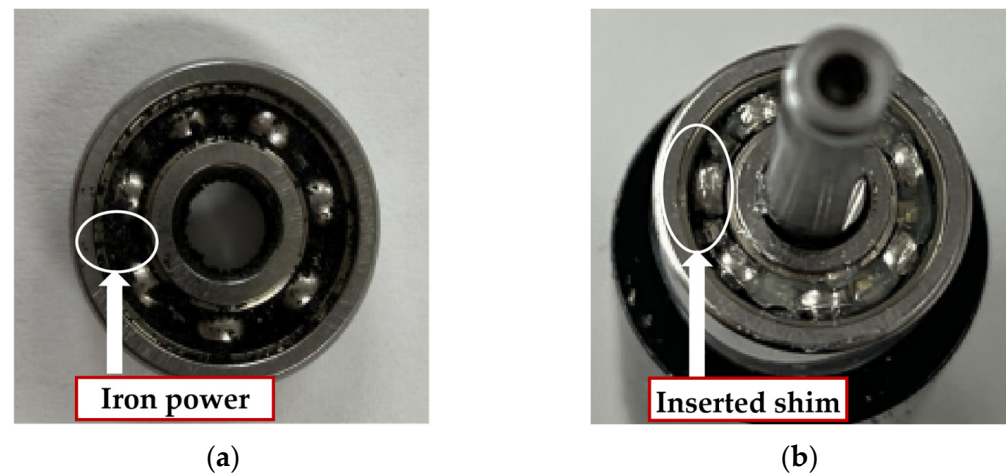


Figure 2. Bearing conditions: (a) bearing fault and (b) eccentricity fault.

- Normal state: The motor operates under healthy conditions with all components properly aligned and no faults present.
- Bearing fault: Iron powder is added to the bearing lubricant to simulate contamination-induced bearing degradation. The powder disrupts the lubricating oil film, thereby increasing surface roughness and inducing abnormal vibrations. This setting is employed as a controlled laboratory emulation of contamination-induced bearing degradation, rather than as an exact reproduction of all naturally evolving bearing failures.
- Eccentricity fault (EF): A spacer is inserted between the bearing housing and the motor shaft, shifting the rotation center of the shaft from its original geometric center. This misalignment induces dynamic eccentricity, causing considerable low-frequency vibration modulation. The spacer thickness considered in this study is 0.5 mm. This configuration is intended to emulate a repeatable eccentric rotating condition caused by geometric offset or misalignment under controlled experimental conditions.

Notably, the fault conditions considered in this study correspond to controlled and detectable fault states established on a laboratory testbed. Despite reflecting physically meaningful degradation mechanisms, these settings fail to cover the full severity progression of naturally evolving faults, especially incipient faults with considerably weak signatures. Therefore, evaluating fault-severity levels and early-stage fault diagnosis is an important direction for future work.

3. Proposed Fault Diagnosis Framework

This section introduces the efficient and reliable diagnosis system for BLDC motor faults proposed in this study. The overall procedure is illustrated in Figure 3. The framework consists of three main stages, namely, sensor optimization, data processing, and model development. The first stage is sensor optimization (Section 3.1). Specifically, a *t*-test is applied to identify the most informative signals from the six-axis IMU data. This process reduces system complexity and data transmission requirements by selecting the most

relevant signals and reducing the original six channels to three channels. The second stage is data processing (Section 3.3). Specifically, a sliding window method is employed to segment continuous vibration signals into individual samples. Subsequently, a strict Z-score normalization procedure is applied to standardize the data while preventing information leakage between the training and testing sets. The final stage involves model development.

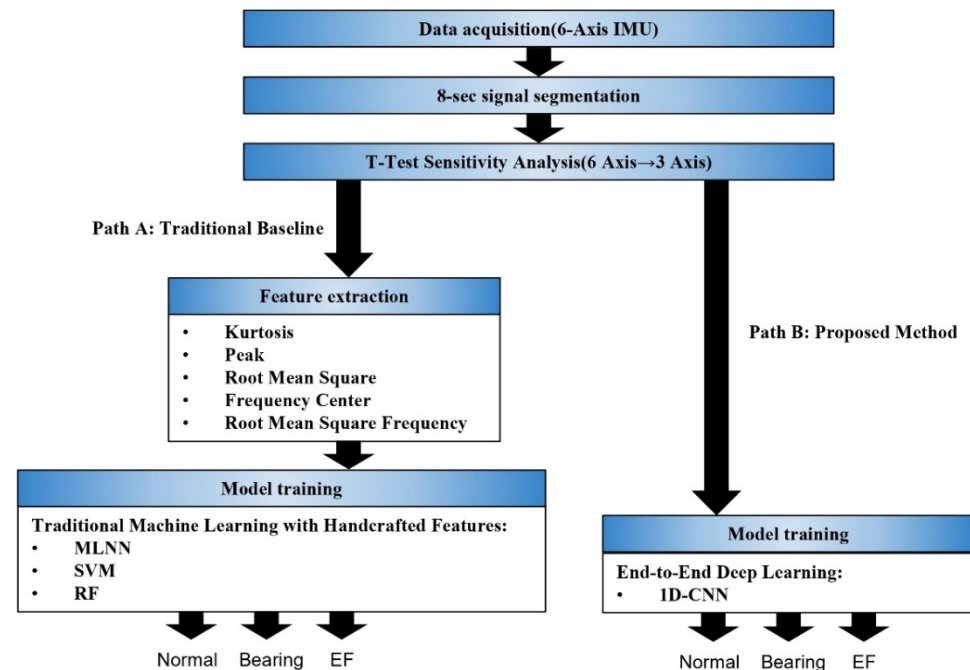


Figure 3. Proposed fault diagnosis procedure.

Two diagnostic routes are designed for comparison to ensure a fair evaluation of the proposed approach. Route A represents the traditional baseline methods. Section 3.2 describes that the statistical features are first manually extracted from the vibration signals and subsequently used as inputs to traditional classifiers such as SVMs, MLNNs, and random forests (RFs). Route B represents the method proposed in this study. Section 3.3 describes that an optimized end-to-end 1D-CNN is directly applied to the raw time-domain waveforms. This model automatically learns discriminative features without requiring manual feature engineering and provides highly robust fault diagnosis under conditions of signal drift. The overall framework representing the proposed method is illustrated in Figure 3.

3.1. Sensor Selection Strategy

The initial data collection included six channels consisting of a three-axis accelerometer and a three-axis gyroscope. To improve computational efficiency, an independent two-sample *t*-test was conducted to evaluate the discriminative capability of each channel. Moreover, the visual ranking of feature sensitivity is presented in Figure 4, indicating that gyroscope features generally exhibit low statistical significance for stationary fault classification, i.e., they provide limited discriminative information. In contrast, accelerometer signals demonstrate strong separability across different fault conditions.

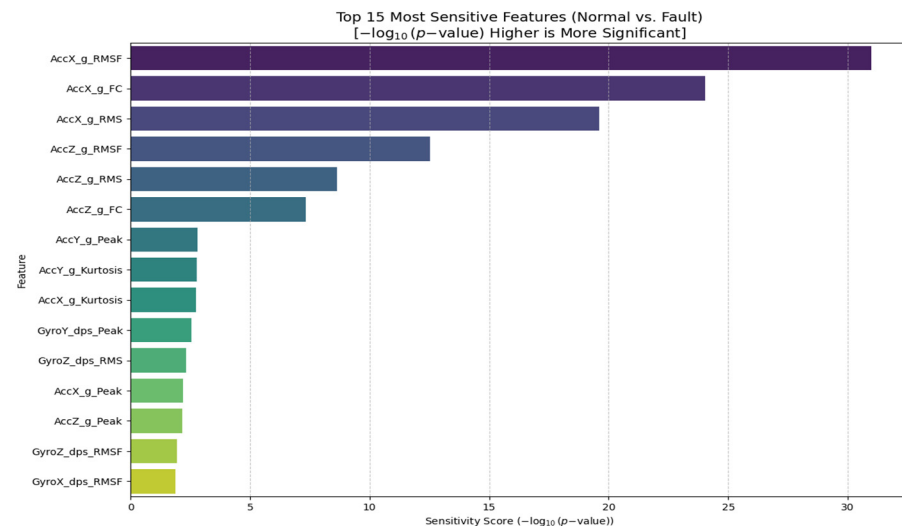


Figure 4. Sensitivity analysis using the independent *t*-test.

The detailed statistical metrics of the representative features are listed in Table 1 to further quantify this observation. The data in this table reveal that the Z-axis gyroscope produced a *p*-value of 0.9094, indicating that no meaningful fault-related patterns were present; therefore, the useful diagnostic information obtained was limited.

Table 1. Statistical *t*-test results of representative features across different sensor channels.

Sensor Channel	Feature Name	T-Statistic	<i>p</i> -Value	Decision
<i>Acc_x</i> (dominant)	Root mean square (RMS) frequency	13.90	1.01×10^{-31}	Retain
	Frequency center	11.62	9.34×10^{-25}	Retain
<i>Acc_z</i> (strong)	RMS frequency	8.57	3.02×10^{-13}	Retain
<i>Acc_y</i> (valid)	Kurtosis	3.25	0.0017	Retain
	Peak value	3.24	0.0016	Retain
<i>Gyro</i> (noise)	<i>Gyro_x</i> peak	0.11	0.9094	Discard
	<i>Gyro_x</i> RMS	0.16	0.8769	Discard

Crucially, *Acc_x* and *Acc_z* exhibited the most dominant fault signatures, with T-statistics greater than 13; additionally, *Acc_y* demonstrated clear statistical significance. Important indicators, such as kurtosis ($T = 3.25$, $p = 0.0017$) remain well below the significance threshold of 0.01, although it appears less prominent in Figure 4. These results indicate that the Y-axis contains meaningful fault-related information and contributes to preserving the spatial characteristics of the vibration data. Consequently, all three accelerometer channels (*Acc_x*, *Acc_y*, and *Acc_z*) were retained, and the redundant gyroscope channels were removed to reduce input dimensionality.

Although the statistical *t*-test is an effective preliminary screening method, relying solely on linear statistical measures may cause complex nonlinear relationships within the data to be overlooked. Therefore, in this study, the *t*-test was employed as an interpretable first-pass sensor screening tool, rather than as the sole basis for the final sensor selection. To further validate the sensor selection, a comprehensive ablation study was conducted using the 1D-CNN model to compare the performance of the six- and three-axis inputs. The results presented in Section 4.4 further support employing the three-axis accelerometer configuration under the present dataset and protocol. These findings support the decision

to employ only the three-axis accelerometer signals as the input to the proposed framework for fault diagnosis.

3.2. Feature Engineering and Baseline Model Settings

This study considers SVM, MLNN, and RF as comparative baseline models to objectively evaluate the performance of the proposed 1D-CNN model. Unlike the 1D-CNN, which directly processes raw waveform signals, these traditional machine learning models require manual feature extraction to operate effectively. Feeding raw high-dimensional time-series data directly into shallow models often results in the curse of dimensionality and poor model convergence.

Different input strategies were adopted for the deep learning model and the traditional machine learning models to ensure a fair and meaningful comparison.

- **Proposed 1D-CNN:** The proposed 1D-CNN was trained using raw time-domain waveforms obtained from the three accelerometer channels (*Acc x/y/z*). This design allows the deep learning model to automatically learn shift-invariant morphological features that may be more informative than handcrafted indicators.
- **Baseline models (SVM, RF, and MLNN):** These were trained using a 15-dimensional statistical feature vector derived from the same three accelerometer channels, consistent with the sensor optimization results discussed in Section 3.1.

Specifically, for each sliding window containing 1024 signal samples, five statistical indicators were calculated for each of the three channels, yielding a total of 15 features (3 channels $\times$ 5 indicators). These features include both time- and frequency-domain characteristics commonly used in vibration analysis, such as root mean square (RMS), peak value, and kurtosis. The mathematical definitions of these features are summarized in Table 2. This design ensures that any observed performance differences are primarily attributable to the model architecture, rather than inconsistencies in the input data.

Table 2. Mathematical definitions of the extracted statistical and spectral features.

Category	Feature Name	Formula
Time domain	RMS	$X_{\text{rms}} = \sqrt{\frac{1}{L} \sum_{i=1}^L x_i^2}$
	Peak value (peak)	$X_{\text{peak}} = \max(x_i)$
	Kurtosis	$K = \frac{\frac{1}{L} \sum_{i=1}^L (x_i - \bar{x})^4}{\left(\frac{1}{L} \sum_{i=1}^L (x_i - \bar{x})^2\right)^2}$
Frequency domain	Frequency center (FC)	$f_c = \frac{\sum_{j=1}^K f_j \cdot S(f_j)}{\sum_{j=1}^K S(f_j)}$
	RMS frequency (RMSF)	$f_{\text{rms}} = \sqrt{\frac{\sum_{j=1}^K f_j^2 \cdot S(f_j)}{\sum_{j=1}^K S(f_j)}}$

Three baseline classifiers were implemented for comparison using these feature vectors as inputs. The SVM was configured with a radial basis function kernel to capture nonlinear decision boundaries. In addition, an MLNN with a fully connected architecture was constructed using the 15-dimensional handcrafted feature vector as input. Additionally, an RF classifier consisting of 100 decision trees was implemented owing to its robustness to noise and its ability to handle high-dimensional feature spaces through ensemble voting mechanisms [28]. A 15-dimensional feature vector was constructed for each time window by concatenating the statistical and spectral indicators extracted from all three sensor channels. These vectors serve as the input features for the baseline models, i.e., SVM, MLNN,

and RF, enabling a systematic evaluation of traditional feature engineering approaches in comparison with the proposed end-to-end deep learning method.

3.3. Data Processing and Proposed 1D-CNN Architecture

The proposed framework for fault diagnosis integrates rigorous data preprocessing with an optimized deep learning model. Each operating condition was recorded for 60 s per day, yielding nine independent vibration recordings in total (3 days $\times$ 3 categories) with an overall duration of 540 s. The sampling frequency of 126 Hz ensured that each recording contained 7560 sampling points. In this study, the three acquisition days (D1, D2, and D3) were treated as independent domains in the LODO evaluation protocol.

First, the continuous time-series data were segmented using a sliding window approach with a window length of 1024 sampling points. Each segment, therefore, corresponds to approximately 8.13 s. At the motor operating speed of 1200 RPM (20 Hz), this time window captured approximately 325 full mechanical rotations. This duration provided sufficient temporal context to preserve low-frequency periodic fault signatures that may not have been captured by short windows. The number of segments generated from each recording is expressed as

$$N_{seg} = \left\lceil \frac{N - L}{S} \right\rceil + 1, \quad (1)$$

where N is the number of sampling points, W is the window length, and S is the stride.

A stride of 256 sampling points was used, causing a 75% overlap between adjacent segments. Therefore, each 60 s recording produced 26 overlapping segments, and a total of 234 segments were obtained with a balanced distribution across the three fault categories, corresponding to 78 segments per class. The data acquisition conditions and window segmentation settings are summarized in Table 3.

Table 3. Dataset composition and window distribution under the LODO protocol.

Acquisition Day	Health State	Recording Duration/Sampling Rate	Window Setting	Number of Windows
D1	Normal	60 s/126 Hz	1024/256 (75% overlap)	26
D2	Bearing fault			
D3	Eccentricity fault			

Despite overlapping segmentation increasing local similarity between adjacent windows from the same recording, no windows from the test day were included in the training set under the LODO protocol. Therefore, the present windowing strategy did not introduce direct train–test leakage.

Subsequently, Z-score normalization was applied to all segments. The mean and standard deviation were calculated exclusively from the training data to strictly prevent information leakage during inference on unseen target domains.

After preprocessing, the normalized segments are input to the proposed end-to-end 1D-CNN, as illustrated in Figure 5. The network was designed to process the 1024×3 raw time-domain waveforms directly without manual feature engineering. The architecture first contained an input layer that accepts multivariate time-series signals, followed by three stacked convolutional blocks containing 32, 64, and 128 filters, respectively. Each convolution layer employed a relatively wide kernel size of nine to effectively capture the morphological characteristics of vibration signals collected at a moderate sampling rate.

The detailed hyperparameters and structural configuration of the neural network are listed in Table 4.

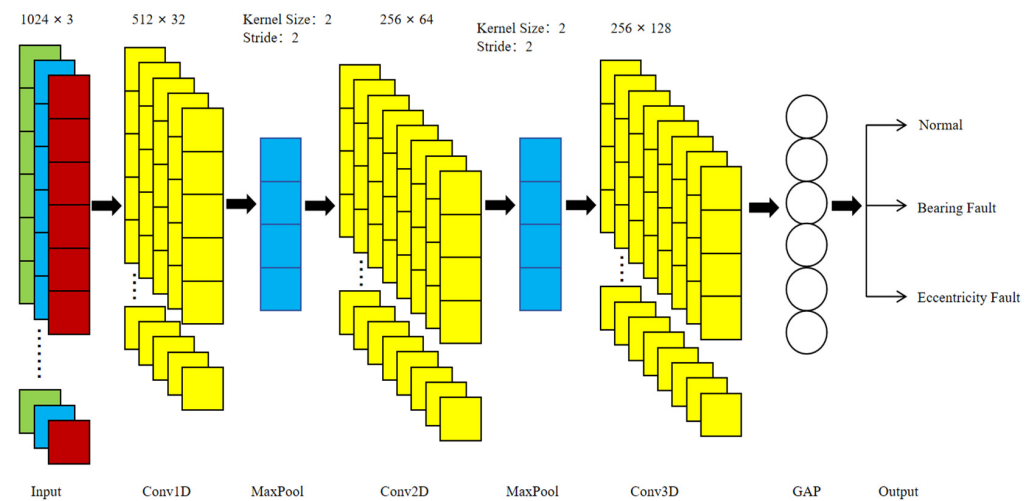


Figure 5. Architecture of the proposed three-axis optimized 1D-CNN. The arrows indicate the forward data flow, and the colors are used only to distinguish different network components.

Table 4. Detailed parameters of the proposed 1D-CNN.

Layer Type	Output	Kernel	Stride	Filters
Input	(3, 1024)	-	-	-
Conv1D + BN + ReLU	(32, 1024)	9	1	32
MaxPool1D	(32, 512)	2	2	-
Conv1D + BN + ReLU	(64, 512)	9	1	64
MaxPool1D	(64, 256)	2	2	-
Conv1D + BN + ReLU	(128, 256)	9	1	128
Global AvgPool	(128)	-	-	-

The process of core feature extraction is implemented through the one-dimensional convolution operation. Mathematically, the output of the k -th filter in the l -th layer, denoted as y_k^l , is obtained by convolving the input feature map x_i^{l-1} with a learnable kernel w_k^l and adding a bias b_k^l :

$$y_k^l(j) = \sum_{m=1}^{C_{in}} \sum_{i=0}^{K-1} w_k^l(m, i) x_i^{l-1}(j+m) + b_k^l, \quad (2)$$

where C_{in} represents the number of input channels and K denotes the kernel size. To mitigate internal covariate shift and accelerate training convergence, Batch normalization (BN) [29] is applied immediately after each convolution layer. Subsequently, the rectified linear unit activation function is used to introduce nonlinearity into the model. Unlike traditional CNN architectures that rely on fully connected layers, the proposed network employs global average pooling (GAP) [30] to compress the final feature maps into a compact feature vector. This design considerably reduces the number of trainable parameters while preserving important global information. Finally, a softmax classifier is used to predict the motor health state, which is categorized into normal, bearing fault, or eccentricity fault. The detailed configuration of the CNN architecture is summarized in Table 4.

4. Experimental Results and Discussion

4.1. Signal Analysis

Figure 6 presents the raw time-domain signals recorded over three experimental days. The EF produced a noticeable increase in vibration amplitude in the radial directions, whereas the bearing fault generated distinct impulsive patterns. Although slight baseline drift can be observed across different days, the primary fault signatures remained clearly distinguishable.

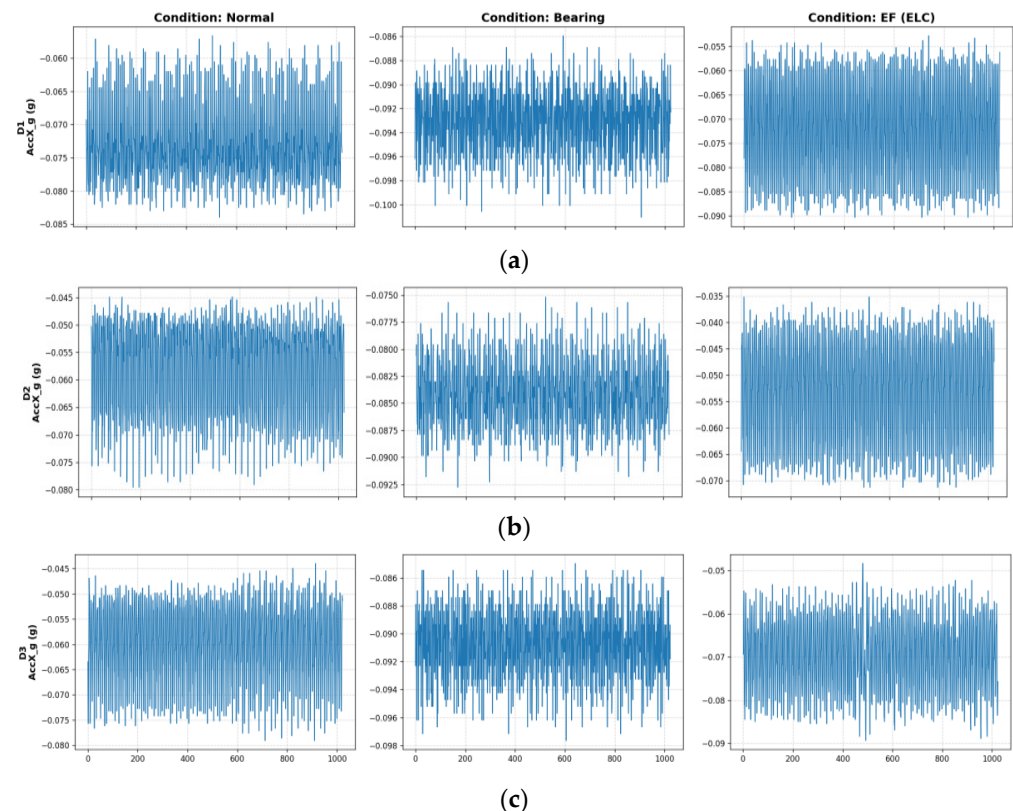


Figure 6. Raw time-domain vibration signals of the AccX channel across different health states and experimental days: (a) D1, (b) D2, and (c) D3. In each row, the signals correspond to the normal, bearing fault, and eccentricity fault conditions.

Next, an LODO cross-validation protocol was adopted to rigorously evaluate the generalization capability of the proposed model under nonstationary environmental conditions [17,26]. Figure 7 depicts that the dataset consists of three independent data collection sessions (D1, D2, and D3). Each session represents a distinct domain in which the sensors were reinstalled and the ambient conditions varied. This three-fold evaluation design provided a statistically meaningful framework for assessing the ability of the model to learn shift-invariant features. Additionally, this design ensures that the diagnostic results depend on the intrinsic physical fault characteristics, rather than on environmental noise or sensor baseline variations.

4.2. Model Training and Convergence

The proposed 1D-CNN model was trained using the Adam optimizer [31]. Figure 8 illustrates the training dynamics across the three validation folds. The model results demonstrate rapid convergence, with the training loss stabilizing after approximately 15 epochs. The close alignment between the curves representing training and validation accuracies indicates stable learning behavior. This observation suggests that the use of BN and GAP effectively mitigates overfitting while maintaining a strong generalization capability.

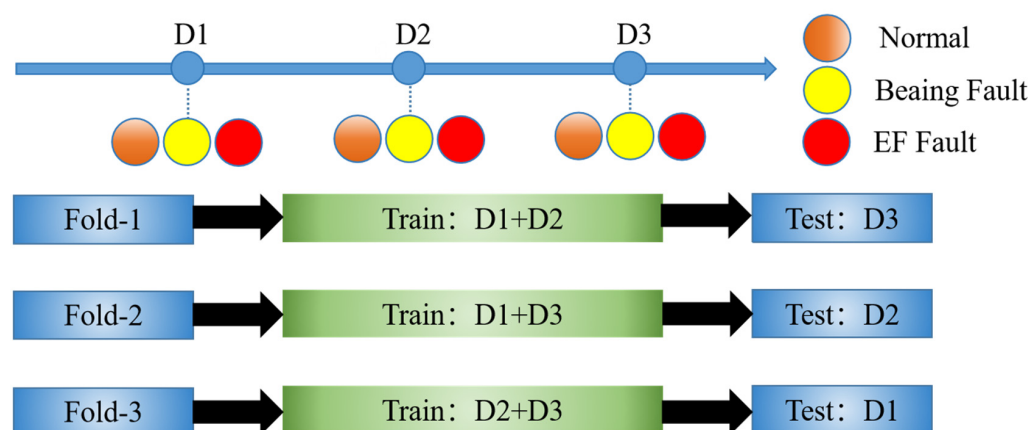


Figure 7. Experimental protocol diagram.

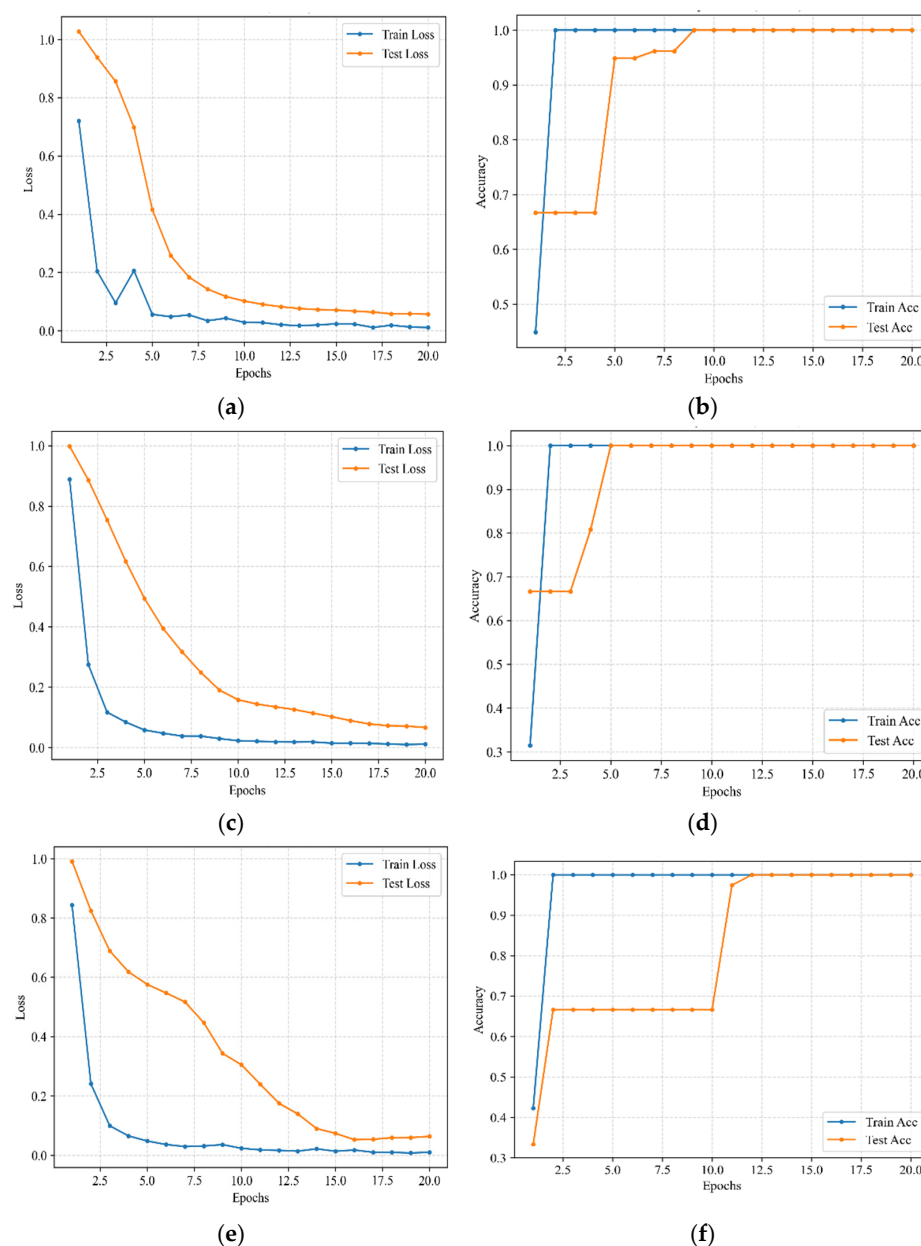


Figure 8. Training and validation curves of the 1D-CNN: (a) Fold 1 loss curve, (b) Fold 1 accuracy curve, (c) Fold 2 loss curve, (d) Fold 2 accuracy curve, (e) Fold 3 loss curve, and (f) Fold 3 accuracy curve.

4.3. Diagnosis Performance

Subsequently, a three-fold LODO cross-validation strategy was implemented to rigorously assess model robustness under day-to-day operating variability. In this protocol, data from two experimental days were used for training, whereas the remaining day was reserved exclusively for testing in each fold. This evaluation strategy ensured that no temporal information from the test day was involved during model training and, therefore, facilitated a realistic assessment of the model's ability to generalize to unseen daily operating conditions. Notably, the use of overlapping windows did not cause train–test leakage in this study because the LODO protocol separated the data at the acquisition-day level, rather than at the window level. The detailed results corresponding to all three folds, along with the overall average window-level performance, are summarized in Table 5.

Table 5. Detailed accuracy comparison of each validation fold.

Model	Fold 1 (Test D3)	Fold 2 (Test D2)	Fold 3 (Test D1)	Average Accuracy	Operating Time (ms)
Random Forest	56.41%	47.44%	100.00%	67.95%	30.74
MLNN	74.36%	33.33%	100.00%	69.23%	0.09
SVM	69.23%	44.87%	100.00%	71.37%	0.18
1D-CNN	100.00%	100.00%	100.00%	100.00%	0.96

To ensure an effective comparison, the average diagnostic accuracy across all models is illustrated in Figure 9. To further analyze the classification performance under nonstationary conditions, the normalized confusion matrices of the three validation folds are presented in Figure 10.

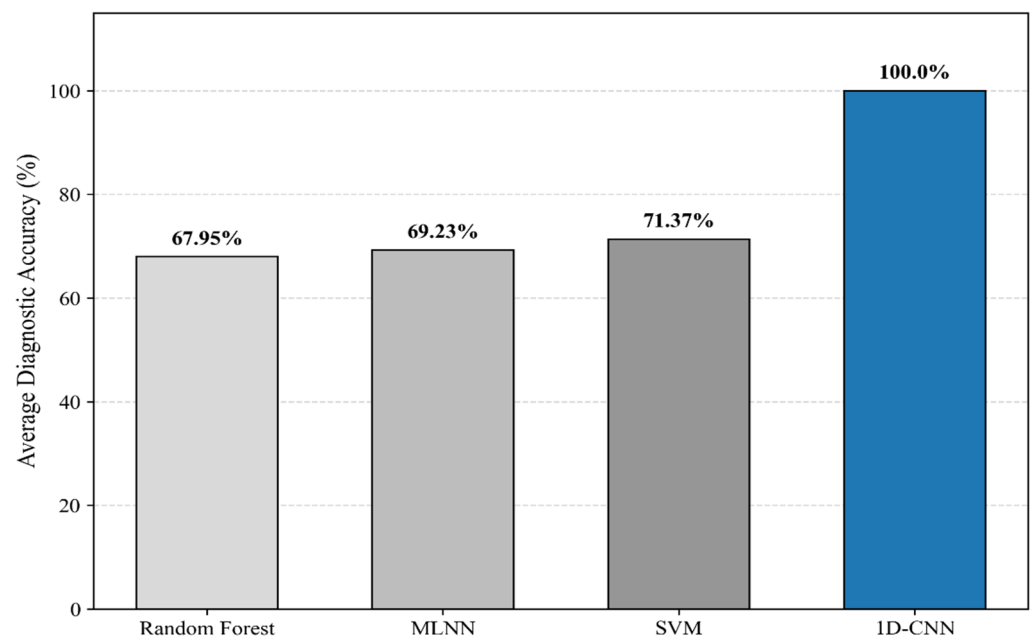


Figure 9. Comparison of average diagnostic accuracies across four models.

Figure 10 reveals that the proposed 1D-CNN achieves highly consistent classification performance across all folds. In Folds 2 and 3, the diagnostic accuracy reached 100%, indicating strong class separability when the training data adequately captured the primary distribution variations.

To further illustrate the separability of the learned feature representations, the t-distributed stochastic neighbor embedding (SNE) technique is used to project the high-dimensional features extracted by the GAP layer into a 2D space, as displayed in Figure 11. This visualization indicates that even under the unseen test domain (D3), samples from the three health conditions formed compact and clearly separated clusters with minimal overlap. This observation suggests that the proposed model learns feature representations that remain stable under the present cross-day evaluation setting, consistent with the strong performance of this model relative to the baseline models.

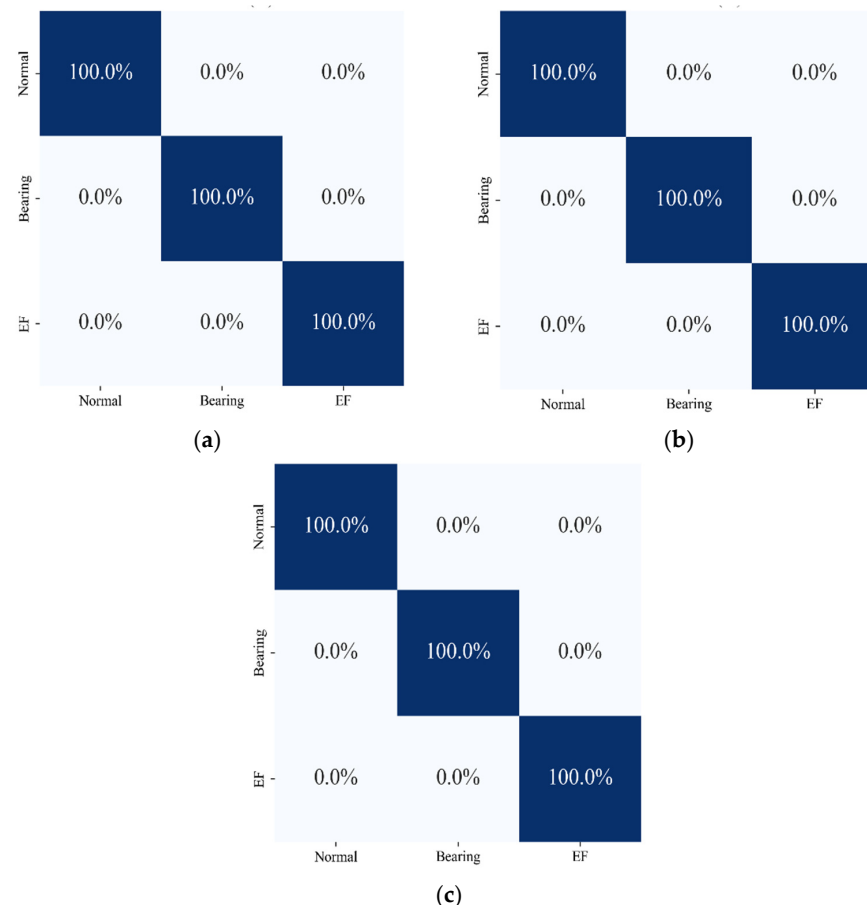


Figure 10. Normalized confusion matrices of the proposed 1D-CNN corresponding to each experimental fold: (a) Fold 1 (tested on D3), (b) Fold 2 (tested on D2), and (c) Fold 3 (tested on D1).

4.4. Ablation Study and Discussion

The experimental results revealed clear differences in robustness among the evaluated models under nonstationary operating conditions. The MLNN experienced severe performance degradation, with accuracy reducing to 33.33% in Fold 2. This substantial decline highlights the sensitivity of handcrafted statistical features, such as RMS, kurtosis, and peak values, to sensor baseline drift caused by reinstallation or daily environmental variations. Even minor shifts in signal offset or amplitude scaling across different days can considerably reduce the effectiveness of these rigid, domain-dependent features. In contrast, the RF model achieved an average accuracy of 67.95%, outperforming both SVM and MLNN. The ensemble learning mechanism of RF reduced variance and improved robustness across folds. However, the model continued to exhibit clear limitations when encountering challenging domain shifts. For example, the accuracy decreased to 56.41% in Fold 1, indicating that feature-based shallow models remain vulnerable to variations in signal distribution. The proposed end-to-end 1D-CNN demonstrated substantially su-

perior performance and robustness. By directly learning shift-invariant morphological patterns from raw time-domain waveforms, the model avoids the limitations associated with handcrafted statistical or frequency-domain features. This capability enables the network to maintain stable performance, even when environmental conditions and sensor baselines vary across data collection sessions. A targeted ablation study was conducted to further validate the sensor selection strategy described in Section 3.1. The analysis compared two input configurations under the same LODO cross-validation protocol, i.e., a three-axis configuration using only accelerometer signals and a six-axis configuration that combines accelerometer and gyroscope signals. The key results are summarized in Table 6. Consistent with the statistical findings obtained from the *t*-test analysis, the inclusion of gyroscope signals in the six-axis configuration introduced additional variations that failed to offer meaningful diagnostic information under domain-shift conditions. This effect was particularly evident in the most challenging validation scenario, Fold 1, in which 15.4% of normal samples were incorrectly classified as eccentricity faults. Therefore, the six-axis model produced a worst-case accuracy of 94.87% in this fold. In contrast, the three-axis configuration eliminated these redundant and potentially noisy inputs and achieved perfect classification performance across all folds. The model maintained an accuracy of 100.00% with zero variance across the validation experiments. Furthermore, reducing the number of input channels from six to three improved computational efficiency. Moreover, the inference latency decreased from 1.28 to 0.96 ms on the experimental platform. These results confirm that reducing sensor dimensionality can simultaneously enhance model robustness and computational efficiency. The three-axis accelerometer configuration, therefore, provides an optimal balance between diagnostic accuracy and resource efficiency, rendering this configuration particularly suitable for real-time industrial monitoring systems deployed on resource-constrained Internet-of-Things devices.

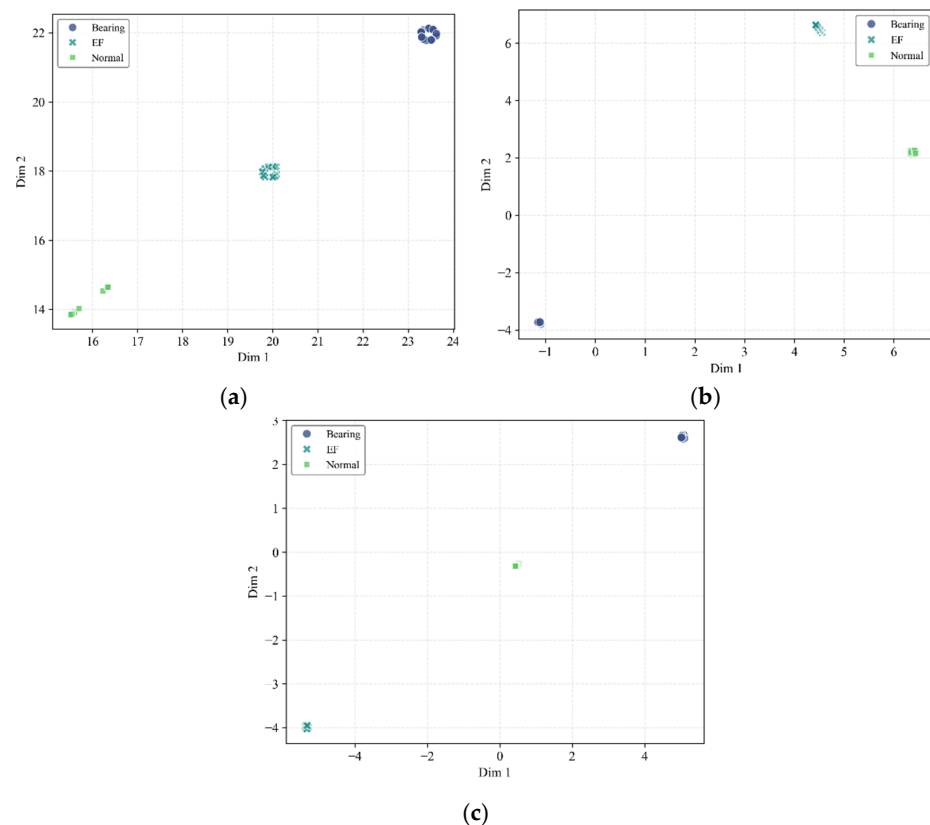


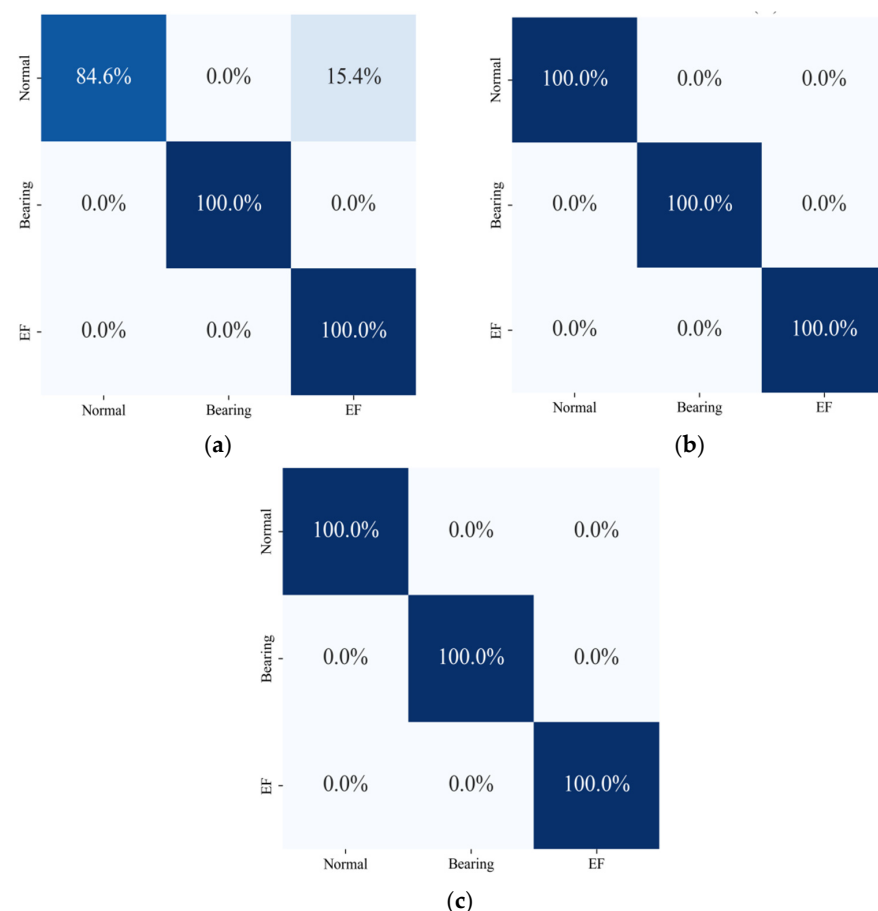
Figure 11. t-SNE visualization of the learned feature representations under LODO cross-validation: (a) Fold 1 (testing on D3), (b) Fold 2 (testing on D2), and (c) Fold 3 (testing on D1).

Table 6. Ablation study and performance comparison of different sensor configurations.

Input Configuration	Average Accuracy	Std. Dev. (σ)	Worst Case (min. Acc)	Inference Time (ms)	Remarks
6-axis (<i>Acc</i> + <i>Gyro</i>)	98.29%	$\pm 2.96\%$	94.87%	1.28	Redundant features caused overfitting in Fold 1.
3-axis (<i>Acc</i> only)	100.00%	$\pm 0.00\%$	100.00%	0.96	Optimal stability and efficiency.

Notably, the fault conditions considered in this study correspond to controlled and detectable fault states established on a laboratory testbed. Despite reflecting physically meaningful degradation mechanisms, these settings do not cover the full severity progression of naturally evolving faults, especially incipient faults with very weak signatures. Therefore, evaluating fault-severity levels and early-stage fault diagnosis will be an important direction for future work.

To provide a detailed interpretation of the six-axis result, the normalized confusion matrix of the six-axis configuration in Fold 1 is additionally presented in Figure 12. The degradation was mainly associated with increased confusion between the normal and eccentricity fault classes, indicating that the additional gyroscope channels introduced variability without improving class separability under cross-day conditions.

**Figure 12.** Normalized confusion matrices of the six-axis 1D-CNN configuration under LODO cross-validation: (a) Fold 1 (Test D3), (b) Fold 2 (Test D2), and (c) Fold 3 (Test D1).

The corresponding feature-space visualization of the six-axis configuration is shown in Figure 13.

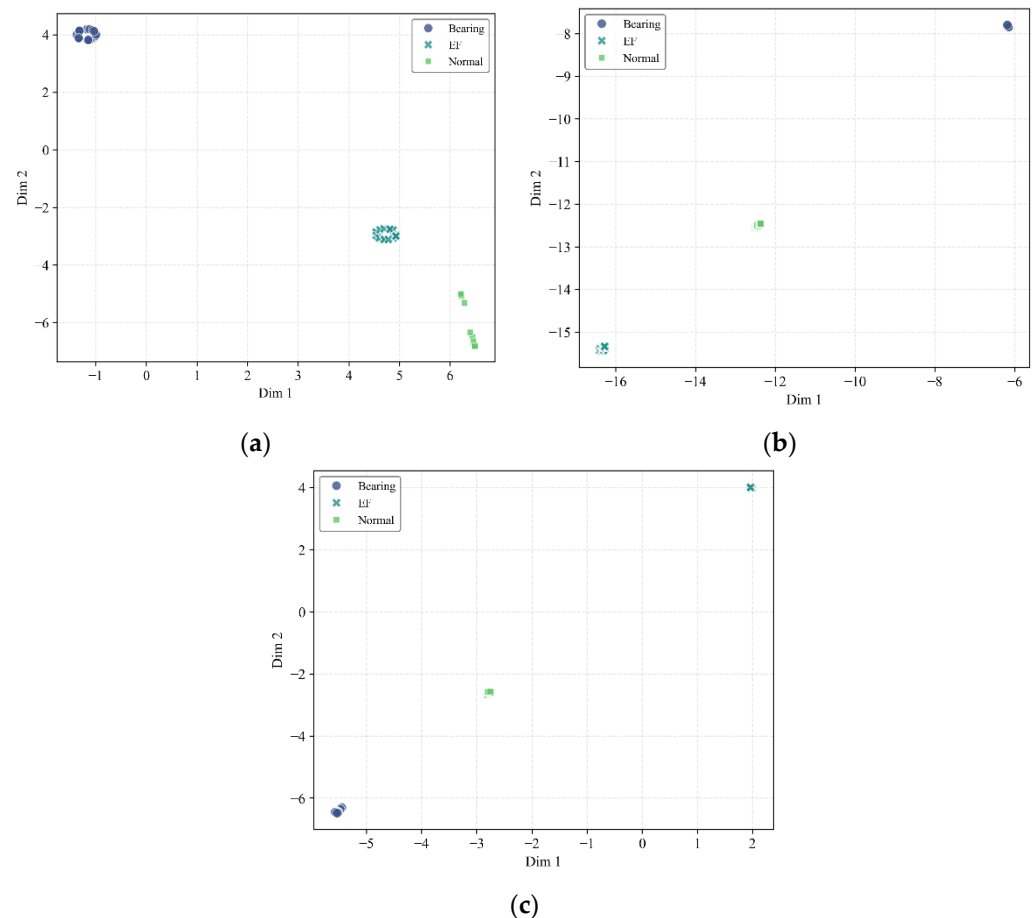


Figure 13. t-SNE visualization of the learned feature representations of the six-axis 1D-CNN configuration under LODO cross-validation: (a) Fold 1 (testing on D3), (b) Fold 2 (testing on D2), and (c) Fold 3 (testing on D1).

Notably, this result does not imply that gyroscope signals are universally unnecessary for all scenarios involving motor fault diagnosis. Rather, the additional gyroscope channels failed to provide sufficient discriminative information to improve robustness and introduced additional variability across acquisition days under the present dataset and LODO protocol. Further analyses, including feature-space visualization of the six-axis configuration and sensitivity evaluation with respect to IMU placement and orientation, remain important directions for future work.

4.5. Additional Cross-Speed Validation

To address the concern that the original experimental platform was mainly evaluated under a single operating speed of 1200 RPM, additional cross-speed experiments were conducted. In this study, two complementary evaluation settings were considered. First, a strict single-speed training setting, in which the model was trained only using data collected at 1200 RPM and subsequently tested under unseen speed conditions, was examined. Second, a mixed-speed training setting, in which data collected at 1100, 1200, and 1300 RPM were jointly used for training, was introduced, and the model was subsequently evaluated on held-out subsets for each target speed.

Table 7 lists the results obtained under the strict single-speed training setting. When the model was trained only at 1200 RPM and tested at 1100 RPM, the fold-wise accuracies

obtained were 66.67%, 100.00%, and 92.31%, yielding an average accuracy of 86.32%. When tested at 1300 RPM, the fold-wise accuracies obtained were 66.67%, 66.67%, and 66.67%, with an average accuracy of 66.67%. These results indicate that the proposed model preserves a certain degree of cross-speed fault-discrimination capability, even when trained under a single-speed condition. However, the generalization performance was not consistently satisfactory, especially in the 1200 and 1300 RPM cases, suggesting that a narrow training-speed distribution limits robustness under speed variation.

Table 7. Cross-speed diagnosis results under single-speed training.

Training Basis	Testing Speed	Fold 1	Fold 2	Fold 3	Average Accuracy
1200 RPM	1100 RPM	66.67%	100.00%	92.31%	86.32%
1200 RPM	1300 RPM	66.67%	66.67%	66.67%	66.67%

An additional mixed-speed training strategy was adopted to further investigate whether the aforementioned limitation mainly originated from the restricted training-speed condition, rather than from the model structure. In this setting, the training distribution included data obtained at 1100, 1200, and 1300 RPM; however, in each fold, the specific held-out test group from the target speed condition was excluded from training and used solely for testing. Therefore, although the training data covered multiple operating speeds, the specific held-out test group in each fold was excluded during training. The corresponding results are listed in Table 8. Under this configuration, the model achieved 100.00%, 100.00%, and 99.57% average accuracies using the held-out 1100, 1200, and 1300 RPM subsets, respectively. These results demonstrate that when the training data cover multiple adjacent operating speeds, the proposed framework can learn highly robust fault-related representations and maintain highly consistent performance across different speed conditions.

Table 8. Diagnosis results under mixed-speed training with held-out testing at each speed.

Training Strategy	Held-Out Test Speed	Fold 1	Fold 2	Fold 3	Average
Mixed-speed training (1100 + 1200 + 1300)	1100 RPM	100.00%	100.00%	100.00%	100.00%
Mixed-speed training (1100 + 1200 + 1300)	1200 RPM	100.00%	100.00%	100.00%	100.00%
Mixed-speed training (1100 + 1200 + 1300)	1300 RPM	98.72%	100.00%	100.00%	99.57%

Overall, comparing the data in Tables 7 and 8 reveals a clear trend. Under the strict single-speed training scenario, the model exhibited only partial cross-speed generalization. In contrast, when speed diversity was incorporated into the training set, the diagnostic performance approached perfection across all held-out speed conditions. This comparison suggests that the proposed framework benefits substantially from multi-speed training and that training-speed diversity is critical for achieving robust fault diagnosis under variable operating speeds.

5. Conclusions

This study presented a robust and cost-effective fault diagnosis framework for BLDC motors, designed to address nonstationary environmental conditions and hardware cost constraints. Specifically, a custom experimental testbed with realistic fault injections, including abrasive bearing wear and dynamic eccentricity, was developed to evaluate the

proposed method. The framework was rigorously assessed using a LODO cross-validation protocol, thereby simulating real-world sensor baseline drift across different operating days.

The experimental analysis produced several important findings. First, the feature analysis confirmed that three-axis accelerometer signals obtained from a low-cost IMU contain sufficient discriminative information for reliable fault identification. Independent *t*-test analysis suggested that gyroscope data provided limited additional diagnostic value under the present dataset and protocol. By contrast, the three-axis accelerometer configuration was sufficient for reliable fault identification. This sensor optimization reduced computational complexity and hardware cost while maintaining high diagnostic accuracy.

Second, the proposed end-to-end 1D-CNN demonstrated considerably improved generalization and robustness relative to traditional machine learning models. The 1D-CNN achieved 100.00% average window-level classification accuracy across the three validation folds under the present controlled dataset and the LODO protocol. The model successfully learned shift-invariant morphological features directly from raw time-domain waveforms. In contrast, the performances of traditional models that rely on handcrafted statistical features, including MLNN and SVM, degraded substantially when the signal distribution changed across different days. These results highlight the limitations of static feature engineering in dynamic industrial environments.

Furthermore, the proposed framework demonstrated strong computational efficiency. The 1D-CNN achieved an inference latency of only 0.96 ms per sample, representing a 32-fold speed improvement relative to the RF baseline (0.96 ms vs. 30.7 ms). This level of efficiency indicates that the proposed approach is well suited for real-time deployment on resource-constrained embedded systems and edge computing devices.

In summary, this study demonstrates that high-performance fault diagnosis is possible through time-domain signal analysis combined with deep learning, without requiring complex frequency-domain transformations. The proposed approach supports rapid fault detection and real-time decision making while maintaining low computational complexity and high diagnostic accuracy. Future work will further extend this evaluation to additional operating-condition variations, including fluctuating load levels, supply variations, and mounting-condition changes. In addition, lightweight model optimization techniques, such as quantization and pruning [32], will be investigated to further improve deployment efficiency using low-power microcontrollers for advanced edge artificial intelligence applications. Although the present study compares the proposed framework with representative conventional machine learning baselines under the same LODO protocol, the inclusion of strong deep learning-based baselines and explicit domain-adaptive models remains an important direction for future work.

Author Contributions: Conceptualization, K.H.W. and I.S.L.; methodology, K.H.W. and I.S.L.; software, K.H.W., H.G.L. and S.M.Y.; validation, K.H.W., H.G.L. and S.M.Y.; investigation, K.H.W.; resources, K.H.W.; writing—original draft preparation, K.H.W.; writing—review and editing, K.H.W. and I.S.L.; supervision, I.S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Gundogdu, T.; Komurgoz, G. Self-tuning PID control of a brushless DC motor by adaptive interaction. *IEEE Trans. Electr. Electron. Eng.* **2014**, *9*, 384–390. [\[CrossRef\]](#)
- Liu, C.; Chau, K.T.; Lee, C.H.T.; Song, Z. A critical review of advanced electric machines and control strategies for electric vehicles. *Proc. IEEE* **2020**, *109*, 1004–1028. [\[CrossRef\]](#)
- Zhang, P.; Du, Y.; Habetler, T.G.; Lu, B. A survey of condition monitoring and protection methods for medium-voltage induction motors. *IEEE Trans. Ind. Appl.* **2011**, *47*, 34–46. [\[CrossRef\]](#)
- Choudhary, A.; Jamwal, D.M.M.; Chelladurai, D. Condition monitoring and fault diagnosis of induction motors: A review. *Arch. Comput. Methods Eng.* **2019**, *26*, 1221–1238. [\[CrossRef\]](#)
- Randall, R.B.; Antoni, J. Rolling element bearing diagnostics—A tutorial. *Mech. Syst. Signal Process.* **2011**, *25*, 485–520. [\[CrossRef\]](#)
- Liang, P.; Deng, C.; Wu, J.; Yang, Z. Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network. *Measurement* **2020**, *159*, 107768. [\[CrossRef\]](#)
- Lei, Y.; Lin, J.; He, Z.; Zuo, M.J. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **2013**, *35*, 108–126. [\[CrossRef\]](#)
- Wang, D.; Tse, P.W.; Tsui, K.L. An enhanced kurtogram method for fault diagnosis of rolling element bearings. *Mech. Syst. Signal Process.* **2016**, *72–73*, 701–720. [\[CrossRef\]](#)
- LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#)
- Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [\[CrossRef\]](#)
- Tao, H.; Jia, P.; Wang, X.; Wang, L. Real-Time Fault Diagnosis for Hydraulic System Based on Multi-Sensor Convolutional Neural Network. *Sensors* **2024**, *24*, 353. [\[CrossRef\]](#)
- Feng, Z.; Tong, Q.; Jiang, X.; Lu, F.; Du, X.; Xu, J.; Huo, J. Deep reconstruction transfer convolutional neural network for rolling bearing fault diagnosis. *Sensors* **2024**, *24*, 2079. [\[CrossRef\]](#)
- Xiong, S.; Zhang, L.; Yang, Y.; Zhou, H.; Zhang, L. Multi-source information-based bearing fault diagnosis using multi-branch selective fusion deep residual network. *Sensors* **2024**, *24*, 6581. [\[CrossRef\]](#) [\[PubMed\]](#)
- Jiang, W.; Gu, X.; Zhao, Y.; Tang, E.; Jiang, X.; Song, Z.; Zeng, L. Intelligent fault diagnosis of hydraulic pumps based on multi-source signal fusion and dual-attention convolutional neural networks. *Sensors* **2025**, *25*, 7018. [\[CrossRef\]](#)
- Ma, C.; Fei, J.; Shi, Z.; Rob, M.A.; Islam, M.A.; Habibullah, M. Multi-Source Fusion CNN-RF Framework for Intelligent Fault Diagnosis of Head Sheave Devices in Mining Hoists. *Machines* **2026**, *14*, 244. [\[CrossRef\]](#)
- Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [\[CrossRef\]](#)
- Qian, W.; Li, S.; Yi, P.; Zhang, K. A novel transfer learning method for robust fault diagnosis of rotating machinery under variable working conditions. *Measurement* **2019**, *138*, 514–525. [\[CrossRef\]](#)
- Guo, L.; Lei, Y.; Xing, S.; Yan, T.; Li, N. Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data. *IEEE Trans. Ind. Electron.* **2019**, *66*, 7316–7325. [\[CrossRef\]](#)
- Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2446–2455. [\[CrossRef\]](#)
- Gong, C.; Li, Y.R.; Zargari, N.R. An Overview of Advancements in Multimotor Drives: Structural Diversity, Advanced Control, Specific Technical Challenges, and Solutions. *Proc. IEEE* **2024**, *112*, 184–209. [\[CrossRef\]](#)
- Shao, X.; Kim, C.-S. Unsupervised Domain Adaptive 1D-CNN for Fault Diagnosis. *Sensors* **2022**, *22*, 4156. [\[CrossRef\]](#)
- Gu, Y.; Zhang, Y.; Yang, M.; Li, C. Motor On-Line Fault Diagnosis Method Research Based on Multi-Sensor Fusion and 1D-CNN. *Appl. Sci.* **2023**, *13*, 4192. [\[CrossRef\]](#)
- Gao, T.; Sheng, W.; Zhou, M.; Fang, B.; Zheng, L. MEMS inertial sensor fault diagnosis using a CNN-based data-driven method. *Int. J. Pattern Recognit. Artif. Intell.* **2020**, *34*, 2059048. [\[CrossRef\]](#)
- Cao, J.; Zhang, J.; Jiao, X.; Yu, P.; Zhang, B. Gearbox fault diagnosis method in noisy environments based on deep residual shrinkage networks. *Sensors* **2024**, *24*, 4633. [\[CrossRef\]](#)
- Łuczak, D. Machine fault diagnosis through vibration analysis: Time series conversion to grayscale and RGB images for recognition via convolutional neural networks. *Energies* **2024**, *17*, 1998. [\[CrossRef\]](#)
- Kim, M.-C.; Lee, J.-H.; Wang, D.-H.; Lee, I.-S. Induction motor fault diagnosis using support vector machine, neural networks, and boosting methods. *Sensors* **2023**, *23*, 2585. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lee, H.G.; Yoo, S.M.; Wang, K.H.; Lee, I.S. Time-domain and neural network-based diagnosis of bearing faults in induction motors under variable loads. *Machines* **2025**, *13*, 1055. [\[CrossRef\]](#)
- Guo, Z.; Du, W.; Li, C.; Guo, X.; Liu, Z. Fault diagnosis of rotating machinery with high-dimensional imbalance samples based on wavelet random forest. *Measurement* **2025**, *248*, 116936. [\[CrossRef\]](#)
- Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the International Conference on Machine Learning (ICML), Lille, France, 6–11 July 2015; pp. 448–456.

30. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
31. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
32. Han, S.; Mao, H.; Dally, W.J. Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding. *arXiv* **2016**, arXiv:1510.00149. [[CrossRef](#)]

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.