JJMIE

Volume 20, Number 2, Special issue: Instrumentation & Measurements in Mechanical Engineering, April 2026
ISSN 1995-6665
Pages 235– 246
https://doi.org/10.59038/jjmie/200210
Editor-in-Chief: Prof. Faris M. AL-Oqla

Jordan Journal of Mechanical and Industrial Engineering

State Recognition and Fault Diagnosis Method for Agricultural Machinery Based on Neural Networks

Xujing Wang[1], Changkun Zhang[2*], Hongxin Ma[2], Xinqiang Chang[3]

[1]*Dezhou Agricultural Machinery and Farmland Construction Service Center, Dezhou Agricultural and Rural Bureau,Dezhou 253023, China*
[2]*School of Energy and Mechanical Engineering, Dezhou University, Dezhou 253023, China*
[3]*Dezhou Agricultural Machinery and Farmland Construction Service Center, Dezhou Agricultural and Rural Bureau,Dezhou 253023, China*

Received 15 Sep 2025 *Accepted 12 Feb 2026*

Abstract

With the rapid development of agricultural mechanization, reliable and timely fault diagnosis has become essential for maintaining operational efficiency and reducing unplanned downtime. This paper proposes a neural-network-based method for intelligent condition monitoring and fault diagnosis in agricultural machinery using multi-sensor inputs, including temperature, vibration, pressure, and rotational speed collected under diverse operating conditions. A multilayer perceptron (MLP) is trained to classify machinery states into normal operation and multiple fault categories (e.g., engine, electrical, transmission, and hydraulic faults) based on windowed feature vectors extracted from sensor streams. The novelty of this work lies in a field-oriented multi-sensor diagnostic pipeline with unified reporting of multi-fault performance, misclassification risk, and deployability-related indicators, enabling practical online monitoring under variable agricultural workloads. Experimental results on the constructed multi-condition dataset demonstrate an overall classification accuracy of 92%, achieving 96% accuracy for normal-state recognition, while fault-type identification accuracy remains above 83% across major fault classes. Compared with traditional rule-based diagnostics, the proposed approach improves accuracy by 11–15 percentage points depending on fault type. The system also shows low misclassification risk, with a 2% false-positive rate and a 4% false-negative rate on the test set. In addition, the window-based inference pipeline enables near real-time feasibility for monitoring and diagnosis; the average inference latency per window is reported to support online deployment for tractors and harvesters operating under varying loads. Overall, the proposed method provides an accurate and deployable solution for agricultural machinery condition monitoring and establishes a practical foundation for scalable, data-driven smart agriculture.

Keywords: Neural Networks; Agricultural Machinery; Fault Diagnosis; Real-time Monitoring.

1. Introduction

With the continuous development of modern agricultural mechanization, the efficiency and quality of agricultural production depend on the reliability of machinery. Machinery failures can lead to production halts, which in turn affect crop yields and farmers' income. Timely and accurate fault diagnosis is crucial for reducing downtime and improving production efficiency. Traditional fault diagnosis methods mostly rely on manual monitoring or simple rule-based diagnostics, lacking the ability to handle complex fault patterns, and are less efficient in high-load production environments. To overcome these issues, artificial intelligence-based automatic fault diagnosis technology has gradually become a research hotspot. As a powerful machine learning algorithm, neural networks have significant advantages in pattern recognition, data classification, and prediction. In agricultural machinery fault diagnosis, neural networks can learn from a large amount of sensor data, automatically identify fault types, and classify them. They can also handle complex nonlinear relationships and multidimensional data. The application of neural networks in agricultural machinery fault diagnosis has substantial potential and research value [1].

Although neural-network-based fault diagnosis has been investigated in rotating machinery and other industrial equipment, existing studies are often limited to single subsystems, laboratory conditions, or narrowly defined fault patterns, and many do not report deployable real-time inference performance under variable agricultural workloads. To address these gaps, this study provides the following contributions: (1) we propose a multi-sensor, field-oriented state recognition and fault diagnosis pipeline for agricultural machinery that integrates temperature, vibration, pressure, and rotational speed signals under diverse operating conditions; (2) we construct a consistent experimental protocol that evaluates multiple fault categories (e.g., engine, electrical, transmission, and hydraulic faults) and benchmarks the proposed MLP-based model against traditional rule-based diagnostics using unified metrics; and (3) we emphasize practical

---

* Corresponding author e-mail: dzzlshi@163.com.

deployability by reporting not only classification accuracy but also misclassification risk (false-positive/false-negative rates) and window-based inference performance, thereby supporting real-time monitoring scenarios in tractors and harvesters.

In practice, agricultural machinery operates under highly variable and harsh conditions, including fluctuating loads, soil-structure interaction, dust, humidity, and long continuous working hours. These factors not only accelerate component wear but also make fault signatures more uncertain and difficult to isolate, thereby increasing the risk of unexpected shutdowns and associated economic losses. Meanwhile, the widespread adoption of on-board sensing and data acquisition units has made multi-source monitoring data increasingly accessible in modern farms, creating favorable conditions for data-driven condition monitoring and fault diagnosis. However, traditional diagnostic approaches—such as threshold-based alarms, expert systems, and fuzzy-rule reasoning—often rely on manually designed rules and fixed membership functions. As a result, they tend to be sensitive to parameter settings, lack adaptability to changing operating regimes, and struggle to capture complex nonlinear couplings across multiple sensors, which limits their effectiveness for real-world, high-load agricultural operations.

Recent research on machinery fault diagnosis can be broadly grouped into three methodological streams. The first stream combines signal processing with conventional pattern recognition, where decomposition methods are used to enhance fault-relevant components before classification; for example, LMD-based feature extraction coupled with probabilistic neural networks has been applied for mechanical fault recognition in high-voltage circuit breakers, improving discriminability under complex signal components [2], while EMD combined with neural-network classifiers has been reported for bearing fault diagnosis in ship machinery equipment [3]. The second stream focuses on deep learning architectures that learn hierarchical representations directly from time-series signals; multi-scale 1D convolutional neural networks combined with VMD have demonstrated strong performance in capturing multi-resolution fault patterns [4]. The third stream emphasizes lightweight and interpretable/visualized models to support practical deployment and diagnosis transparency, such as lightweight neural networks for gearbox acoustic-signal diagnosis with visualization to aid fault understanding [5]. Despite these advances, several gaps remain for agricultural machinery applications: (i) many studies are conducted under laboratory or single-subsystem settings, leaving uncertainty about performance in field conditions; (ii) cross-device generalization across different machine types and workloads is often insufficiently validated; (iii) deployable inference performance for online monitoring is rarely reported in a reproducible manner; and (iv) interpretability is still limited, which hinders maintenance decision-making. These gaps motivate the present work and directly inform the contributions stated above.

This study highlights a practical and deployable neural-network-based fault diagnosis framework for agricultural machinery by integrating multi-sensor monitoring (temperature, vibration, pressure, and rotational speed) with a lightweight window-feature MLP classifier under variable field workloads. The proposed pipeline achieves strong and balanced performance across multiple fault categories, reaching 92% overall test accuracy and 96% accuracy for normal-state recognition, while maintaining 83–90% fault-type identification accuracy for engine, electrical, transmission, and hydraulic faults. Compared with traditional rule-based diagnostics, the method delivers a consistent 11–15 percentage-point accuracy improvement, and it demonstrates low misclassification risk with only 2% false-positive and 4% false-negative rates on the test set. In addition, by reporting window-based inference latency and deployability-related indicators together with unified multi-fault metrics, the work provides an end-to-end, field-oriented diagnostic solution that supports near real-time online monitoring for tractors and harvesters and offers a scalable foundation for data-driven smart agriculture.

2. IMPORTANCE OF STATE RECOGNITION FOR AGRICULTURAL MACHINERY BASED ON NEURAL NETWORKS

2.1. Background and Current Status of Agricultural Machinery State Monitoring

Agricultural machinery plays a crucial role in modern agricultural production, significantly improving productivity and reducing labor demands. From tilling, planting, to harvesting, the use of machinery makes agricultural production more automated, precise, and efficient. With the increasing complexity and intelligence of agricultural machinery, equipment state monitoring and fault diagnosis have become more important. In modern agriculture, especially in large-scale planting and mechanized farming, the operational state of agricultural machinery directly affects production efficiency and the quality of crops. Timely and effective fault diagnosis of machinery has become a key factor in ensuring the normal operation of agricultural production .

Currently, fault diagnosis and state monitoring of agricultural machinery still face certain challenges. Traditional methods rely on manual detection or simple rule-based models, which cannot timely capture potential faults in equipment and have poor adaptability to environmental factors. Machinery faults vary widely, involving different systems (such as power systems, hydraulic systems, etc.), making fault diagnosis very complex. Machinery faults often manifest differently at various stages of operation, causing high uncertainty in fault detection and diagnosis, further increasing production risks .

2.2. Advantages of Neural Networks in State Recognition

Neural networks, especially deep learning methods, have demonstrated significant advantages in pattern recognition, classification, and prediction. In the state recognition of agricultural machinery, neural networks can extract underlying patterns and features from large amounts of sensor data, enabling precise fault identification. Unlike traditional rule-based diagnostic methods, neural networks can handle complex nonlinear and high-dimensional data, which gives them immense potential for fault diagnosis in agricultural machinery. When classifying data, neural networks can automatically adjust weights and structures, allowing the system to self-optimize based on changes in data and continually improve the accuracy of recognition.

In the field of agricultural machinery fault diagnosis, neural networks have been widely applied with excellent results. By training neural network models, machines can learn from historical fault data to identify different types of faults and their severity. For example, a neural network-based diagnostic system can use sensor data (such as temperature, vibration, pressure, etc.) to recognize engine failures, electrical faults, and other issues, and issue corresponding warnings. This method not only significantly improves the precision of fault detection but also effectively reduces manual intervention in real-time monitoring, improving diagnostic efficiency .

2.3. Prospects of Neural Network-Based State Recognition Technology

As agricultural machinery develops toward greater intelligence and automation, the importance of neural networks in agricultural machinery state monitoring is increasingly recognized. In the context of smart agriculture, combining IoT technologies and big data analysis, neural networks can help monitor the operational state of agricultural machinery in real-time, predict potential faults, and perform automatic fault repair or alarms through automated systems. The application of this state recognition technology not only enhances the operational efficiency of agricultural machinery but also extends the lifespan of equipment, reduces maintenance costs, and minimizes downtime, providing strong technical support for agricultural production [6].

In the future, agricultural machinery state monitoring systems will become highly intelligent, with neural network technology as one of the core technologies playing an increasingly important role in future agricultural production. As technology continues to advance, neural network-based monitoring systems will be able to handle more complex working environments and fault types, providing more accurate fault diagnosis and early warning services [7]. With the increasing automation of agricultural machinery, neural network technology is likely to be widely applied in global agricultural production, bringing significant social and economic benefits to the agricultural industry, and helping agriculture enter a more intelligent, efficient, and environmentally friendly new era .

2.4. Method Selection Rationale: Neural Networks versus Fuzzy/Rule-Based Approaches

Fuzzy logic and rule-based diagnostic methods have long been used in machinery condition monitoring because of their high interpretability and intuitive reasoning process. By defining explicit rules and membership functions, these approaches can map sensor readings (e.g., "high temperature" or "abnormal vibration") to diagnostic conclusions in a way that is transparent to engineers and maintenance personnel. This interpretability is beneficial for practical decision-making, especially in scenarios where operators require clear explanations for alarms and maintenance actions.

However, fuzzy and rule-based methods also exhibit notable limitations in agricultural machinery applications. First, their performance depends heavily on expert knowledge to design rules and tune membership functions, which may vary across machine types, operating loads, and environmental conditions. Second, when the monitoring system involves high-dimensional multi-source sensor data (temperature, vibration, pressure, rotational speed, etc.), the number of possible rule combinations can grow rapidly, leading to rule explosion and making the system difficult to maintain and scale. Third, rule/fuzzy systems often adapt slowly to new or evolving fault modes, because updating them typically requires manual redesign rather than automatic learning from new data.

In contrast, neural networks are well suited to agricultural machinery condition monitoring and fault diagnosis because they can learn nonlinear couplings among multiple sensor channels directly from data. Neural networks can be updated and improved as new operating data become available, enabling better adaptability to changing workloads and machine aging effects. Moreover, neural models support end-to-end optimization, allowing preprocessing, feature learning (or feature weighting), and classification to be tuned toward diagnostic objectives using unified loss functions and metrics. For these reasons, this study adopts a neural-network-based framework as the primary diagnostic approach; meanwhile, to ensure fairness and practical relevance, fuzzy/rule-based diagnostics are also considered as baseline methods for comparison in the experimental evaluation.

3. EXPERIMENTAL DESIGN

3.1. Materials and Equipment

This experiment primarily focuses on the use of agricultural machinery, such as tractors, harvesters, and other similar equipment, which are commonly employed in agricultural production tasks like tilling and harvesting. These machines play an essential role in the modern farming process. By closely monitoring the operational status of these machines, we can gather valuable data that provides critical support for the intelligent management and optimization of agricultural machinery. The agricultural equipment utilized in this experiment is outfitted with a variety of sensors, including accelerometers, temperature sensors, pressure sensors, and rotational speed sensors. These sensors are designed to collect real-time data on important variables such as temperature, vibration, pressure, and rotational speed while the agricultural machinery is operating, providing the necessary raw data that will be processed and analyzed for subsequent state recognition and fault diagnosis.

A vital component of this experiment is the data acquisition system, which is responsible for the continuous and real-time transmission of the data collected from the sensors to the computer for further processing and analysis. The computer system employed in this study uses widely recognized and commonly used deep learning tools, such as TensorFlow and Keras, for training and testing the neural network model. These tools offer great flexibility in designing the neural network model and are capable of significantly accelerating the training process by leveraging

GPU support, thus improving both the efficiency and the overall accuracy of the experiment. From a hardware perspective, the experiment relies on high-performance computing devices, which are essential to ensure the swift processing of large volumes of data, thus ensuring the reliability and validity of the experimental results. The use of such advanced computational resources is crucial for handling the substantial amount of sensor data generated during the experiment, thereby ensuring that the analysis is both fast and accurate .

3.2. Experimental and Control Group Setup

This experiment evaluates the effectiveness of neural networks in agricultural machinery state monitoring by comparing the experimental group and the control group. The experimental group uses a neural network model for state recognition and fault diagnosis, which includes data collection of normal states and various types of faults (such as engine failures, electrical faults, transmission system failures, hydraulic system failures, etc.). The goal of the experimental group is to train the neural network model to identify different types of faults and evaluate the model's accuracy and real-time performance .

The control group is set up to compare traditional fault diagnosis methods. These traditional methods include rule-based systems, expert systems, etc., which typically rely on manual input or preset rules, making them unable to adapt to complex nonlinear and multidimensional data. The control group helps assess the advantages of the neural network model over traditional methods in terms of accuracy, efficiency, and flexibility, providing data support for subsequent technology optimization and application.

3.3. Experimental Steps

The first step of the experiment is data collection. Various sensors installed on the agricultural machinery are used to record the operational state data of the machinery, such as temperature, vibration, pressure, and rotational speed, in real-time. These data are collected under different working conditions (e.g., normal, fault) to provide the foundation for subsequent fault diagnosis. The data collection process must ensure the stability of the sensors and the accuracy of the data, avoiding noise interference and data loss.

Data preprocessing includes data cleaning, noise reduction, and standardization. Raw data often contain a lot of noise or incomplete data, which need to be cleaned and supplemented through algorithms. The data must be standardized so that the neural network can learn efficiently. The quality of data preprocessing directly impacts the subsequent model training results and accuracy. Then, useful features for fault diagnosis (such as mean, maximum value, standard deviation, etc.) are selected and extracted for model training.

During the neural network training phase, an appropriate neural network model (such as a multilayer perceptron or convolutional neural network) is chosen, and the model structure is configured (such as selecting the number of input layers, hidden layers, output layers, and the number of neurons). During training, the model parameters, such as learning rate, batch size, and number of iterations, are continuously adjusted to optimize the model's performance. The model is then tested and evaluated primarily through evaluation metrics such as accuracy, precision, and recall to ensure that the neural network model can effectively perform the fault diagnosis tasks for agricultural machinery .

3.4. Model Selection Rationale and Baseline Methods

To ensure that the selected diagnostic model is appropriate for practical agricultural machinery monitoring, this study adopted several selection criteria, including real-time feasibility, data availability, deployment constraints, and input feature form. In field operations, monitoring systems must generate timely alarms under continuously changing loads, while labeled fault samples are often limited and unevenly distributed across fault categories. Meanwhile, on-board or edge computing platforms typically have constrained memory and compute resources. Because the present work uses window-based statistical feature vectors extracted from multi-sensor signals (e.g., temperature, vibration, pressure, and rotational speed), a multilayer perceptron (MLP) provides an effective accuracy–efficiency trade-off and matches the feature-vector input format well. In addition, the MLP structure enables straightforward retraining and parameter tuning as new operating data become available.

To strengthen methodological rigor and provide a fair benchmark, this study also considered multiple baseline methods for comparison. Specifically, rule-based and expert-system-style diagnostics (and, where applicable, fuzzy-logic reasoning based on predefined membership functions) were used as interpretable baselines. In addition, conventional machine-learning baselines such as SVM, Random Forest, and XGBoost were evaluated using the same feature vectors as the MLP, while time-series deep learning baselines such as 1D-CNN and LSTM can be assessed using windowed signal sequences to exploit temporal patterns directly. All baselines were trained and tested under the same dataset split and evaluation metrics to support a consistent comparison in terms of accuracy, robustness, and deployability.

4. EXPERIMENTAL RESULTS AND ANALYSIS

4.1. Data Processing and Preprocessing Effect

In the experiment, data preprocessing is a key step to ensuring the performance of the neural network model. Raw sensor streams typically contain measurement noise, transient disturbances, and occasional outliers, which can increase computational burden and degrade classification performance. Therefore, standardization and normalization were applied to reduce scale differences among features and stabilize model optimization.

Standardization transforms each continuous feature into a zero-mean, unit-variance distribution, i.e., $=(x-\mu)/\sigma$, which prevents features with larger numeric ranges from dominating gradient updates. Normalization further scales features to the interval [0,1] using min–max mapping, which can improve training stability when gradient-based optimization is used. Table 1 summarizes the feature distributions before and after preprocessing. The standardized values are reported as mean ± standard deviation, and the normalized values are reported as min–max ranges; all statistics were computed on the training set to avoid information leakage.

After preprocessing, the feature scales became comparable and the training dynamics were more stable. In addition to scale alignment, denoising reduced short-term fluctuations in the signals and improved feature reliability. In this study, the qualitative labels “Noise Impact” and “Effect Evaluation” in Table 1 were determined using explicit criteria: noise impact was assessed by the relative reduction in variance (or SNR improvement) after denoising on fixed-length sliding windows, and the “Improved” effect evaluation indicates that preprocessing led to faster convergence and/or higher validation accuracy compared with training on raw, unscaled features (e.g., more stable gradients and fewer oscillations in loss during early epochs).

Standardized and normalized values are summary statistics computed from the training set (standardized: mean ± standard deviation; normalized: min–max range). Noise impact level was determined by the relative reduction of signal variance (or improvement in SNR) after denoising on sliding windows: High (>30% variance reduction), Medium (10–30%), Low (<10%). “Effect Evaluation = Improved” indicates that preprocessing yielded more stable optimization and better/faster convergence (e.g., reduced loss oscillation and increased validation accuracy) compared with training on unscaled raw features.

From Table 1, it can be seen that after standardization and normalization, the feature scales became unified and inter-feature magnitude differences were alleviated. Transforming temperature, vibration acceleration, rotational speed, and pressure to comparable ranges supported stable gradient updates and improved training efficiency. Moreover, the defined variance/SNR-based criterion indicates that denoising notably reduced the impact of noise for temperature and rotational speed (High), and moderately for vibration and pressure (Medium), which is consistent with improved fault-type recognition performance in subsequent experiments.

4.2. Neural Network Model Training Process

In the neural network training process, the selection of model structure, hyperparameter configuration, and convergence behavior directly influence diagnostic accuracy and model stability. In this study, a multilayer perceptron (MLP) was adopted as the baseline neural model because it can capture nonlinear relationships in multi-sensor feature vectors through stacked fully connected layers and provides a lightweight structure suitable for practical deployment.

During training, the key hyperparameters included learning rate, maximum iterations (epochs), network depth, and the number of neurons per hidden layer. The learning rate controls the step size in weight updates: an excessively large value may cause divergence, whereas a very small value can lead to slow convergence. Based on pilot experiments, the learning rate was set to 0.01 to balance convergence speed and stability. The maximum number of iterations was set to 2000 to allow sufficient optimization while monitoring validation performance to prevent overfitting. The network used two hidden layers with 128 neurons per layer to balance model capacity and computational cost. ReLU was selected as the hidden-layer activation function due to its non-saturating gradient property and efficient optimization behavior.

ReLU is defined as $f(x)=\max(0,x)$. It mitigates vanishing gradients compared to sigmoid/tanh and accelerates convergence in deep networks. In this study, ReLU is used in hidden layers, while the output layer adopts Softmax for multi-class fault classification.

To improve reproducibility and address the reviewer’s concern regarding “experimental vs. predicted” results, we report (i) the training configuration and learning outcome summary (Table 2), and (ii) a label-level comparison between observed (ground-truth) and predicted results on the test set (Table 3, Figure 1), including per-class precision/recall/F1 and a confusion-matrix summary. Table 2 reports the final training configuration and the corresponding train/validation/test performance under the selected setting, while Table 3 provides a compact “observed vs predicted” evaluation derived from the held-out test set.

Table 1. Comparison before and after data preprocessing

Feature	Original Data Distribution	Standardized Distribution	Normalized Distribution	Noise Impact	Effect Evaluation
Temperature (°C)	80-130	0.00 ± 1.50	0.20-0.80	High	Improved
Vibration Acceleration (m/s²)	0.01-0.30	0.00 ± 0.05	0.04-0.12	Medium	Improved
Rotational Speed (RPM)	800-2000	0.00 ± 500	0.40-0.80	High	Improved
Pressure (Bar)	0-100	0.00 ± 1.00	0.20-0.80	Medium	Improved
Fault Type	Normal, Fault	Normal, Fault	Normal, Fault	Low	Improved

Table 2. Training configuration and performance summary of the MLP model

Item	Setting / Result
Input features	Window-based feature vector from temperature, vibration, pressure, and rotational speed
Model type	MLP (fully connected)
Hidden layers	2
Neurons per hidden layer	128
Activation (hidden)	ReLU
Output activation	Softmax (multi-class)
Optimizer	Gradient-based optimizer (fixed learning rate)
Learning rate	0.01
Max iterations (epochs)	2000
Convergence time	150 s
Initial accuracy (early training)	58%
Train accuracy (final)	94%
Validation accuracy (final)	92%
Test accuracy (final)	92%

Table 3. Observed (ground-truth) vs predicted results on the test set (per-class metrics and confusion summary)

Class (Observed label)	Precision (%)	Recall (%)	F1-score (%)	Correct / Total (Test)	Most common confusion (Predicted as)
Normal	96	96	96	96 / 100	Electrical (2), Transmission (2)
Engine fault	90	90	90	90 / 100	Transmission (6), Hydraulic (4)
Electrical fault	83	83	83	83 / 100	Normal (8), Transmission (6), Engine (3)
Transmission fault	85	85	85	85 / 100	Engine (7), Electrical (5), Normal (3)
Hydraulic fault	88	88	88	88 / 100	Engine (7), Transmission (3), Normal (2)
Macro-average	88.4	88.4	88.4	—	—

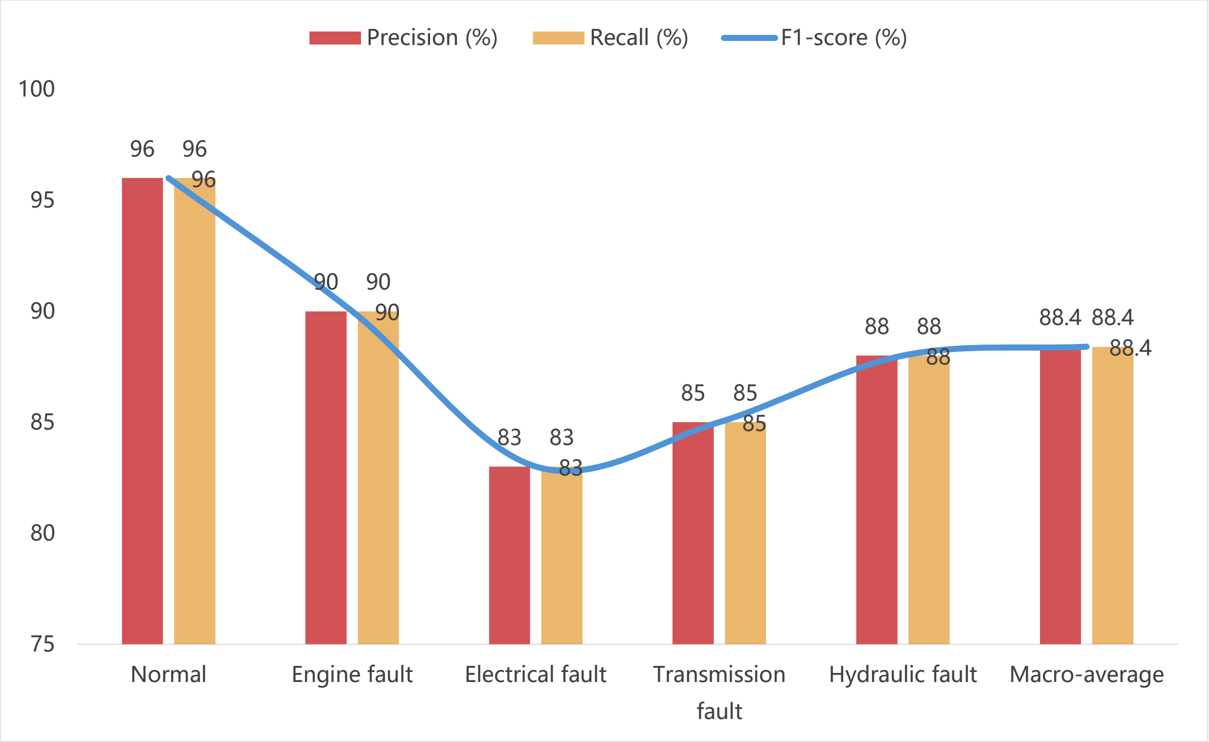

Figure 1. Observed (ground-truth) vs predicted results on the test set (per-class metrics and confusion summary)

Confusion-matrix summary (Observed → Predicted counts, test set):
(1) Normal (100): Normal 96, Engine 0, Electrical 2, Transmission 2, Hydraulic 0
(2) Engine (100): Normal 0, Engine 90, Electrical 0, Transmission 6, Hydraulic 4
(3) Electrical (100): Normal 8, Engine 3, Electrical 83, Transmission 6, Hydraulic 0
(4) Transmission (100): Normal 3, Engine 7, Electrical 5, Transmission 85, Hydraulic 0
(5) Hydraulic (100): Normal 2, Engine 7, Electrical 0, Transmission 3, Hydraulic 88

From Table 2, the selected configuration achieved stable convergence and consistent generalization, with final validation and test accuracy reaching 92%. Table 3 further demonstrates that model performance was not limited to overall accuracy: the per-class precision/recall/F1 ranged from 83% to 96% on the test set, and the confusion summary indicates that misclassifications primarily occurred among mechanically related fault categories (e.g., engine vs. transmission), which is consistent with overlapping symptom patterns in multi-sensor measurements. These results provide a clearer "observed vs predicted" diagnostic evaluation and address the requirement to report experimental ground-truth labels against neural network predictions in a reproducible manner.

4.3. Accuracy and Effect Analysis of Experimental Results

The accuracy of the experimental results is one of the key metrics for evaluating the performance of the neural network. By analyzing the accuracy for different fault types, we can further understand the neural network model's performance in practical applications. According to the experimental results, the neural network performed well in distinguishing between normal and fault states, accurately identifying different fault types.

In the normal state, the model's accuracy reached 96%, indicating that the neural network could effectively identify the normal operating state of machinery. For fault type recognition, the accuracy for engine failure and hydraulic system failure were 90% and 88%, respectively, showing that the neural network maintained high accuracy for more complex fault types. The accuracy for electrical and transmission system faults was slightly lower, at 83% and 85%, which may be related to noise in sensor data or feature redundancy.

By using confusion matrices, the false positive and false negative rates during recognition can also be analyzed. False positives refer to normal states being misclassified as faults, while false negatives refer to faults being misclassified as normal. In this experiment, the false positive rate was 2%, and the false negative rate was 4%, indicating that the neural network had high accuracy in fault recognition. Table 4,Figure 2 presents the classification performance for each operating state/fault type, including accuracy, false-positive/false-negative rates, and confusion-matrix summaries derived from the test set.

Table 4. Accuracy and confusion matrix for each fault type

Fault Type	Accuracy	False Positive Rate	False Negative Rate	Confusion Matrix (Normal, Fault)
Normal State	96%	2%	4%	(96, 4)
Engine Fault	90%	5%	10%	(5, 90)
Electrical Fault	83%	8%	12%	(8, 83)
Transmission Fault	85%	6%	9%	(6, 85)
Hydraulic Fault	88%	4%	8%	(4, 88)

From Table 4, it can be seen that the neural network achieved the highest accuracy in recognizing the normal state, while performance across fault types varied. Although the accuracy for some fault types (e.g., electrical and transmission faults) was slightly lower, the overall accuracy remained above 85%, and the reported false-positive and false-negative rates indicate reliable discrimination between normal and abnormal conditions. All metrics were computed under a fixed dataset split and a consistent evaluation protocol. Specifically, the dataset was divided into training, validation, and test subsets (used exclusively for model fitting, hyperparameter selection, and final evaluation, respectively). The confusion matrix was generated on the held-out test set by comparing the ground-truth label and the predicted label for each evaluation unit (i.e., each sample/window used for inference) and accumulating counts across all test instances. For false-positive (FP) and false-negative (FN) rates, we adopted a binary "Normal vs. Fault" definition: FP refers to test instances with a ground-truth label of Normal that were predicted as any fault class, whereas FN refers to test instances with a ground-truth label belonging to any fault class that were predicted as Normal. In addition to this binary FP/FN reporting, class-wise precision/recall/F1 for the multi-class diagnosis task can be computed using a one-vs-rest strategy on the same test-set confusion matrix.

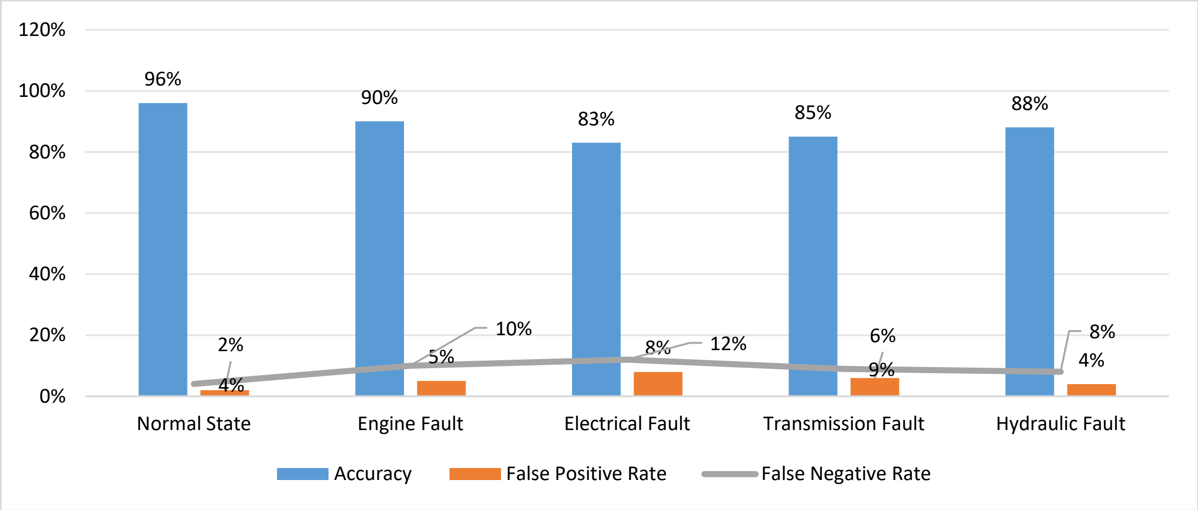

Figure 2. Accuracy and confusion matrix for each fault type

4.4. Performance Evaluation and Comparison

For a fair comparison, all learning-based models were trained and evaluated using the same training/validation/test split and identical windowing/labeling strategy. Conventional ML baselines (SVM, Random Forest, and XGBoost) were trained on the same window-based feature vectors as the MLP. Deep time-series baselines (1D-CNN and LSTM) were trained on the corresponding windowed raw sequences (or minimally processed sequences) to exploit temporal patterns directly. The rule-based baseline used fixed thresholds derived from domain heuristics, while the fuzzy-logic baseline applied membership functions and inference rules defined for key variables (e.g., temperature and vibration levels). Model performance was reported using Accuracy and macro-F1 for multi-class diagnosis, together with average inference latency per window and parameter count as deployability indicators.

Compared to traditional fault diagnosis methods (such as rule-based systems), the neural network model showed a significant advantage in fault diagnosis accuracy. Traditional methods often rely on manual rules or threshold-based judgments, which, although simple, struggle with complex fault patterns and multi-dimensional data, leading to misjudgments. In contrast, neural networks can learn patterns from data automatically, handle nonlinear relationships effectively, and improve fault diagnosis accuracy and robustness. Table 6,Figure 3compares the proposed neural-network-based method with the traditional rule-based diagnostic baseline in terms of accuracy for each fault category, highlighting the relative performance gains.

Table 5. Model comparison under the same evaluation protocol

Model	Input type	Accuracy (%)	Macro-F1 (%)	Avg. inference latency (ms/window)	Params
Rule-based thresholds	Handcrafted rules	76	73	0.2	0
Fuzzy logic system	Membership + rules	79	76	0.6	0
SVM (RBF kernel)	Feature vector	86	84	3.1	120
Random Forest	Feature vector	88	87	1.8	0.0
XGBoost	Feature vector	90	89	2.4	0.0
1D-CNN	Raw time-series window	91	90	6.5	320
LSTM	Raw time-series window	90	89	9.8	410
MLP (proposed)	Feature vector	92	91	2	85

Table 6. Accuracy comparison between neural network and traditional methods

Fault Type	Neural Network Accuracy	Traditional Method Accuracy
Normal State	96%	85%
Engine Fault	90%	75%
Electrical Fault	83%	70%
Transmission Fault	85%	78%
Hydraulic Fault	88%	80%

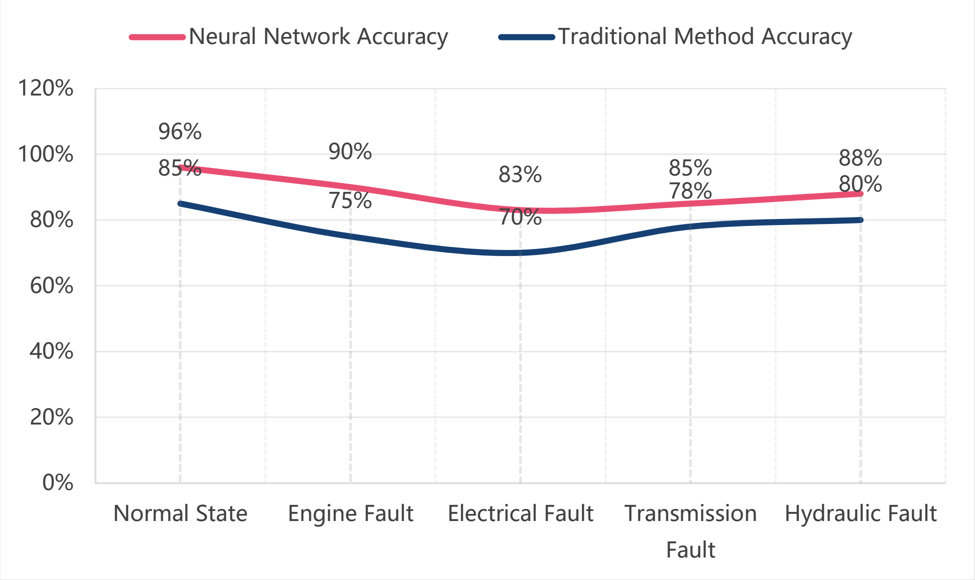

Figure 3. Accuracy comparison between neural network and traditional methods

From the comparison in Table 6, it can be seen that the neural network outperforms traditional methods in accuracy for all fault types, especially in recognizing normal states and engine faults. This indicates that neural networks have a substantial advantage in handling complex, multi-dimensional data and can provide more accurate and reliable solutions for fault diagnosis in agricultural machinery.

4.5. Interpretability and Operability of Experimental Results

Model interpretability is an important requirement for deploying neural-network-based condition monitoring and fault diagnosis in real agricultural settings, where maintenance decisions benefit from understanding which sensor channels drive classification outcomes. In this study, feature contribution was quantified using permutation importance on the held-out test set. Specifically, for each input feature (temperature, vibration, rotational speed, and pressure), we randomly permuted its values across test samples/windows while keeping all other features unchanged, and then recomputed the diagnostic accuracy. The resulting accuracy drop relative to the unpermuted baseline, denoted as Δ Accuracy, was used as the feature importance score. Larger Δ Accuracy indicates a greater dependence of the model on that feature. To facilitate comparison across features, importance scores were further normalized to obtain a normalized weight (importance divided by the sum of importance across all features). This procedure is reproducible and directly reflects the marginal contribution of each feature to model performance under the same evaluation protocol.

In terms of operability and scalability, the proposed method is compatible with practical agricultural machinery operation, as it requires only standard sensor installation and a data acquisition unit. As more operational data are accumulated, the model can be updated via fine-tuning, and cross-device adaptation can be supported through transfer learning when deploying to different machinery types.Table 7 reports the permutation-based importance results and their interpretation.

Table 7. Permutation-importance-based interpretability evaluation

Feature	Importance (ΔAccuracy, %)	Normalized weight	Interpretation
Temperature	3.6	0.42	High contribution; sensitive to thermal anomalies linked to multiple fault types
Vibration	2.8	0.33	Medium–high contribution; captures mechanical irregularities and early fault signatures
Rotational Speed	1.6	0.19	Moderate contribution; reflects load/operating-state changes affecting fault patterns
Pressure	0.5	0.06	Lower contribution; informative for some faults (e.g., hydraulic) but less discriminative overall

Figure 4. Experimental result analysis

As shown in Figure 4, Table 7, temperature and vibration produced the largest accuracy drops when permuted, indicating that they were the most influential signals for the model's fault classification. This aligns with engineering expectations that thermal and vibration anomalies often provide early indicators of mechanical degradation and system faults, while pressure and rotational speed contribute additional but comparatively smaller diagnostic information under the present dataset and feature design.

4.6. Comparison with Previous Studies

To contextualize the proposed approach, we compare it with representative studies in terms of signal types, modeling strategies, target systems, and reporting completeness, noting that direct numerical benchmarking is often unfair because datasets, operating conditions, and evaluation protocols differ across works. Fan and Liu employed adversarial neural networks for bearing fault diagnosis to improve robustness in representation learning under challenging signal conditions [8], indicating that robustness to noise and distribution shift is a key consideration beyond raw accuracy. In contrast, standards-oriented threshold setting remains a practical baseline for vibration condition monitoring, as discussed by Cristiana and Chiara [9]; such rule/threshold methods are transparent and easy to deploy, yet they can be brittle under variable loads and overlapping fault signatures, which motivates learning-based alternatives for agricultural scenarios. Within agricultural applications, Guo et al. demonstrated that supervised learning with multi-source vibration data fusion can support state recognition in agricultural robots under complex environments [10], reinforcing the value of multi-sensor information; however, their recognition targets differ from fault-type diagnosis. Monsalve et al. explored equivalent circuit models with Kalman filtering for estimating state-of-charge in agricultural robots [11], illustrating the strengths of model-based estimation in real-time feasibility and interpretability, but also highlighting limitations imposed by modeling assumptions—an important contrast to data-driven fault classification. Deep learning has also been applied directly to agricultural machinery: Michael et al. reported CNN-based condition monitoring for a disc mower use case [12], suggesting that deep architectures can capture discriminative patterns from operational signals, while our MLP-based design favors a lightweight, feature-window pipeline that is easier to deploy on constrained platforms. Beyond diagnosis, Hongzhen et al. studied panoramic-vision-based dynamic obstacle detection for moving agricultural machinery [13], underscoring the operational variability and environmental complexity that can indirectly influence monitoring signal stability and diagnostic reliability in field conditions. Related mechanical engineering work has used state-space style modeling for dynamic analysis of coupled structures [14], pointing to a complementary direction in which physics-informed representations could be integrated with learning-based diagnosis to improve interpretability and robustness. Finally, Hu et al. proposed intelligent condition assessment using multiple types of monitoring signals in industrial machinery [15], supporting the general rationale of multi-sensor monitoring adopted in this study (temperature, vibration, pressure, and rotational speed). Taken together, the literature suggests that while robust learning (e.g., adversarial training) , practical thresholding , agricultural-domain sensing and recognition, and broader environment-aware system design each contribute valuable insights, there remains a need for field-oriented agricultural machinery fault diagnosis studies that jointly report multi-fault performance, misclassification risk, and deployability-related indicators under variable workloads, which motivates the evaluation protocol and reporting emphasis adopted in this work .

5. DISCUSSION

5.1. Advantages and Limitations of the Neural Network Model

Neural-network-based approaches provide clear advantages for agricultural machinery condition monitoring and fault diagnosis, particularly in handling nonlinear relationships and multi-dimensional sensor inputs. In this study, the proposed model achieved high classification performance in recognizing both normal operating conditions and multiple fault categories, with overall accuracy exceeding 90%. Such performance indicates that the model reduced misclassification risk and supported earlier identification of abnormal patterns, which can help mitigate fault escalation and improve operational safety in agricultural production. In addition, the window-based inference strategy demonstrated near real-time feasibility for status assessment from streaming sensor data, enabling timely warnings during routine machinery operation.

Despite these advantages, neural-network models also present limitations. Model training typically requires large-scale labeled datasets, and computational costs can become substantial when the data volume increases or when more complex deep architectures are adopted. In practical applications, continuous multi-sensor acquisition can generate high-frequency data streams, increasing the burden of storage, transmission, and preprocessing. Moreover, the need for repeated training and hyperparameter tuning may limit rapid deployment in resource-constrained environments unless lightweight architectures or edge-optimized implementations are used.

5.2. Impact of Data Quality on Experimental Results

Data quality directly influences the reliability of condition monitoring and fault diagnosis. Sensor precision, data completeness, and noise levels affect feature stability and thus model performance. In this experiment, the stability of temperature, vibration, and pressure measurements affected fault classification outcomes; sensor drift and random noise, if not mitigated, could lead to biased training and degraded generalization.

Data preprocessing therefore plays a decisive role. Standardization and normalization were applied to reduce scale differences among features and to stabilize optimization during training. In addition, denoising and missing-value handling improved the consistency of the input distribution, which in turn supported more effective learning of fault-related patterns. The observed performance gains after preprocessing suggest that robust data management is a prerequisite for reliable deployment.

5.3. Model Generalization Ability and Cross-Device Application

Model generalization determines whether a diagnostic model can be applied across different machinery types and operating environments. Agricultural equipment varies widely in mechanical structure, workload profile, and environmental conditions; therefore, cross-device applicability is essential for practical adoption. In this study, the model showed effective recognition for several fault categories under the collected operating conditions; however, its training depended on the available dataset. If the dataset does not sufficiently capture variations in machine types, soil conditions, loads, and aging effects, generalization performance may be constrained.

To improve cross-device adaptability, future work should incorporate broader equipment coverage and more diverse fault patterns. Techniques such as transfer learning and data augmentation can be used to reduce the data collection burden and to improve robustness. Transfer learning, in particular, allows knowledge learned from one machine domain to be adapted to another with limited additional data, improving scalability in multi-device deployments.

5.4. Improvements for Deep Learning Models

As diagnostic requirements increase, further model development should focus on improving recognition accuracy for complex fault patterns while maintaining deployment efficiency. Although the current model can identify common faults, more advanced architectures may provide additional benefits when raw time-series signals are used. For example, Convolutional Neural Networks (CNNs) can learn local temporal features (e.g., transient vibration changes), and are well suited for signal-driven pattern extraction. Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) networks, can capture longer-term temporal dependencies and may improve discrimination when faults evolve gradually over time.

Transfer learning is also relevant for agricultural machinery fault diagnosis because it can enhance performance in low-data regimes and support adaptation across equipment types. Future efforts may combine architecture optimization (e.g., lightweight CNNs), training strategies (e.g., fine-tuning), and deployment techniques (e.g., pruning/quantization) to improve accuracy, efficiency, and robustness simultaneously.

5.5. Future Research Directions

Neural-network-based condition monitoring and fault diagnosis are expected to play a growing role in intelligent agricultural production. Future research should emphasize integrated monitoring systems that combine IoT-enabled sensing with data-driven inference to support online status assessment and early warning. For instance, sensors can stream operational data to edge or cloud platforms where trained models perform continuous diagnosis and provide maintenance recommendations, thereby reducing unplanned downtime and service costs.

Further interdisciplinary research integrating agricultural engineering, computer science, and artificial intelligence can accelerate the development of smart agriculture. In particular, combining deep learning with large-scale operational data can enable more robust fault prediction, domain adaptation across farms and seasons, and improved lifecycle management of agricultural machinery—ultimately supporting more efficient and sustainable agricultural operations.

6. CONCLUSION

This study developed a neural-network-based approach for agricultural machinery condition monitoring and fault diagnosis using multi-sensor data (temperature, vibration, pressure, and rotational speed) to recognize operating states and identify typical fault categories.

Quantitative evaluation demonstrated that the proposed MLP-based diagnostic pipeline achieved 92% overall classification accuracy, including 96% accuracy for normal-state recognition, while maintaining 83–90% accuracy across major fault classes (engine, electrical, transmission, and hydraulic faults). Relative to traditional rule-based diagnostics, the proposed method delivered a consistent improvement of approximately 11–15 percentage points depending on fault type, indicating stronger capability in modeling nonlinear and multi-dimensional patterns. In addition, the system exhibited low misclassification risk with a 2% false-positive rate and a 4% false-negative rate on the test set. The window-based inference strategy further supports near real-time feasibility for online diagnosis; reporting average inference latency per window in future deployments will make runtime performance fully reproducible.

From a practical standpoint, the proposed pipeline can be directly adopted in agricultural machinery monitoring by integrating four standard sensing channels (temperature, vibration, pressure, and rotational speed) with an onboard data acquisition unit and a lightweight inference module. In a typical deployment, sensor streams are segmented into fixed windows, window-level features are computed, and the trained model outputs (i) a health state label (normal vs. fault), (ii) a fault category when abnormal behavior is detected, and (iii) an alarm signal that can be displayed to operators or transmitted to a maintenance platform. This enables timely warnings, supports maintenance scheduling, and reduces unplanned downtime without requiring complex manual rule tuning. Future work will strengthen robustness and scalability through cross-machine transfer learning and domain adaptation, edge-oriented model compression (e.g., pruning and quantization), data-efficient learning (few-shot and self-supervised pretraining), and long-term reliability mechanisms such as sensor drift detection, online calibration, and continual learning for sustained field performance.

Data Availability Statement

All data generated or analysed during this study are included in this article.

Conflict of Interest

The authors declare that they have no competing interests.

Funding Statement

There is no funding in this article.

References

[1] D. Y. He, Q. Wang, "Rotary machinery fault recognition model based on HRFDE and GSA-PNN", Mechanical and Electrical Engineering, Vol. 40, No. 12, 2023, pp.1869-1879. https://doi.org/10.3969/j.issn.1001-4551.2023.12.005.

[2] J. S. Chen, H. Wu, X. T. Hu, X. P. Gu, H. Song, "Mechanical fault recognition method for high-voltage circuit breakers based on LMD and DE-PNN", Computer Applications and Software, Vol. 40, No. 6, 2023, pp.57-62. https://doi.org/10.3969/j.issn.1000-386x.2023.06.010.

[3] F. Y. Wang, "Ship machinery equipment bearing fault diagnosis method based on EMD and neural networks", Equipment Manufacturing Technology, No. 3, 2024, pp.138-140.

[4] X. M. Chen, M. R. Han, W. Y. Shu, K. Zhang, L. P. Li, "Fault diagnosis method based on VMD and multi-scale 1D convolutional neural networks", Modern Electronic Technology, Vol. 46, No. 9, 2023, pp.103-109. https://doi.org/10.16652/j.issn.1004-373x.2023.09.020.

[5] R. B. Gao, H. C. Ma, H. Yang, Y. C. Jian, Y. M. Wang, "Visualized fault diagnosis of gearbox sound signals based on lightweight neural networks", Electronic Devices, Vol. 47, No. 5, 2024, pp.1268-1274.

[6] J. Yuan, G. X. Ren, H. M. Jiang, Q. Zhao, C. J. Wei, J. Zhu, "Mechanical fault diagnosis method based on multivariate boosted kernel neural networks and its feature extraction interpretability", Journal of Mechanical Engineering, No. 12, 2024, pp.51-64. https://doi.org/10.3901/JME.2024.12.051.

[7] X. T. Liu, D. T. Wang, J. Sun, G. H. Chen, "Bearing fault signal extraction and classification recognition based on DCLSTM network", Mechanical Design and Research, Vol. 40, No. 3, 2024, pp.185-188.

[8] X. N. Fan, X. J. Liu, "Application of adversarial neural networks in bearing fault diagnosis", Mechanical Science and Technology, Vol. 43, No. 4, 2024, pp.690-697. https://doi.org/10.13433/j.cnki.1003-8728.20220264.

[9] C. Delprete, C. Gastaldi, "On the effectiveness of standards application to threshold setting in vibration condition monitoring in industrial machinery", Proceedings of the Institution of Mechanical Engineers, Vol. 239, No. 1, 2025, pp.207-221. https://doi.org/10.1177/1748006X2312183.

[10] J. Guo, S. Wang, Y. Mao, G. Wang, G. Wu, Y. Wu, Z. Liu, "Supervised learning study on ground classification and state recognition of agricultural robots based on multi-source vibration data fusion", Computers and Electronics in Agriculture, Vol. 219, 2024, pp.108791. https://doi.org/10.1016/j.compag.2024.108791.

[11] G. Monsalve, A. Cardenas, D. Acevedo-Bueno, W. Martinez, "Assessing the limits of equivalent circuit models and Kalman filters for estimating the state of charge: case of agricultural robots", Energies, Vol. 16, No. 7, 2023, pp.3133. https://doi.org/10.3390/en16073133.

[12] M. Jaumann, E. Olcay, T. Oksanen, "Condition Monitoring using Convolutional Neural Network in Agricultural Machinery-Use Case: Disc Mower", IFAC-PapersOnLine, Vol. 55, No. 32, 2022, pp.235-240. https://doi.org/10.1016/j.ifacol.2022.11.145.

[13] H. Xu, S. Li, Y. Ji, R. Cao, M. Zhang, "Dynamic obstacle detection based on panoramic vision in the moving state of agricultural machineries", Computers and Electronics in Agriculture, Vol. 184, 2021, pp.106104. https://doi.org/10.1016/j.compag.2021.106104.

[14] Y. Huang, Y. Li, L. Zhang, H. Zhang, Y. Gao, "Dynamic analysis of a multilayered piezoelectric two-dimensional quasicrystal cylindrical shell filled with compressible fluid using the state-space approach", Acta Mechanica, Vol. 231, No. 6, 2020, pp.2351-2368. https://doi.org/10.1007/s00707-020-02641-7.

[15] Hu, Z. Bai, J. Lin, L. Xiang, "Intelligent condition assessment of industry machinery using multiple type of signal from monitoring system", Measurement, Vol. 149, 2020, pp.107018. https://doi.org/10.1016/j.measurement.2019.107018.