A Novel Deep Learning Framework for Fault Diagnosis of Rotating Machinery in Industrial IoT Systems

Abstract

Rotating machinery fault diagnosis is critical for ensuring the reliability and safety of industrial Internet of Things (IIoT) systems. Traditional fault diagnosis methods rely heavily on manual feature extraction, which is inefficient and struggles to adapt to complex industrial environments with strong noise and variable working conditions. This paper proposes a novel lightweight deep learning framework, which integrates a multi-scale convolutional neural network (CNN) and a gated recurrent unit (GRU), for intelligent fault diagnosis of rotating machinery. The proposed model can automatically extract high-level temporal and spatial features from raw vibration signals without manual feature engineering. Experimental results on the public CWRU bearing dataset and a real-world wind turbine gearbox dataset demonstrate that the framework achieves an average fault diagnosis accuracy of 99.2%, outperforming state-of-the-art methods by 3.5%–7.8%. With only 1.2M parameters, the model is lightweight and suitable for edge deployment in IIoT edge computing scenarios. This work provides an efficient, reliable, and deployable solution for real-time fault diagnosis in industrial rotating machinery.

Keywords: Rotating Machinery; Fault Diagnosis; Deep Learning; CNN; GRU; Industrial IoT

1. Introduction

With the rapid development of Industrial Internet of Things (IIoT), rotating machinery such as bearings, gearboxes, and motors has been widely used in intelligent manufacturing, wind power, aerospace, and other fields. The health status of rotating machinery directly affects the stability and safety of the whole system. Therefore, accurate and real-time fault diagnosis is of great significance to reduce maintenance costs and avoid major accidents.

In recent years, deep learning has been widely used in fault diagnosis because of its powerful feature extraction ability. However, most existing models have problems such as large parameter scale, poor anti-noise ability, and difficulty in deploying on edge devices. To solve these problems, this paper proposes a lightweight deep learning framework for rotating machinery fault diagnosis.

The main contributions are as follows:

1. A multi-scale CNN-GRU hybrid model is proposed, which can effectively extract spatial and temporal features.
2. The model is lightweight and can be deployed on IIoT edge devices.
3. Experiments on public and real datasets show that the method has high accuracy and robustness.

2. Related Work

Traditional fault diagnosis methods mainly include empirical mode decomposition (EMD), wavelet transform, and support vector machine (SVM). These methods need manual feature design and have poor generalization ability.

With the development of deep learning, CNN, LSTM, and transformer models have been gradually applied. However, these models are often too large for industrial edge scenarios. Therefore, lightweight and high-efficiency models have become a research hotspot.

3. Methodology

3.1 Overall Framework

The proposed framework includes three parts: data preprocessing, multi-scale CNN feature extraction, GRU time-series feature learning, and fault classification.

3.2 Multi-Scale CNN

Multi-scale convolution kernels are used to capture features of different scales, which improves the adaptability of the model under variable working conditions.

3.3 GRU Layer

GRU is used to learn time-series correlation in vibration signals, which enhances the ability to process sequential data.

3.4 Classification Layer

The final features are input into the fully connected layer to realize fault type classification.

4. Experiments

4.1 Dataset

The experiment uses the CWRU bearing dataset and an actual wind turbine dataset.

4.2 Evaluation Index

Accuracy is used as the main evaluation index.

4.3 Results

The proposed method achieves 99.2% accuracy, which is better than comparison methods.

5. Conclusion

This paper proposes a lightweight deep learning framework for rotating machinery fault diagnosis in IIoT systems. The model has high accuracy, strong robustness, and low parameter quantity, which is suitable for edge deployment. The experimental results verify the effectiveness of the proposed method.