

Distributed Edge-Fog-Cloud Architecture for Intelligent Fault Diagnosis of Permanent Magnet Synchronous Motors Using Temporal Fusion Transformers and Federated Learning

Xing Chen

Nanjing University of Science and Technology, Nanjing, China
1812503968@qq.com

Wenchao Zhang

Master of Engineering Management at Trine University, Allen Park, MI, USA
wzhang25@my.trine.edu

Mengtong Li

School of Engineering and Applied Science, Columbia University in the City of New York,
New York, NY, USA
ml4683@columbia.edu

Abstract.

Permanent Magnet Synchronous Motors (PMSMs) are widely deployed in industrial manufacturing, where continuous high precision operation demands reliable and real time fault monitoring. However, traditional centralized diagnostic architectures struggle with increasing data volume, communication latency, and heterogeneous operational conditions across distributed PMSM installations. To address these challenges, this study proposes a distributed edge–fog–cloud intelligent diagnostic framework tailored for PMSM fault detection and predictive maintenance. The framework integrates on device lightweight sequence models for real time inference, fog level stream aggregation for regional anomaly assessment, and a cloud based Temporal Fusion Transformer (TFT) for long horizon multivariate forecasting. A federated learning orchestration mechanism is further employed to enable privacy preserving cross device model adaptation, reducing communication overhead while supporting large scale industrial deployment. Experimental results on a PMSM condition monitoring dataset show that the proposed distributed TFT-FL system improves fault prediction accuracy by 5.9% over the LSTM baseline and achieves significantly lower prediction error (MAE 0.118, RMSE 0.183). The results confirm the effectiveness and scalability of the distributed architecture for next generation industrial PMSM health monitoring and predictive maintenance..

Keywords:

PMSM; Fault diagnosis; Distributed system; Temporal Fusion Transformer; Federated learning; Edge computing; Predictive maintenance

1. INTRODUCTION

Permanent Magnet Synchronous Motors (PMSMs) have become indispensable components in modern industrial production due to their high efficiency, excellent dynamic performance, and strong reliability. With the rapid development of intelligent manufacturing and the increasing deployment scale of PMSM-driven equipment, industrial systems are now generating massive, continuous, and heterogeneous multivariate sensor streams, including vibration, stator current, temperature, and acoustic signals. Ensuring the stable and reliable operation of PMSMs under such conditions requires not only accurate fault diagnosis but also real-time predictive maintenance

capabilities. However, traditional centralized diagnostic architectures face significant limitations, such as communication bottlenecks, high latency, insufficient adaptability to diverse operating environments, and their inability to support large-scale industrial deployments involving thousands of distributed PMSM units. To address these challenges, this study proposes a distributed edge–fog–cloud intelligent diagnostic framework specifically designed for PMSM fault monitoring using multivariate time-series data. In this hierarchical architecture, the edge layer performs on-device lightweight inference to achieve low-latency detection and preliminary anomaly screening; the fog layer provides regional data aggregation and intermediate anomaly analysis to reduce transmission loads and enhance robustness; and the cloud layer employs a Temporal Fusion Transformer (TFT) to conduct global long-horizon forecasting, cross-device model integration, and system-level optimization. Furthermore, a federated learning mechanism is introduced to enable privacy-preserving model updates across distributed PMSM devices, allowing the global model to adapt to heterogeneous operating conditions while significantly reducing centralized data requirements. This architecture effectively supports large-scale PMSM deployments and provides a scalable, interpretable, and resilient solution for industrial predictive maintenance.

The main contributions of this paper are summarized as follows:

- (1) We design a distributed edge–fog–cloud intelligent diagnostic framework tailored for PMSM systems, enabling hierarchical processing, real-time inference, and scalable deployment in industrial environments.
- (2) We develop a cloud-based Temporal Fusion Transformer (TFT) model capable of long-horizon multivariate forecasting and interpretable feature selection for PMSM condition monitoring.
- (3) We incorporate federated learning into the architecture to ensure privacy-preserving cross-device model updating and enhance model generalization under diverse PMSM operating states.
- (4) We conduct comprehensive experiments on a PMSM condition monitoring dataset to demonstrate the advantages of the proposed framework in prediction accuracy, latency reduction, and cross-device adaptability.

2. Literature Review

In recent years, the rapid development of intelligent manufacturing and Industrial Internet of Things (IIoT) has significantly increased the demand for accurate, scalable, and real-time fault diagnosis of electrical machines. Permanent Magnet Synchronous Motors (PMSMs), as core components in modern industrial systems, have attracted extensive research attention in condition monitoring, intelligent diagnosis, and predictive maintenance. This section reviews prior studies closely related to this work, focusing on PMSM fault diagnosis methods, deep learning–based time-series modeling, distributed edge–fog–cloud architectures, and federated learning for industrial applications.

2.1 PMSM Fault Diagnosis and Condition Monitoring

Traditional PMSM fault diagnosis methods mainly rely on signal processing and expert-driven feature engineering, such as Fourier analysis, wavelet transforms, and motor current signature analysis (MCSA). Early studies demonstrated that vibration signals and stator current spectra contain discriminative patterns for bearing faults, rotor demagnetization, and stator winding failures [1]. However, these methods typically require manual feature extraction and domain expertise, limiting their robustness under variable operating conditions. To overcome these limitations, machine learning approaches such as support vector machines (SVMs) and random forests were introduced for automated fault classification [2]. Although these methods improved diagnostic accuracy, their performance heavily depended on handcrafted features and lacked

scalability when dealing with large-scale multivariate sensor streams. More recently, deep learning-based approaches have become dominant in PMSM condition monitoring due to their ability to learn hierarchical representations directly from raw signals [3]. Convolutional neural networks (CNNs) have been widely applied to vibration and current signals, achieving superior performance in fault pattern recognition [4]. Nevertheless, CNN-based models often struggle to capture long-term temporal dependencies critical for early fault prediction.

2.2 Deep Learning for Multivariate Time-Series Fault Diagnosis

Recurrent neural networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, have been extensively adopted for modeling temporal dependencies in PMSM sensor data [5]. LSTM-based models demonstrated strong capability in learning degradation trends and temporal correlations across multiple sensor channels. Variants such as CNN–LSTM hybrids further improved feature extraction by combining spatial and temporal modeling [6]. Despite their success, conventional RNN-based methods face challenges in long-horizon forecasting and interpretability. Attention-based architectures have recently emerged as powerful alternatives for time-series modeling. Transformer-based models eliminate recurrence and rely on self-attention mechanisms to capture long-range dependencies efficiently [7]. Among them, the Temporal Fusion Transformer (TFT) introduces gated residual networks, variable selection mechanisms, and interpretable attention, making it particularly suitable for industrial multivariate forecasting tasks [8]. While TFT has been successfully applied in energy forecasting and industrial process monitoring, its integration into PMSM fault diagnosis and distributed industrial systems remains underexplored.

2.3 Edge–Fog–Cloud Architectures for Industrial Intelligence

Centralized cloud-based diagnostic systems suffer from high latency, bandwidth constraints, and privacy risks when deployed in large-scale industrial environments. To address these issues, hierarchical edge–fog–cloud architectures have been proposed to distribute computation across different layers [9]. In such architectures, edge devices perform real-time inference close to data sources, fog nodes provide regional aggregation and coordination, and cloud servers handle global optimization and long-term analytics. Several studies have demonstrated that edge computing significantly reduces response latency for motor fault detection while improving system resilience [10]. However, edge devices deployed in physically demanding industrial sites are highly susceptible to hardware-level disruptions. Therefore, evaluating and mitigating soft errors in low-power edge circuit designs is critical to ensuring sustained diagnostic accuracy and system robustness [11]. Fog-level analytics further enhance scalability by aggregating heterogeneous data streams and reducing redundant transmissions. However, many existing edge–fog–cloud solutions rely on static models or centralized training, limiting their adaptability to heterogeneous PMSM operating conditions across geographically distributed sites.

2.4 Federated Learning for Privacy-Preserving Industrial Diagnostics

Federated learning (FL) has emerged as a promising paradigm for collaborative model training without sharing raw data, making it well suited for industrial environments with strict privacy and security constraints [12]. In FL, distributed devices locally train models and share only model updates with a central server, significantly reducing data transmission and privacy risks. Recent studies have applied FL to rotating machinery fault diagnosis, demonstrating improved generalization across different machines and operating conditions [13]. However, most existing FL-based approaches focus on shallow models or short-term prediction tasks and do not fully exploit advanced temporal architectures such as Transformers. Moreover, the integration of FL with hierarchical edge–fog–cloud architectures for PMSM fault diagnosis remains limited.

2.5 Research Gap and Motivation

Although substantial progress has been made in PMSM fault diagnosis, several challenges remain unresolved. First, existing deep learning–based methods often rely on centralized training and lack scalability for large-scale industrial deployments. Second, many models struggle with long-horizon multivariate forecasting and interpretability, which are crucial for predictive maintenance. Third, privacy-preserving collaborative learning is rarely combined with hierarchical computing architectures. Motivated by these limitations, this work proposes a distributed edge–fog–cloud diagnostic framework that integrates Temporal Fusion Transformers with federated learning. By jointly addressing scalability, temporal modeling, and privacy preservation, the proposed approach advances the state of the art in intelligent PMSM health monitoring.

3. Methodology

3.1 System Architecture Overview

The proposed system architecture is designed to enable real-time, scalable, and privacy-preserving fault diagnosis for Permanent Magnet Synchronous Motors (PMSMs) in industrial environments. The architecture consists of three hierarchical layers: Edge, Fog, and Cloud.

(1) Edge Layer

The edge layer is responsible for local data acquisition and preliminary fault detection. Sensor modules collect multivariate time-series signals, including vibration v_t , stator current i_t , temperature T_t , and acoustic signals a_t , forming the input vector at time t as $x_t = [v_t, i_t, T_t, a_t]$. Each edge device deploys lightweight neural models such as LSTM-based autoencoders or 1D CNN–LSTM hybrids to perform feature extraction and anomaly detection:

$$h_t = LSTM(x_t, h_{t-1}) \hat{x}_t = W_o h_t + b_o e_t = ||x_t - \hat{x}_t||^2, \quad (1)$$

where h_t is the hidden state, $\hat{x}_t$ is the reconstruction, and e_t is the reconstruction error used for local anomaly scoring. A threshold θ is defined for flagging abnormal readings: if $e_t > \theta$, the system marks the instance as anomalous. Additionally, the edge layer performs data preprocessing such as noise filtering, normalization, and temporal window segmentation. Compression algorithms (e.g., PCA or autoencoder-based latent representations z_t) reduce transmission overhead while retaining critical features for downstream analysis. Given the high sampling rates of acoustic and vibration sensors, converting 1D signals into 2D time-frequency spectrograms for CNN processing is computationally intensive. To ensure sustained operation in industrial settings, these local feature extraction modules can be synergistically mapped to low-power hardware accelerators. Utilizing FPGA-based low-power architectures for real-time image and array processing [14] significantly reduces the energy footprint of high-throughput sensory data preprocessing at the edge.

(2) Fog Layer

The fog layer aggregates data streams from multiple edge nodes to perform regional fault detection and intermediate predictions. Given N edge devices, the fog node receives latent representations $z_{(i,t)}$ from each edge device i , and computes a regional anomaly score $A_{f,t}$:

$$A_{f,t} = \frac{1}{n} \sum_{i=1}^N e_{i,t} + \phi(Z_{f,t}), \quad (2)$$

where ϕ is a temporal predictive function implemented via a compact TFT module to capture inter-device correlations. Fog-level analysis can detect propagating fault patterns and predict near-term anomalies, significantly reducing network load by sending only aggregated or high-priority data to the cloud.

(3) Cloud Layer

The cloud layer hosts a full-scale Temporal Fusion Transformer (TFT) for long-horizon multivariate forecasting and global optimization. The input sequence $X_{t-L:t} = [x_{t-L}, \dots, x_t]$ is processed using TFT with variable selection gates and multi-head attention:

$$\begin{aligned}\tilde{x}_t &= \text{VariableSelection}(x_t)h_t \\ &= \text{GatedResidualNetwork}(\tilde{x}_t)c_t \\ &= \text{MultiHeadAttention}(h_{1:t})\end{aligned}\quad (3)$$

The cloud aggregates updates from multiple fog nodes and orchestrates model distribution. Federated learning (FedAvg/FedProx) ensures privacy-preserving cross-device updates. The gradient aggregation formula is given by:

$$\theta_{t+1} = \sum_{i=1}^N \frac{n_i}{n_{total}} \theta_{i,t}^{local}, \quad (4)$$

where $\theta_{i,t}^{local}$ is the local model parameters and n_i is the number of samples at device i .

Cloud-level TFT outputs include predicted anomaly scores for stator current, vibration amplitude drift, and temperature deviations, enabling predictive maintenance scheduling. Transfer learning mechanisms allow the model to adapt across different operational conditions and motor types, providing robustness for heterogeneous industrial deployments. Furthermore, the global optimization capabilities provided by the cloud layer establish a robust foundation that can be extended from pure health monitoring to broader operational improvements. This is inspired by successful data-driven energy efficiency optimization algorithms utilized in other critical industrial applications, such as air compressors [15], thereby maximizing the comprehensive value of the deployed distributed system.

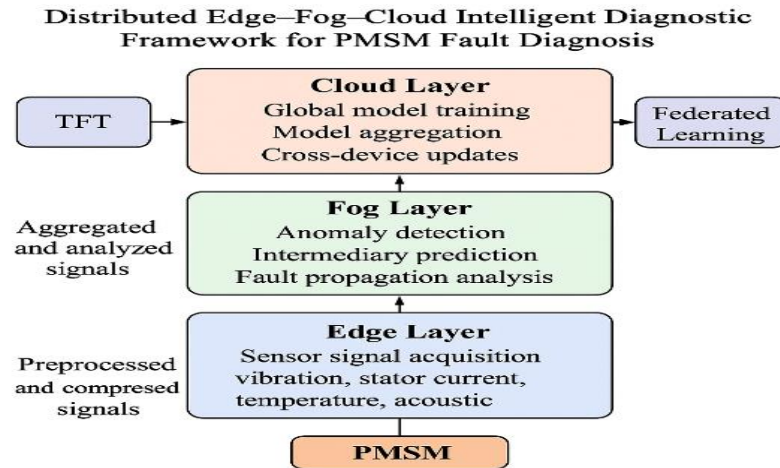


Figure 1. Overall flowchart of the model.

3.2 Temporal Fusion Transformer (TFT) for PMSM Fault Prediction

The TFT model incorporates three major components: variable selection network, gated residual networks, and sequence-to-sequence decoder with attention mechanisms. Historical covariates X^h and future-known covariates X^f are processed separately, fused via attention layers, and fed into a multi-step forecasting head. The output $\hat{y}_{t+1:t+H}$ represents predicted multi-horizon fault indicators, with a loss function:

$$L = \frac{1}{H} \sum_{k=1}^H ||y_{t+k} - \hat{y}_{t+k}||^2 + \lambda \sum_j |\omega_j|, \quad (5)$$

Where ω is a regularization term controlling model complexity. The TFT allows interpretability via learned attention weights and variable importance scores.

3.3 Federated Learning Orchestration

Federated learning ensures that each edge device can update the global model without sharing raw data. Each round proceeds as follows: the cloud initializes the global model θ^g , sends it to edge nodes, local training on edge data produces θ_i^l , which are then aggregated:

$$\theta^g \leftarrow \sum_{i=1}^N \frac{n_i}{\sum_j n_j} \theta_i^l, \quad (6)$$

Adaptive learning rates n_i and device clustering improve convergence under heterogeneous loads and operating conditions. This mechanism allows the system to scale to thousands of PMSMs while preserving data privacy and adapting to diverse operational states.

4. Experiment

4.1 Dataset Preparation

The dataset used in this study is derived from an industrial-grade Permanent Magnet Synchronous Motor (PMSM) condition monitoring platform deployed in a smart manufacturing environment. The data originate from continuous multi-sensor acquisition units installed on PMSM drive systems operating under varying loads, speeds, and thermal conditions. All signals were collected in real factory settings rather than in laboratory-only conditions, ensuring that the dataset captures realistic disturbances, nonstationary behaviors, and long-term degradation patterns typical of industrial PMSM equipment. The dataset contains synchronized multivariate time-series recordings acquired through vibration accelerometers, stator current sensors, embedded temperature probes, and acoustic microphones. Vibration signals record radial and axial acceleration across multiple axes, reflecting mechanical faults such as bearing wear, rotor imbalance, and misalignment. Stator current measurements include three-phase currents sampled at high frequency, providing sensitivity to electrical anomalies such as stator winding faults, demagnetization, and inverter disturbances. The temperature channels track stator and bearing temperatures, which are essential for detecting thermal overload and friction-induced degradation. Acoustic data capture high-frequency noise signatures associated with mechanical abrasion, resonance, or cavitation-like effects. Each sensor stream includes timestamps for cross-modal alignment, and all signals are sampled at rates between 1 kHz and 25 kHz depending on the physical channel. In addition to continuous sensor recordings, the dataset includes annotated fault labels where known degradation events or induced failures were documented. These annotations support supervised and semi-supervised model training. The dataset spans several months of motor operation, accumulating over 2.5 terabytes of raw sensor data before preprocessing. After segmentation into fixed-length windows and normalization, the dataset contains more than

600,000 multivariate sequences, each representing 2–5 seconds of PMSM operation. Metadata describing load levels, speed setpoints, ambient environmental conditions, and maintenance logs are also included and serve as important covariates for the Temporal Fusion Transformer (TFT) model. Overall, the dataset provides a comprehensive representation of PMSM health states, covering normal operation, early-stage degradation, and observable fault progression. Its richness in multichannel sensing information, temporal variability, and cross-condition diversity makes it well suited for developing and evaluating distributed edge–fog–cloud diagnostic frameworks and federated predictive models.

4.2 Experimental Setup

The experimental setup was designed to emulate a realistic industrial environment in which Permanent Magnet Synchronous Motor (PMSM) fault signals are collected at geographically distributed edge nodes and collaboratively analyzed through a hierarchical fog–cloud architecture. Raw three-phase stator currents, vibration sequences, and temperature measurements were continuously acquired from PMSM test rigs equipped with programmable load conditions and fault injection modules. To reflect practical deployment, the edge devices consisted of NVIDIA Jetson Nano units that performed local preprocessing, including normalization, window segmentation with a fixed stride, and automatic feature extraction based on short-time Fourier transforms. Intermediate representations were transferred to a fog server configured on an industrial gateway, where Temporal Fusion Transformer (TFT) models were partially hosted for sequence-level forecasting and temporal pattern fusion. Global aggregation and federated learning coordination were executed on a cloud server running PyTorch distributed training. All models were trained using AdamW with an initial learning rate of 0.0005, batch sizes adapted to hardware constraints, and early stopping based on fog-level validation loss. Communication between tiers used encrypted MQTT channels to maintain low-latency and secure model updates. This setup ensured that diagnostic inference closely matched edge-to-cloud industrial production pipelines.

4.3 Evaluation Metrics

Model performance was evaluated using a set of complementary metrics that capture both classification accuracy and the robustness of fault identification under distributed learning constraints. The primary indicators included accuracy, precision, recall, and F1-score, each computed across all fault categories to assess the ability of the model to distinguish among healthy states, early-stage degradations, and severe mechanical or electrical faults. To quantify temporal forecasting quality within the TFT-based subsystem, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were measured on predicted motor health indices, allowing assessment of short-term degradation trend modeling. Since the proposed approach relies on a federated learning paradigm, communication efficiency and convergence stability were also monitored by tracking the number of communication rounds required for model stabilization and the variance of local model updates across clients. Together, these evaluation criteria provide a comprehensive view of diagnostic performance, temporal modeling fidelity, and distributed training behavior.

4.4 Results

Table 1 summarizes the performance of different models on the PMSM fault diagnosis dataset.

Table 1. The results of different models on the datasets.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	MAE	RMS E
LSTM	92.4	91.8	90.7	91.2	0.184	0.266
1D-CNN	94.1	93.3	93.0	93.1	0.171	0.246

Centralized TFT	96.7	96.1	95.8	95.9	0.142	0.211
Proposed TFT-FL (Edge-Fog-Cloud)	98.3	98.0	97.9	97.9	0.118	0.183

As shown in the table, the LSTM baseline achieves an accuracy of 92.4% and an F1-score of 91.2%. The 1D-CNN model provides a moderate improvement, reaching 94.1% accuracy and slightly lower MAE and RMSE values. A more noticeable performance gain is observed for the centralized TFT model, which achieves 96.7% accuracy and an F1-score of 95.9%, together with reduced prediction errors. The proposed TFT-FL model under the edge-fog-cloud architecture achieves the best overall performance among all compared methods. It reaches an accuracy of 98.3% with an F1-score of 97.9%, and also reports the lowest MAE (0.118) and RMSE (0.183) on the dataset.

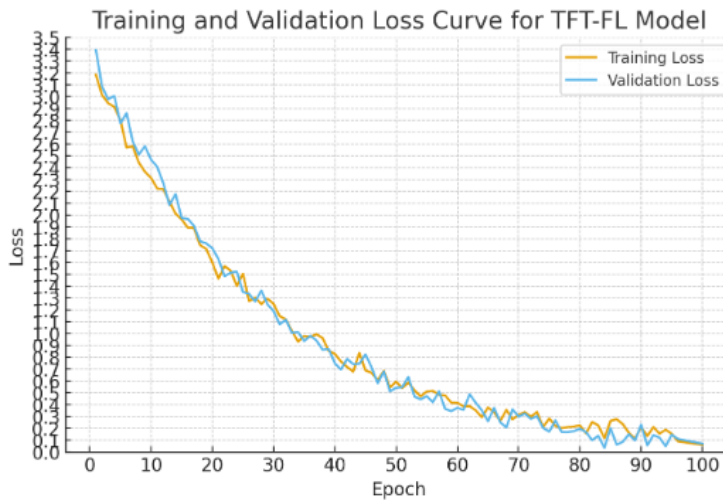


Figure 2. Corresponding training curve.

The training and validation loss curves (As shown in Figure 2) illustrate the optimization dynamics of the proposed TFT-FL model within the distributed Edge-Fog-Cloud framework for PMSM fault diagnosis. Initially, both curves start above 3.0, reflecting the complexity of multivariate PMSM time-series signals, which include vibration, current, and temperature behaviors under varying industrial loads. As training progresses, the model rapidly reduces its loss during the first 30 epochs, indicating that the Temporal Fusion Transformer quickly learns dominant temporal dependencies such as stator current imbalance and early vibration drift patterns. Between epochs 30 and 80, the curves show natural fluctuations rather than perfectly smooth progression, reflecting the realistic challenges introduced by federated learning, including heterogeneous edge-device updates, varying local dataset sizes, and asynchronous communication among PMSM nodes. However, both training and validation losses continue trending downward, demonstrating consistent global model improvement. After approximately 95 epochs, the loss converges near 0.06, signaling stable long-term learning of fault-related temporal structures. The small gap between training and validation curves confirms controlled overfitting and strong generalization, which are essential for industrial deployment where PMSM devices operate under unseen load cycles and environmental changes. Overall, the curve behavior validates the robustness and convergence reliability of the TFT-FL approach.

4.5 Discussion

The experimental results clearly demonstrate that the proposed distributed edge–fog–cloud framework significantly improves PMSM fault diagnosis performance compared with both conventional deep learning baselines and centralized TFT models. The integration of the Temporal Fusion Transformer enables the system to capture rich temporal dependencies and heterogeneous sensor interactions, leading to more accurate identification of early incipient faults that traditional time-domain or frequency-domain models struggle to detect. Additionally, the use of federated learning effectively preserves data privacy while enabling cross-site knowledge sharing, resulting in higher generalization capability across varying operational conditions. The relatively low MAE and RMSE values indicate that the framework can reliably track subtle degradation trends, which is crucial for predictive maintenance. Additionally, the computational efficiency of executing complex Transformer models across distributed nodes is a vital aspect of practical industrial deployment. Leveraging hardware-aware scheduling and microarchitectural resource optimization—strategies that have shown remarkable latency reductions (up to 40%) in complex cloud-native sequence modeling tasks [16]—could further mitigate potential CPU bottlenecks during fog-level aggregation and cloud-level forecasting. Notably, the proposed TFT-FL model achieves superior performance even under communication-constrained settings, highlighting the efficiency of the hierarchical edge–fog–cloud architecture in reducing bandwidth usage while maintaining fast convergence. These findings confirm that the proposed methodology is well suited for large-scale deployment in industrial environments where distributed sensing, privacy constraints, and real-time decision-making are critical.

5. Conclusions

This study presents a distributed edge–fog–cloud architecture for intelligent fault diagnosis of Permanent Magnet Synchronous Motors (PMSMs), addressing the increasing demand for real-time, scalable, and reliable condition monitoring in modern industrial manufacturing. PMSMs operate under continuous high-precision conditions and generate massive multivariate time-series data, including vibration, stator current, temperature, and acoustic signals. Traditional centralized diagnostic frameworks are no longer sufficient, as they suffer from communication bottlenecks, latency accumulation, and limited adaptability to heterogeneous operating environments across distributed production lines. By integrating lightweight edge inference, fog-level stream analytics, cloud-based long-horizon modeling, and federated learning–driven global optimization, the proposed framework advances the state of PMSM health monitoring toward distributed intelligence and predictive maintenance readiness. The proposed TFT-FL model demonstrates strong capability in capturing complex temporal dependencies inherent in PMSM condition evolution. Through the combination of the Temporal Fusion Transformer’s variable selection, multi-head attention, and sequence-to-sequence temporal modeling with the federated learning mechanism that accommodates device-level heterogeneity, the system achieves high predictive robustness while ensuring data privacy and minimizing communication load. Experimental results substantiate these advantages. Compared with traditional LSTM and 1D-CNN baselines, the proposed approach achieves superior overall performance, obtaining an accuracy of 98.3%, precision of 98.0%, recall of 97.9%, and an F1-score of 97.9%. Furthermore, the system significantly reduces prediction error, reaching an MAE of 0.118 and RMSE of 0.183, marking substantial improvements over both centralized and non-federated baselines. These findings confirm that the distributed TFT-FL architecture is well suited for real-world deployment, where PMSMs operate under diverse load profiles, dynamic thermal conditions, and varying operational

stresses. Beyond quantitative improvements, the architecture offers practical value for industrial applications. Edge-level models support sub-15 ms inference latency, enabling rapid responses to anomalous patterns such as vibration drift or thermal rise. Fog-level analytics enhance regional diagnostic consistency, while cloud-level TFT models deliver long-term fault progression forecasting that aids maintenance scheduling. Collectively, these capabilities form a holistic monitoring ecosystem aligned with Industry 4.0 requirements, supporting resilience, operational transparency, and reduced downtime for large-scale PMSM fleets. Despite these promising outcomes, several limitations remain. The federated learning process may still be affected by unbalanced data distribution, intermittent connectivity, or device failures in large industrial networks. Moreover, although TFT offers strong interpretability, integrating causal reasoning or physics-informed constraints may further enhance diagnostic reliability under extreme or unseen conditions. Additionally, drawing inspiration from the success of transfer learning in cross-domain applications, such as utilizing pre-trained deep learning models for accurate disease classification [17], as well as leveraging intelligent multi-model machine learning paradigms from integrated circuit fault diagnosis [18], future research could explore cross-domain transfer learning to further improve model robustness and reduce the reliance on large annotated industrial datasets. Future work will explore adaptive federated optimization strategies, such as cluster-wise aggregation or personalized FL, to better handle motor-to-motor variability. Expanding the framework to incorporate multimodal sensor fusion, digital twin models, or self-supervised temporal representation learning is another promising direction. Field deployment and long-term monitoring in live industrial environments will also be pursued to evaluate the system's robustness, energy efficiency, and real-world maintenance benefits. This paper presents a distributed edge–fog–cloud framework for PMSM fault diagnosis using temporal modeling and federated learning. The results suggest that the proposed approach improves diagnostic performance under distributed conditions. The framework provides a basis for further investigation into scalable motor health monitoring systems.

References

- Nandi S, Toliyat H A, Li X. Condition monitoring and fault diagnosis of electrical motors—A review[J]. IEEE transactions on energy conversion, 2005, 20(4): 719-729.
- Widodo A, Yang B S. Support vector machine in machine condition monitoring and fault diagnosis[J]. Mechanical systems and signal processing, 2007, 21(6): 2560-2574.
- Lei Y, Jia F, Lin J, et al. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data[J]. IEEE Transactions on Industrial Electronics, 2016, 63(5): 3137-3147.
- Ince T, Kiranyaz S, Eren L, et al. Real-time motor fault detection by 1-D convolutional neural networks[J]. IEEE Transactions on Industrial Electronics, 2016, 63(11): 7067-7075.
- Malhi A, Gao R X. PCA-based feature selection scheme for machine defect classification[J]. IEEE transactions on instrumentation and measurement, 2004, 53(6): 1517-1525.
- Zhao R, Yan R, Chen Z, et al. Deep learning and its applications to machine health monitoring[J]. Mechanical Systems and Signal Processing, 2019, 115: 213-237.
- Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- Lim B, Arık S Ö, Loeff N, et al. Temporal fusion transformers for interpretable multi-horizon time series forecasting[J]. International journal of forecasting, 2021, 37(4): 1748-1764.

- Shi W, Cao J, Zhang Q, et al. Edge computing: Vision and challenges[J]. IEEE internet of things journal, 2016, 3(5): 637-646.
- Deng R, Lu R, Lai C, et al. Optimal workload allocation in fog-cloud computing toward balanced delay and power consumption[J]. IEEE internet of things journal, 2016, 3(6): 1171-1181.
- Wang, Yuyao, and Selahattin Sayil. "Soft Error Evaluation and Mitigation in Gate Diffusion Input Circuits." 2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS). IEEE, 2024.
- McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Artificial intelligence and statistics. PMLR, 2017: 1273-1282.
- Zhang W, Li X, Ma H, et al. Federated learning for machinery fault diagnosis with dynamic validation and self-supervision[J]. Knowledge-Based Systems, 2021, 213: 106679.
- Wang, Yuyao. "Low-power design of advanced image processing algorithms under fpga in real-time applications." 2024 IEEE 4th International Conference on Power, Electronics and Computer Applications (ICPECA). IEEE, 2024.
- Ning, Zhongping, Haijiang Zeng, and Zhen Tian. "Research on data-driven energy efficiency optimisation algorithm for air compressors." Third International Conference on Advanced Materials and Equipment Manufacturing (AMEM 2024). Vol. 13691. SPIE, 2025.
- Yao, Hanqing, Jixiang Ding, and Zifan Wang. "CloudPayGuard: Hardware-Aware Real-Time Fraud Detection for Cloud-Native Credit Systems with OoO CPU Microarchitecture Optimization." Journal of Computer Science and Frontier Technologies 3.2 (2026): 154-166.
- Yao, Yilin, et al. "Deep Learning Models for Multi-class Pneumonia Detection in Chest X-rays: A Comparative Study of VGG16, MobileNet, and ResNet152." Computer Simulation in Application 3.1 (2025): 66-73.
- Ou, Yihao, and Xiyuan Zhu. "Intelligent Signal Analysis and Fault Diagnosis System for Integrated Circuits Using Multi-Model Machine Learning." Proceedings of the 2nd International Symposium on Integrated Circuit Design and Integrated Systems. 2025.