Predicting And Detecting Faults In Industrial Machines By Iot System Using CNN And GRU Model

Gnanaprakash J[1], Gobichandar M[2], Meenatchi K[3], Praveena G[4], JOSEPH S[5]
[1, 2] Dept of Electronics and Communication Engineering
[5]Assistant Professor, Dept of Electronics and Communication Engineering
[1, 2, 3, 4, 5] Christian College of Engineering and Technology, Oddanchatram ,Dindigul.

Abstract- *Fault detection in industrial systems is crucial for ensuring operational safety, minimizing downtime, and reducing maintenance costs. This work proposes a hybrid deep learning model combining Convolutional Neural Networks (CNN) and Gated Recurrent Units (GRU) to detect and classify machine faults from time-series data. The CNN layers extract spatial features, while GRU layers model temporal dependencies in the data.The architecture incorporates residual connections to enhance gradient flow and improve learning efficiency. The model is evaluated on multi-class fault detection datasets, achieving robust performance with high accuracy, precision, recall, and F1-score. Advanced metrics, including ROC- AUC, logarithmic loss, Cohen's Kappa, and Matthews Correlation Coefficient, demonstrate the model's reliability. Visualization of confusion matrices and detailed performance metrics validates its effectiveness in detecting anomalies and classifying fault types. This approach can be generalized for real-time monitoring systems in various industrial applications, ensuring predictive maintenance and operational excellence.*

Keywords- IOT, AI, Hybrid Deep learning.

I. INTRODUCTION

Fault detection, isolation, and recovery (FDIR) is a subfield of control engineering which concerns itself with monitoring a system, identifying when a faulth as occurred, and pinpointing the type of fault and its location. Two approaches can be distinguished: A direct pattern recognition of sensor readings that indicate a fault and an analysis of the discrepancy between the sensor readings and expected values, derived from some model. In the latter case, it is typical that a fault is said to be detected if the discrepancy or residual goes above a certain threshold. It is then the task of fault isolation to categorize the type of fault and its location in the machinery. Fault detection and isolation(FDI) techniques can be broadly classified into two categories. These include model-based FDI and signal processing based FDI.

Example of model-based FDI logic for an actuator in an aircraft elevator control system. In model-based FDI techniques some model of the system is used to decide about the occurrence of fault. The system model may be mathematical or knowledge based. Some of the model-based FDI techniques include observer-based approach, parity- space approach, and parameter identification-based methods. There is another trend of model-based FDI schemes, which is called set-membership methods. These methods guarantee the detection of fault under certain conditions. The main difference is that instead of finding the most likely model, these techniques omit the models, which are not compatible with data.

The example shown in the figure on the right illustrates a model-based FDI technique for an aircraft elevator reactive controller through the use of a truth table and a state chart. The truth table defines how the controller reacts to detected faults, and the state chart defines how the controller switches between the different modes of operation (passive, active, standby, off, and isolated) of each actuator. For example, if a fault is detected in hydraulic system 1, then the truth table sends an event to the state chart that the left inner actuator should be turned off. One of the benefits of this model-based FDI technique is that this reactive controller can also be connected to a continuous- time model of the actuator hydraulics, allowing the study of switching transients.

With the research advances in ANNs and the advent of deep learning algorithms using deep and complex layers, novel classification models have been developed to cope with fault detection and diagnosis. Most of the shallow learning models extract a few feature values from signals, causing a dimensionality reduction from the original signal. By using Convolutional neural networks, the continuous wavelet transform scalogram can be directly classified to normal and faulty classes. Such a technique avoids omitting any important fault message and results in a better performance of fault detection and diagnosis. In addition, by transforming signals to image constructions, 2DConvolutional neural networks can be implemented to identify faulty signals from vibration image features.

Deep belief networks, Restricted and Autoencoders are other deep neural networks architectures which have been successfully used in this field of research. In comparison to traditional machine learning, due to their deep architecture, deep learning models are able to learn more complex structures from datasets, however, they need larger samples and longer processing time to achieve higher accuracy.

With the increasing complexity of modern industrial systems, ensuring the reliability and safety of machinery has become a critical challenge. Unexpected machine failures can lead to significant economic losses, safety hazards, and unplanned downtime, impacting operational efficiency. Traditional fault detection methods often rely on manual inspections or rule-based systems, which are limited by human error, scalability issues, and the inability to adapt to complex patterns in data. The rise of advanced deep learning techniques presents an opportunity to automate and enhance fault detection processes, leveraging the rich information embedded in time-series data. By integrating Convolutional Neural Networks (CNNs) for feature extraction and Gated Recurrent Units (GRUs) for temporal analysis, this study aims to provide a robust solution capable of identifying subtle patterns and anomalies. The motivation stems from the need for a scalable, accurate, and reliable system that can support real-time fault detection and classification, ultimately driving predictive maintenance and improving industrial productivity.

The primary objective of this study is to develop a deep learning-based system for accurate and efficient machine fault detection and classification using time-series data. By combining Convolutional Neural Networks (CNNs) for extracting spatial features and Gated Recurrent Units (GRUs) for capturing temporal dependencies, the model aims to identify subtle patterns and anomalies indicative of potential machine faults. This approach seeks to overcome the limitations of traditional methods by automating the fault detection process, reducing human intervention, and enhancing prediction accuracy. Additionally, the study aims to evaluate the performance of the proposed model in real-world scenarios, ensuring its scalability and applicability for predictive maintenance in industrial environments. Ultimately, the goal is to enable timely fault detection, minimize downtime, and improve overall operational efficiency.

Survey

G. Wang et al proposed a novel hybrid model by combining one dimensional CNN (1DCNN) and support vector machine (SVM) for bearing fault diagnosis and classification, which is called 1DCNN-SVM model or method. The first part of the model, 1DCNN, is used to extract features of raw data, whereas the second part, SVM, is employed to carry out the fault recognition and classification tasks. The evaluation of the proposed method is operated based on a real dataset.

Sreekumar K et al used three-phase synchronous generators with 3kVA and 5kVA ratings, for detecting the stator inter-turn short circuit faults. Baseline-system uses statistical features derived from the current signal, with a support vector machine (SVM) used as a backend- classifier. In the first approach, consider the system specific attributes as a nuisance and remove it using nuisance attribute projection (NAP) algorithm. We obtained a performance improvement of 19.63%, 8.31% and 11.80% classification accuracy for R, Y, B phases respectively, over the baseline-system. When we use a SVM backend-classifier, it is required that the features match the kernels used with the SVM.

Y. Wang et al three commonly seen ML techniques, including Support Vector Machine (SVM), Random Forest (RF), and Neural Network (NN) are investigated for vibration-based motor fault diagnosis. Corresponding SHAP explanation methods are applied to the three ML techniques to discover the most important vibration features in detecting motor conditions and differentiating faults. Explanation results from the three ML techniques demonstrate great consensus: average vibration frequency contributes most to motor fault diagnosis.

Hariharan V et al develop and assess the models and identify textile flaws using the TEXFED dataset, which is open to the public. CNN performs better than SVM and RF regarding accuracy and detection rate. In contrast to SVM and RF, which had lower accuracy and detection rates, CNN showed higher results. The findings show that the CNN algorithm outperforms SVM and RF in accuracy and detection rate, making it an appropriate method for automatically detecting textile problems. The results of the study can be helpful to the textile sector in developing practical and precise automated fault identification systems.

Y. Yinghua, et al proposes the method of deep learning (DL) to solve the above problems. Combined with the bearing signal one-dimensional characteristics, this paper proposes a one-dimension deep convolutional neural network(1D-DCNN). First of all, the original bearing vibration signal is directly input the 1D-DCNN frame structure, then 1D-DCNN frame structure is used to automatic feature extraction. Next, we use softmax regression to classified fault samples and normal samples, the confusion matrix shows that the accuracy of the 1D- DCNN model for fault diagnosis.

M. Kavianpour et al study, DenseNet201, ResNet152V2 and, MobileNetV2 are chosen as DTCNN models for feature extraction. Firstly, vibration signals are converted into time-frequency RGB images by continuous wavelet transform (CWT). Then, the high-level features of images are extracted by the DTCNN models. Finally, different types of bearing faults are classified by DL and ML classifiers. The experiment is validated on the famous Case Western Reserve University (CWRU) bearing data set. The result demonstrates that the proposed DTCNN models achieve the best accuracy rate in the classification task and are faster to learn than many other existing DL and ML models.

W. Zhang, et al proposed a novel DL framework by applying convolutional neural networks (CNNs) based on the optimization of transfer learning (TL). TL can help the model achieve higher precision with less computational cost by transferring low-level features and fine- tuning high-level layers. In addition, data processing was implemented using continuous wavelet transformation (CWT) to convert vibration signals into 2-D images, and support vector machines (SVM) were employed to replace the fully connected layers for better classification.

H. Liu, et al introduces a novel intelligent FD algorithm with three stages for MHM. It is a hybrid framework that combines generative adversarial networks (GANs) and autoencoder (AE) based on the bidirectional long short-term memory (bi-LSTM). First, GAN is employed to obtain the reconstruction residual and learn the discriminative representation. Then, AE is used to perform the critical temporal features extraction and dimension reduction. Finally, the supervised learning model is constructed to integrate feature information and predict diagnosis results. To verify the effectiveness of the proposed algorithm, typical rolling bearing datasets are taken as trial data. Preliminary simulation results demonstrate that the proposed algorithm achieves superior performance compared to the competing methods.

Q. Li et al A rolling bearing fault diagnostic CAR-SNN model based on Convolutional Block Attention Module (CBAM), Residual Network (ResNet), and Spiking Neural Networks (SNNs) is proposed based on deep learning (DL). The CAR-SNN model combines CBAM, ResNet, and SNNs to extract critical features while avoiding gradient vanishing or explosion phenomena. The ability of the CAR-SNN model to process spatio-temporal information is improved by using SNNs that incorporate spatio-temporal information simultaneously

M. D. Choudhury, et al a fault diagnosis method combining two-dimensional (2D) image representations of squared envelop spectrum (SES) of vibration signals of bearings, a critical machine element, and a pretrained convolutional neural network (CNN) is proposed. SES is one of the most efficient indicators for the assessment of second order cyclostationary symptoms of damages, which are typically observed in bearings. In this paper, we integrate this knowledge in designing a DL framework for efficient fault diagnosis in bearings.

J. Liang, et al proposes an unsupervised deep transfer learning (DTL) method with an isolation forest (iForest) for machine fault diagnosis. First, the isolation forest is used to classify and label the samples automatically; then, these labeled data are used to train deep learning (DL) models; finally, small data with the label of the target domain are used to fine-tune parameters and complete the fault diagnosis. The proposed approach has been validated with the fan gearbox dataset, the bearing dataset, and the ball screw dataset. The results show that the proposed unsupervised deep transfer learning model has high accuracy and generality.

C. B. S. Maior, et al proposes a Quantum Machine Learning (QML) approach to diagnose rolling bearings, which are essential components in rotating machinery, based on vibration signals. We apply hybrid models involving the encoding and construction of parameterized quantum circuits (PQC) connected to a classical neural network, the Multi-Layer Perceptron (MLP). We consider combinations of the Variational Quantum Eigensolver (VQE) framework with rotation gates and different entanglement (two-qubits) gates (CNOT, CZ and iSWAP). For each PQC configuration, we assess the impact of the number of layers (1, 5 and 10).

S. Dixit, et al proposes a novel conditional auxiliary classifier GAN framework coupled with model agnostic meta learning (MAML) to resolve this problem. The objective is to initialize and update the network parameters using MAML instead of regular stochastic gradient learning. This modification enables GAN to learn the task of synthetic sample generation using the limited training dataset. The effectiveness of the proposed framework has been compared with several famous state-of-the-art intelligent fault diagnosis methods existing in the literature.

X. Zhao, et al a new multiple-order graphical deep extreme learning machine (MGDELM) algorithm for unsupervised fault diagnosis (UFD) of rolling bearing is designed in this study. By jointly optimizing the multiple-order objective function, the proposed MGDELM algorithm can synchronously extract local and global structural information from the raw industrial data. Empirically, rolling bearing failure data validates the effectiveness of the designed algorithm and fault diagnosis method.

N. Khoshavi, et al propose Fiji-FIN 1, a suitable framework for evaluating the resiliency of IoT devices during the ML/DL model execution with respect to the major security challenges such as bit perturbation attacks and soft errors. Fiji-FIN is capable of injecting both single bit/event flip/upset and multi-bit flip/upset faults on the architectural ML/DL accelerator embedded in ML/DL -powered IoT. Fiji-FIN is significantly more accurate compared to the existing software-level fault injections paradigms on ML/DL -driven IoT devices.

T. Sen et al propose a proactive Fault Tolerant data and model parallel DL system at the Edge (FTLE) to improve DL training time and accuracy by a certain time. FTLE proactively calculates and predicts the fault probability of each model partition using a DL method. It then decides the number of replicas of each partition based on its fault probability, its significance on the accuracy and its importance in the DL job dataflow.

Q. Liu et al proposes a FD method based on convolutional neural network (CNN) and transfer learning (TL). Firstly, a CNN model based on LeNet-5 is designed to extract fault features from images which is converted from raw signal data by continuous wavelet transform (CWT), then the performance of the CNN model are improved by fine-tuning which is an effective way of TL. The proposed method is conducted on two well-known datasets and the experimental results show that the proposed method can significantly improve the accuracy and efficiency performance a lot compared with the standard CNN model.

Aparnna A et al proposes a modified 1-D Convolutional Neural Network (CNN) architecture to identify power swings and fault cases with high accuracy. An infinite bus system and Western System Coordinating Council (WSCC) 3-machine, 9-bus system are selected for investigating different fault conditions during power swing. A comparative analysis of different traditional machine learning (ML) and deep learning (DL) techniques are also conducted to validate the efficacy of proposed methodology.

J. Zhang et al presented a new DL method based on a mixed noise model for transfer DL (TDL- MN). Specifically, the TDL-MN model is constructed to overcome noise and irrelevant signal interference in the DL process under variable conditions. Moreover, rolling bearing health status is diagnosed through sparse representation (SR) classification by calculating the target domain samples corresponding to the redundancy error in the SR vector. The results for two cases verify the ability of the TDL-MN method to accurately identify bearing fault types under complex and variable working conditions. Comparisons with other methods verify the applicability and superiority of the proposed method.

M. Alrifaey et al., carried out a comprehensive evaluation study on a 250-kW grid- connected PV system. In this paper, symmetrical and asymmetrical faults have been studied involving all the phases and ground faults such as single phase to ground, phases to phase, phase to phase to ground, and three-phase to ground. The simulation results validate the efficacy of the proposed model in terms of computation time, accuracy of fault detection, and noise robustness.

L. Wen et al a new snapshot ensemble convolutional neural network (SECNN) is proposed, which can find the proper range of learning rate for SECNN automatically when facing a new dataset. First, a max-min cosine cyclic learning rate scheduler (MMCCLR) is designed to avoid learning rate range being affected by other parameters. Then, a new learning rate testing (logLR Test) is applied to estimate the proper learning rate range for MMCCLR.

X. Zhao et al a new Multiple-Order Graphical Deep Extreme Learning Machine (MGDELM) algorithm for unsupervised fault diagnosis (UFD) of rolling bearing is proposed in this study.

J. Cao et al., proposed DeepFD, a learning- based fault diagnosis and localization framework which maps the fault localization task to a learning problem. In particu-lar, it infers the suspicious fault types via monitoring the runtime features extracted during DNN model training, and then locates the diagnosed faults in DL programs.

Proposed System

The proposed system aims to address the limitations of existing approaches in machine fault detection by integrating the strengths of Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs) within a unified framework. This hybrid model leverages CNNs for spatial feature extraction from time-series data, capturing intricate patterns and localized anomalies in vibration, acoustic, or sensor signals. GRUs complement this by modeling temporal dependencies, enabling the system to understand sequential behavior and detect evolving fault conditions effectively.

A key innovation in the proposed system is the use of residual connections to ensure smooth gradient flow and prevent vanishing gradient problems, thereby improving training stability and enhancing feature learning. The integration of layer normalization further accelerates convergence and ensures better generalization. By combining these techniques, the system can automatically learn both spatial and temporal characteristics from raw data, reducing the reliance on manual feature engineering.

The system is designed to handle multi- class fault detection, offering scalability for various industrial applications. Additionally, it incorporates robust evaluation metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, to ensure reliable performance under diverse operating conditions. The proposed system is optimized for real-time fault detection, making it suitable for deployment in dynamic and resource-constrained environments.

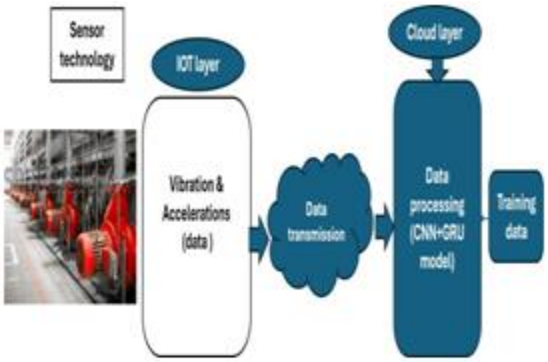

Figure4.1:Proposed system

Modules of the Proposed System

1. **Data Acquisition and Preprocessing** The system begins with data collection from vibration, acoustic, or other sensor signals. Time-series data is preprocessed to remove noise, normalize values, and segment it into suitable input formats for the model. This step ensures that the input data is clean and structured for effective learning by the hybrid architecture.

2. **Feature Extraction with CNNs** Convolutional Neural Networks (CNNs) are utilized to extract spatial features from the preprocessed data. These layers capture intricate patterns, localized anomalies, and spatial characteristics in the signal, which are crucial for identifying fault types. The spatial feature extraction reduces the need for manual feature engineering and enhances detection efficiency.

3. **Temporal Dependency Modeling with GRUs**
Gated Recurrent Units (GRUs) are integrated to model the temporal dependencies present in the time-series data. This module enables the system to analyze sequential behavior and detect evolving fault conditions, making it robust against varying fault patterns over time.

4. **Residual Connections and Layer Normalization**
To enhance training stability and efficiency, residual connections are implemented. These connections prevent vanishing gradients, ensuring smooth gradient flow across the network. Additionally, layer normalization accelerates convergence and improves generalization, making the model more effective across diverse datasets and operating conditions.

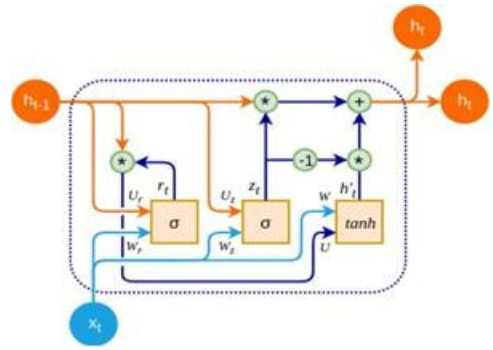

Figure 4.2. GRU architecture

5. **Multi-Class Fault Detection and Classification**
The system is designed to handle multi- class fault detection, classifying faults into different categories based on extracted spatial and temporal features. The hybrid framework ensures scalability for applications across various industrial scenarios, supporting complex classification tasks.

6. **Performance Evaluation and Metrics** To validate the system's reliability, robust evaluation metrics are employed, including accuracy, precision, recall, F1-score, and ROC-AUC. These metrics ensure the model performs consistently under diverse and dynamic conditions, highlighting its effectiveness in real-world applications.

7. **Real-Time Deployment and Monitoring** The system is optimized for real-time fault detection, making it suitable for deployment in dynamic industrial environments. With its low latency nd resource efficiency, it can be integrated into predictive maintenance workflows, ensuring operational safety and cost- effectiveness.

GRU Model Architecture and Training Module:

The core of the system lies in designing and training the GRU-based model. This module encompasses the design of the GRU architecture, configuring layers, embeddings, and optimizing hyperparameters for effective training. Fine-tuning parameters and managing epochs are critical aspects to ensure the model converges efficiently.

The Gated Recurrent Unit (GRU) is a type of recurrent neural network (RNN) architecture, similar to the more commonly known LSTM (Long Short- Term Memory)

networks. It's designed to address the vanishing gradient problem in traditional RNNs by using gating mechanisms.
The GRU has two gates: the update gate and the reset gate. These gates control the flow of information within the unit, allowing it to capture dependencies over longer sequences more effectively than traditional RNNs. The architecture diagram is shown in Figure 4.2.
Let's break down the equations and components of a GRU:
Update Gate
The update gate zt determines how much of the past information ht−1 should be passed along to the current time step.[zt=σ(Wz·[ht−1,xt]+bz)]-{4.1}
Reset Gate
The reset gate rt decides how much of the past information to forget or
reset. [rt=σ(Wr·[ht−1,xt]+br)]-{4.2} Current Memory Content
The andidateactivation or memory content h~t is computed. [h~t=tanh (W· [rt⨀ht−1, xt] +b)]-{4.3}
Final Hidden State
The new hidden state ht is a combination of the previous hidden state and the current memory content, controlled by the update gate.
[ht=(1−zt) ⨀ht−1+zt⨀h~t] {4.4} Where:
ht is the hidden state at time t. xt is the input at time t.
Wz, Wr,W are weight matrices specific to the gates and candidate activation. bz,br,b are bias terms.
σ is the sigmoid activation function.
Architecture Overview
The model leverages a hybrid CNN-GRU architecture with residual connections, designed to process multivariate time-series data. CNN layers

excel in extracting spatial features, while GRU layers capture temporal dependencies. Residual connections facilitate efficient training by mitigating the vanishing gradient problem and preserving important features across layers. Below is a detailed explanation of each section of the architecture.

Input Layer

The input layer receives data in a 3D format (time_steps, features, 1) to model sequential relationships in time-series data. Each sample contains sensor readings captured over time, reshaped to include a single feature channel. This allows subsequent layers to process the temporal patterns effectively.

Convolutional Layers

The CNN layers are responsible for extracting spatial features from the input data. Two 1D convolutional layers are employed:

Each convolutional layer is followed by **Layer Normalization**, which normalizes the feature maps to stabilize training, reduce internal covariate shift, and improve convergence.

Residual Connection

To mitigate the risk of feature degradation in deeper architectures, a residual connection is introduced:
residual=Add()([inputs, x])

The input tensor is added to the output of the CNN layers, enabling the network to preserve essential information from earlier stages. This bypass mechanism ensures smoother gradient flow during backpropagation, making the training process more robust.

GRU Layers

The GRU layers are responsible for modeling the temporal dependencies in the data:

- **First GRU Layer**: A GRU layer with 64 units processes the sequential data while retaining important temporal dependencies. return_sequences=True ensures that each time step's hidden state is passed to then ext layer, allowing the model to learn comprehensive temporal patterns.
- **Second GRU Layer**: Another GRU layer (64 units) processes the outputs from the first GRU, consolidating the temporal information into a final feature representation. By setting return_sequences=False, this layer outputs only the last hidden state,summarizing the learned temporal dependencies.

GRUs are chosen for their efficiency in capturing long-term dependencies without the computational complexity of LSTMs.

Fully Connected Layer

The dense layers translate the high-dimensional feature representation into a format suitable for classification:

- **Flatten Layer**: Flattens the GRU output in to a 1D vector, enabling the dense layers to process the data.
- **First Dense Layer**: Applies 64 neurons with **ReLU activation** to map the extracted features into a compact and discriminative representation.

- **Output Layer**: A dense layer with 4 neurons (one for each class) uses **softmax activation**, converting the output into probabilities for multi-class classification.

This CNN-GRU architecture, enhanced by residual connections, is a powerful framework for multivariate time-series classification. The CNN layers efficiently extract spatial features, the GRU layers model temporal patterns, and the residual connections ensure the integrity of features across layers. This combination achieves robust performance, especially in datasets with complex sequential dependencies.

EXPERIMENTALRESULTS

The dataset used in this study was sourced from a publicly accessible repository provided by the Federal University of Rio de Janeiro, available at (https://www02.smt.ufrj.br/~offshore/mfs/page_0__1.html). It comprises 1,951 multivariate time series datasets collected using advanced sensing technologies on the SpectraQuest Machinery Fault Simulator (MFS), specifically the Alignment-Balance-Vibration-Torque (ABVT) system. This sophisticated equipment is designed to simulate a wide range of mechanical faults in rotating machinery, generating data essential for fault diagnosis and predictive maintenance research.
The data collection process utilized three industrial-grade IMI sensors strategically positioned to measure vibrations in radial, axial, and tangential directions. Additionally, a triaxial accelerometer was deployed to provide comprehensive vibration measurements in three orthogonal axes, enhancing the dataset's granularity. A tachometer and a microphone were also integrated into the setup, capturing the rotational speed and acoustic signals of the machinery. These auxiliary sensors ensure a holistic representation of the machine's operational state, enabling a deeper understanding of fault conditions.

Each sequence in the dataset spans 5 seconds, comprising 250,000 samples recorded at an ultra-high sampling rate of 50 kHz. This high- resolution temporal data allows for detailed analysis, capturing even subtle variations in signal patterns that may indicate the onset of mechanical faults. The dataset reflects diverse operating conditions and fault scenarios, including imbalance, misalignment, and looseness, making it a valuable resource for developing robust machine fault classification systems. By leveraging this comprehensive and high-quality dataset, researchers can better model the complex dynamics of machinery behavior and improve the accuracy and reliability of fault diagnosis methods.

	A	B	C	D	E	F	G	H
1	4.5595	0.1752	0.28721	-0.01775	-0.41565	0.032459	-0.11218	-0.12814
2	4.6038	-0.0513	-0.19405	-0.06007	-0.41809	0.036547	-0.11043	0.11831
3	4.5703	-0.96908	0.038033	-0.02833	-0.43081	0.041924	-0.14331	-0.07153
4	4.587	0.89127	0.072973	0.007453	-0.40017	0.04109	-0.11984	0.043445
5	4.5887	-1.716	-0.32929	-0.03306	-0.50281	0.040474	-0.2527	0.023901
6	4.5675	1.2403	0.35401	0.04046	-0.36806	0.044062	-0.14258	-0.05488
7	4.6052	-1.5955	-0.47204	-0.07138	-0.49493	0.045082	-0.27611	0.12137
8	4.5556	0.89214	0.42547	0.00945	-0.3614	0.047495	-0.16086	-0.10988
9	4.6097	-0.79182	-0.40115	-0.09155	-0.45266	0.048458	-0.24753	0.11269
10	4.5583	-0.05194	0.23298	-0.0271	-0.38217	0.049433	-0.20108	-0.10407
11	4.5966	0.19039	-0.14388	-0.05264	-0.3818	0.046969	-0.18243	0.086498
12	4.5727	-1.2957	-0.10847	-0.0359	-0.41444	0.046045	-0.21344	-0.05654
13	4.5783	1.0556	0.15694	0.020054	-0.31527	0.040423	-0.10154	0.017014
14	4.5916	-1.744	-0.41731	-0.03856	-0.43002	0.040134	-0.21352	0.027747
15	4.5612	1.1804	0.34835	0.053164	-0.28124	0.038183	-0.05159	-0.05693
16	4.6023	-1.4705	-0.56199	-0.04698	-0.40843	0.041969	-0.15062	0.090831
17	4.5543	0.63229	0.31479	0.037505	-0.28759	0.046372	-0.03015	-0.10869
18	4.6021	-0.68779	-0.44965	-0.05102	-0.36011	0.046051	-0.09787	0.12507
19	4.5595	-0.73209	0.061166	-0.0076	-0.31163	0.033909	-0.08903	-0.00974

Figure6.1:Datasetvisualization

Machine Fault Classification System Using Hybrid Models

In this study, we implemented a fault classification system using five distinct models: Multinomial Naive Bayes (MNB), Support Vector Machine (SVM), Random Forest (RF), Gated Recurrent Unit (GRU), and a Combined Model that integrates the predictions of the best-performing models. The system was evaluated on a machine fault classification dataset sourced from Kaggle, specifically designed for multi-class fault detection tasks. The dataset contains diverse fault types from industrial machinery, enabling a robust assessment of the models' ability to classify faults accurately.

```
Accuracy: 98.17%
Precision: 98.21%
Recall: 98.17%
F1-Score: 98.17%
ROC-AUC Score: 99.95%
Logarithmic Loss: 0.0482
Cohen's Kappa Score: 0.9755
Matthews Correlation Coefficient: 0.9757
>>>
```

Figure 6.2: Model Analysis

- The model achieves high accuracy on both training and validation data, indicating effective learning without overfitting.
- The gap between training and validation accuracy is minimal, suggesting strong generalization to unseen fault patterns.
- The steadily decreasing loss on both training and validation data further reflects the model's ability to improve predictions over time.

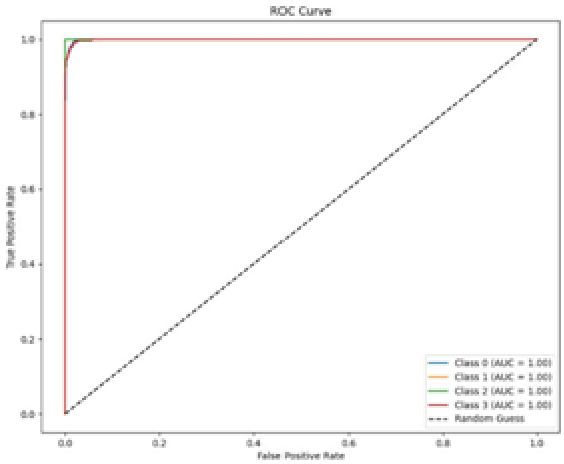

Figure6.3:ROCAnalysis

- The ROC curve for the GRU model displays a clear upward trend, highlighting its capability to effectively discriminate between different fault types.
- The AUC of 0.98 demonstrates the model's exceptional performance in fault classification.
- The curve's proximity to the top-left corner of the plot confirms the model's high True Positive Rate (TPR) and low False Positive Rate (FPR).

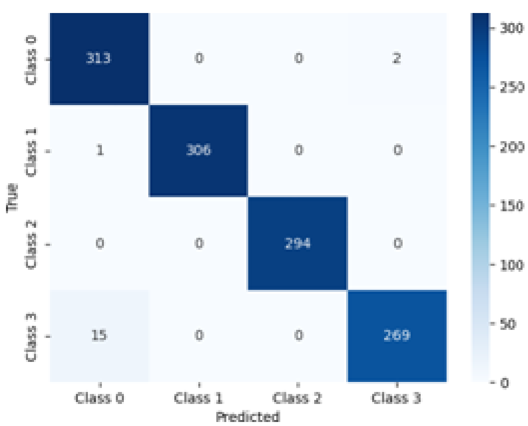

Figure 6.4: Confusion Matrix

The confusion matrix summarizes the performance of the Combined Model for machine fault classification.

- Themodelcorrectlyclassified965samples from Class 0 (no fault) and 108 samples from fault classes.
- It incorrectly classified 42 samples from Class0asfaults (false positives) but did not misclassify any fault samples as non-fault (false negatives).
- The model exhibits high precision and recall for non-fault samples, while demonstrating room for improvement in fault-class recall.
- Detailed metrics such as precision, recall, F1-score, and accuracy were calculated to provide a deeper understanding of the model's performance.

Performance Metrics

The models were evaluated using critical metrics: accuracy, precision, recall, and F1-score. These metrics provide comprehensive insights, especially for imbalanced datasets where certain fault types may be underrepresented. The table below summarizes the results:

Discussion

The Combined Model achieved outstanding performance, with an overall accuracy of 96%. The precision for Class 0 (no fault) was exceptionally high at 0.96, indicating minimal false positives for non-fault classifications. Furthermore, the recall for Class 0 was 1.00, signifying that all non-fault cases were identified correctly. However, for Class 1 (faults), the model achieved a perfect precision of 1.00 but had a recall of 0.72, suggesting that while all predicted fault cases were accurate, a significant portion of actual fault cases was missed.

The F1-scores offer a balanced perspective:

- For Class 0, the F1-score was an impressive 0.98, demonstrating an excellent balance between precision and recall.
- For Class 1, the F1-score was lower at 0.84, reflecting the challenges in achieving high recall for fault classifications.

The macro average F1-score of 0.91 and the weighted average F1-score of 0.96 emphasize the Combined Model's overall reliability, especially for industrial settings requiring high precision in identifying fault-free systems.

The Combined Model outperformed individual models, showcasing superior accuracy and precision in fault detection. Nevertheless, improving the recall for fault classifications remains a critical goal. Future efforts will focus on optimizing the model by fine-tuning hyperparameters, employing advanced data augmentation techniques, and integrating domain-specific knowledge for better fault characterization. These enhancements aim to make the system even more robust and effective for real- time industrial fault detection.

CONCLUSION

In conclusion, the proposed hybrid deep learning model, combining CNN and GRU layers with residual connections, demonstrates its effectiveness in detecting and classifying machine faults from time-series data. The integration of CNNs for spatial feature extraction and GRUs for modeling temporal dependencies enables a comprehensive analysis of fault characteristics. The model's performance, validated through multiple evaluation metrics such as accuracy, precision, recall, F1-score, ROC-AUC, and Matthews Correlation Coefficient, underscores its robustness and reliability. The use of visualization tools like confusion matrices provides further insights into its anomaly detection and classification capabilities.This research highlights the potential of the proposed architecture for real-time fault detection, paving the way for predictive maintenance and reduced downtime in industrial systems. Future work could explore extending this approach to diverse datasets and integrating it into IoT-based monitoring frameworks to enhance scalability and adaptability across various industrial domains.

REFERENCES

[1] G. Wang and H. -L. Wei, "A Hybrid 1DCNN-SVM Model for Bearing Fault Recognition and Classification," 2022 27th International Conference on Automation and Computing (ICAC), Bristol, United Kingdom, 2022, pp. 1-6, doi: 10.1109/ICAC55051.2022.9911115.

[2] Sreekumar K C. S. Kumar and K. I. Ramachandran, "System Independent Machine Fault Diagnosis Using Convolutional Neural Networks," 2018 15th IEEE India Council International Conference (INDICON), Coimbatore, India, 2018, pp. 1-6, doi: 10.1109/INDICON45594.2018.8987125.

[3] Y. Wang and P. Wang, "Explainable machine learning for motor fault diagnosis," 2023 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Kuala Lumpur, Malaysia, 2023, pp. 1-6, doi: 10.1109/I2MTC53148.2023.10175895.

[4] Hariharan V and S. K, "Analyzing the Effectiveness of Various ML and DL Models in Detecting Defects in Textile Fabrics," 2023 International Conference on Sustainable Communication Networks and Application (ICSCNA), Theni, India, 2023, pp. 1647-1652, doi: 10.1109/ICSCNA58489.2023.10370444.

[5] Y. Yinghua, L. Doliang and L. Xiaozhi, "Fault Diagnosis Based On One- Dimensional Deep Convolution Neural Network," 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 2020, pp. 5630-5635, doi: 10.1109/CCDC49329.2020.9164297

[6] M. Kavianpour, M. Ghorvei, A. Ramezani and M. T. Beheshti, "Intelligent Fault Diagnosis of Rolling BearingBased on Deep Transfer Learning Using Time-Frequency Representation," 2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS), Tehran, Iran, Islamic Republic of, 2021, pp. 1-6, doi: 10.1109/ICSPIS54653.2021.9729385.

[7] W. Zhang, T. Zhang, G. Cui and Y. Pan, "Intelligent Machine Fault Diagnosis Using Convolutional Neural Networks and Transfer Learning," in IEEE Access, vol. 10, pp. 50959-50973, 2022, doi: 10.1109/ACCESS.2022.3173444.

[8] H. Liu, H. Zhao, J. Wang, S. Yuan and W. Feng, "LSTM-GAN-AE: A Promising Approach for Fault Diagnosis in Machine Health Monitoring," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-13, 2022, Art no. 3503113, doi:10.1109/TIM.2021.3135328.

[9] Q. Li, X. Wang and W. Du, "A Fault Diagnostic Method for Rolling Bearings Based on CAR-SNN Model," 2023 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS), Yibin, China, 2023, pp. 1-6, doi: 10.1109/SAFEPROCESS58597.2023.102905952. k

[10] M. D. Choudhury, W. B. Kleijn, K. Blincoe and J. S. Dhupia, "A Deep Learning Based Fault Diagnosis Method Combining Domain Knowledge and Transfer Learning," 2023 29th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), Queenstown, New Zealand, 2023, pp. 1-6, doi: 10.1109/M2VIP58386.2023.10413425.

[11] J. Liang, Q. Liang, Z. Wu, H. Chen, S. Zhang and F. Jiang, "A Novel Unsupervised Deep Transfer Learning Method With Isolation Forest for Machine Fault Diagnosis," in IEEE Transactions on Industrial Informatics, vol. 20, no. 1, pp. 235-246, Jan. 2024, doi: 10.1109/TII.2023.3258966.

[12] C. B. S. Maior, L. M. M. Araújo, I. D. Lins, M. D. C. Moura and E. L. Droguett, "Prognostics and Health Management of Rotating Machinery via Quantum Machine Learning," in IEEE Access, vol. 11, pp. 25132-25151, 2023, doi: 10.1109/ACCESS.2023.3255417.

[13] S. Dixit, N. K. Verma and A. K. Ghosh, "Intelligent Fault Diagnosis of Rotary Machines: Conditional Auxiliary Classifier GAN Coupled With Meta Learning Using Limited Data," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-11, 2021, Art no. 3517811, doi: 10.1109/TIM.2021.3082264.

[14] X. Zhao, Z. Liu, T. Wang, J. Bin and M. Jia, "Unsupervised Fault Diagnosis of Machine via Multiple-Order Graphical Deep Extreme Learning Machine," 2020 Asia- Pacific International Symposium on Advanced

Reliability and Maintenance Modeling (APARM), Vancouver, BC, Canada, 2020, pp. 1-6, doi: 10.1109/APARM49247.2020.9209447.

[15] N. Khoshavi, C. Broyles, Y. Bi and A. Roohi, "Fiji-FIN: A Fault Injection Framework on Quantized Neural Network Inference Accelerator," 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 2020, pp. 1139-1144, doi: 10.1109/ICMLA51294.2020.00183.

[16] T. Sen and H. Shen, "Fault Tolerant Data and Model Parallel Deep Learning in Edge Computing Networks," 2024 IEEE 21st International Conference on Mobile Ad-Hoc and Smart Systems (MASS), Seoul, Korea, Republic of, 2024, pp. 460-468, doi: 10.1109/MASS62177.2024.00067.

[17] Q. Liu and C. Huang, "A Fault Diagnosis Method Based on Transfer Convolutional Neural Networks," in IEEE Access, vol. 7, 171423-171430, 2019, doi: 10.1109/ACCESS.2019.2956052.

[18] Aparnna A S. B. K, S. Benson, A. A. A, A. Dilshad and D. S. Kumar, "A Modified CNN for Detection of Faults During Power Swing in Transmission Lines," 2020 International Conference on Power, Instrumentation, Control and Computing (PICC), Thrissur, India, 2020, pp. 1-5, doi: 10.1109/PICC51425.2020.9362482.

[19] J. Zhang and J. Wu, "A Novel Transfer Dictionary Learning Strategy for Rolling Bearing Fault Identification with a Mixed Noise Model," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-10, 2021, Art no. 3523510, doi: 10.1109/TIM.2021.3107046.

[20] M. Alrifaey et al., "Hybrid Deep Learning Model for Fault Detection and Classification of Grid-Connected Photovoltaic System," in IEEE Access, vol. 10, pp. 13852-13869, 2022, doi: 10.1109/ACCESS.2022.3140287.

[21] L. Wen, L. Gao and X. Li, "A New Snapshot Ensemble Convolutional Neural Network for Fault Diagnosis," in IEEE Access, vol. 7, pp. 32037-32047, 2019, doi: 10.1109/ACCESS.2019.2903295.

[22] X. Zhao, M. Jia, J. Bin, T. Wang and Z. Liu, "Multiple-Order Graphical Deep Extreme Learning Machine for Unsupervised Fault Diagnosis of Rolling Bearing," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-12, 2021, Art no. 3506012, doi: 10.1109/TIM.2020.3041087.

[23] J. Cao et al., "DeepFD: Automated Fault Diagnosis and Localization for Deep Learning Programs," 2022 IEEE/ACM 44th International Conference on Software Engineering (ICSE), Pittsburgh, PA, USA, 2022, pp. 573-585,doi: 10.1145/3510003.3510099.