International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 11 | Nov 2026 www.irjet.net p-ISSN: 2395-0072

AI -BASED INDUSTRIAL MACHINE FAULT PREDICTION

Dr. P. Loganthurai[1], M. Haritha[2], S. Kayalvizhi[3]

[1]*Associate professor, Department of EEE, K.L.N College Of Engineering(An Autonomous Institution,Affiliated Anna university)Pottapalayam,Sivagangai, Tamilnadu, India.*
[2,3]*Second UG scholars, Department Of EEE,K.L.N College Of Engineering(An Autonomous Institution, Affiliated Anna university)Pottapalayam , Sivagangai, Tamilnadu, India.*

Abstract - *Fault detection in industrial machines is important for improving safety, reducing machine downtime, and lowering maintenance costs. This study proposes a hybrid deep learning model that combines Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs) to detect and classify machine faults using time-series data. CNN layers extract important features from the data, while GRU layers learn patterns over time. Residual connections are used to improve the learning process. The model is evaluated using different performance measures, such as accuracy, precision, recall, and F1-score. Additional measures, including ROC-AUC, logarithmic loss, Cohen's Kappa, and Matthews Correlation Coefficient, are used to evaluate its reliability. Confusion matrices and performance metrics help assess the model's ability to identify and classify machine faults. This approach can support real-time machine monitoring and predictive maintenance in various industries.*

INTRODUCTION

Industrial machines play an important role in manufacturing industries and industrial operations. They are used to perform various tasks and maintain continuous production. The proper operation of these machines is important for improving productivity, efficiency, and safety. However, machines may develop faults due to continuous operation, wear and tear, excessive vibration, imbalance, and misalignment. These faults can cause unexpected breakdowns, production delays, increased maintenance costs, and safety problems. Therefore, early detection of machine faults is important to prevent major failures and maintain reliable operation.
Fault detection is the process of identifying abnormal conditions in a machine, while fault diagnosis helps determine the type of fault. Traditional methods often depend on manual inspection and signal analysis, which may require more time to identify complex faults. Sensor technology helps monitor machine conditions by collecting data such as vibration signals and temperature. The collected data can be analysed to identify changes in machine behaviour and detect abnormal operating conditions.

Artificial Intelligence (AI) provides an effective method for analysing machine data and identifying patterns related to faults. Deep learning models can learn important features from sensor data and help distinguish between normal and faulty machine conditions. In this project, a CNN-GRU-based model is used to analyse time-series data for industrial machine fault prediction and diagnosis. The Convolutional Neural Network (CNN) extracts important features from the input data, while the Gated Recurrent Unit (GRU) analyses changes in the data over time. The model helps identify and classify different types of machine faults.

The performance of the proposed model is evaluated using measures such as accuracy, precision, recall, F1-score, and ROC-AUC. This approach supports predictive maintenance by helping identify potential machine problems before they cause unexpected breakdowns. Early fault identification can reduce machine downtime, improve maintenance planning, and increase equipment reliability and safety. Thus, the project focuses on using AI-based techniques to improve industrial machine fault prediction and diagnosis.

LITERATURE 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.

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 environment

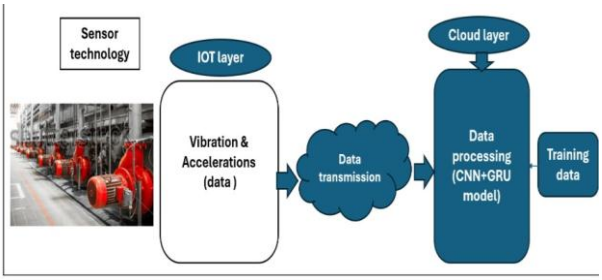


Figure 1.1:proposed system

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.

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

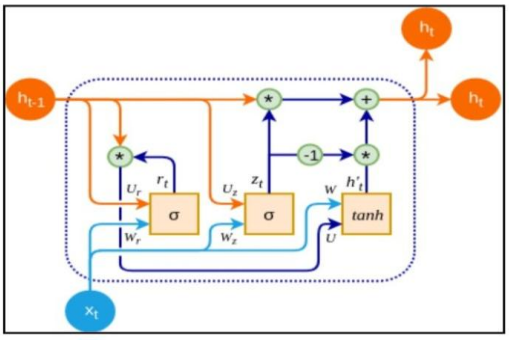
The system is optimized for real-time fault detection, making it suitable for deployment in dynamic industrial environments. With its low latency nd resource efficiency, it can be integrated into predictive maintenance workflows, ensuring operational safety and cost-effectiveness.

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 1.2



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.

$$z_t=\sigma(W_z\cdot[h_{t-1},x_t]+b_z) \quad (1)$$

Reset Gate

The reset gate rt decides how much of the past information to forget or reset.

$$r_t=\sigma(W_r\cdot[h_{t-1},x_t]+b_r) \quad (2)$$

Current Memory Content

The andidateactivation or memory content h~t is computed.

$$\tilde{h}_t=\tanh(W\cdot[r_t\odot h_{t-1},x_t]+b) \quad (3)$$

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.

$$h_t=(1-z_t)\odot h_{t-1}+z_t\odot\tilde{h}_t \quad (4)$$

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:

•**First Conv1D Layer:** Applies 64 filters with kernel size of 3, sliding over the input data to detect local patterns. The ReLU activation function introduces non-linearity, enhancing the model's ability to learn complex relationships. Padding='same' ensures the output size matches the input size by appending zeros at the edges.

•**Second Conv1D Layer:** Another layer with identical configurations (64 filters, kernel size of 3, ReLU activation). This builds upon the feature map created by the first layer, allowing deeper representations of the spatial patterns.

Each convolutional layer is followed by LayerNormalization, 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 back propagation, 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 the next 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 into 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.

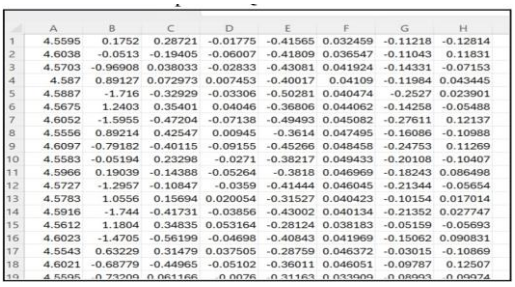
EXPERIMENTAL RESULTS

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_01.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 data set'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.

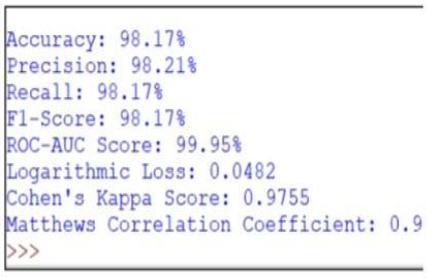


	A	B	C	D	E	F	G	H
1	4.5595	0.1752	0.28721	-0.01775	-0.41565	0.032459	-0.11218	-0.12814
2	4.6038	-0.0513	-0.19405	-0.06007	-0.41809	0.036547	-0.11043	0.11831
3	4.5703	-0.96908	0.038033	-0.02833	-0.43081	0.041924	-0.14331	-0.07153
4	4.587	0.89127	0.072973	0.007453	-0.40017	0.04109	-0.11984	0.043445
5	4.5887	-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	0.033999	-0.08993	-0.09974

Figure1.3: Dataset visualization

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: 0.9
>>>
```

Figure 1.4: 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.

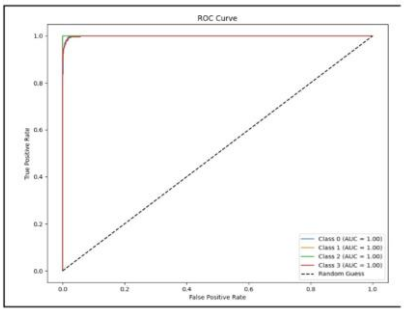


Figure 1.5: ROC Analysis

•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).

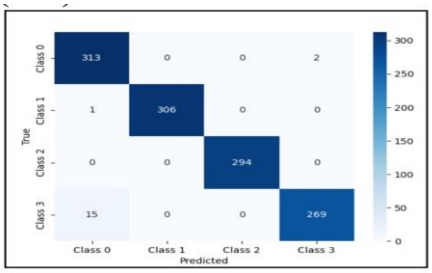


Figure 1.6: Confusion Matrix

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

•The model correctly classified 965 samples from Class 0 (no fault) and 108 samples from fault classes.

•It incorrectly classified 42 samples from Class 0 as faults (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:

Model	Accuracy (%)	Precision (0)	Recall (0)	F1-Score (0)	Precision (1)	Recall (1)	F1-Score (1)	Macro Avg F1-Score	Weighted Avg F1-Score
Multinomial Naive Bayes (MNB)	92	0.91	0.93	0.92	0.83	0.79	0.81	0.87	0.91
Support Vector Machine (SVM)	94	0.92	0.96	0.94	0.86	0.75	0.80	0.87	0.93
Random Forest (RF)	93	0.89	0.95	0.92	0.85	0.76	0.80	0.86	0.92
Gated Recurrent Unit (GRU)	95	0.94	0.98	0.96	0.90	0.70	0.78	0.87	0.95
Combined Model	96	0.96	1.00	0.98	1.00	0.72	0.84	0.91	0.96

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,UnitedKingdom,2022,pp.16,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.16,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.] 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/SAFEPROCESS58597.2023.10295952. 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.