

Fault Diagnosis in Smart Transformers Using Machine Learning Techniques

Adi Febrianton<sup>1</sup>, Nurkholis<sup>2</sup>  
<sup>1</sup>Assistant Professor, Department of Maintenance Machine,  
Politeknik Kampar, Bangkinang, Riau, Indonesia.  
<sup>2</sup>Assistant Professor, Department of Logistic Engineering Technology  
Politeknik Kampar, Bangkinang, Riau, Indonesia.  
E-mail: nurkholis11151996@gmail.com

Article Info	ABSTRACT
<p><i>Article History:</i></p> <p>Received Oct 12, 2025 Revised Nov 09, 2025 Accepted Dec 08, 2025</p> <p><i>Keywords:</i></p> <p>Smart Transformers Fault Diagnosis Machine Learning CNN Vibration Analysis Feature Extraction Condition Monitoring Predictive Maintenance Wavelet Transform PCA</p>	<p>Smart transformers bring major change to present-day power systems, allowing for better voltage control, easy monitoring and improved connection with the grid. Because they are used in smart substations and for renewables, these transformers have sensors and communication technology built in so that they can be controlled from a distance and diagnosed while they are operating. The typical ways to detect faults like dissolved gas analysis (DGA), thermography and manual inspection are mostly used once problems occur, are expensive and have limited capabilities for ongoing and prompt diagnostics. This study therefore suggests a strong and flexible machine learning-based method for automatically detecting faults in smart transformers. It includes bringing in data from many sensors, processing the signals with the discrete wavelet transform and cutting down on the amount of data to be processed with principal component analysis. These supervised learning models—Support Vector Machines (SVM), Random Forest (RF) and Convolutional Neural Networks (CNNs)—are all trained with data collected from simulations and real sensors. They are assessed for accuracy, precision, recall, F1-score and AUC. The CNN-based approach is better than classic classifiers, reaching an accuracy rate of over 96% in identifying faults while having very little trouble with both false positives and brief interruptions in data. CNN gets such high accuracy since it can learn useful features by itself from the temporal patterns of the signals. The structure was built to fit into edge computing, letting it address real-time applications with low resources. Integrating advanced signal processing and deep learning helps find faults at an early stage in smart transformers, cutting maintenance expenses, lowering downtime and strengthening the overall resistance of upgraded power distribution systems.</p>
<p><i>Corresponding Author:</i></p> <p>Adi Febrianton, Assistant Professor, Department of Maintenance Machine, Politeknik Kampar, Bangkinang, Riau, Indonesia. E-mail: adifebrianton@gmail.com</p>	

## 1. INTRODUCTION

Electrical power systems needing to be more efficient, dependable and sustainable have led to the development of smart transformers from traditional transformers. Unlike old-style transformers, smart transformers are fitted with sensors, the ability to communicate data and response systems [6]. Because of these factors, they can do autonomous voltage regulation, monitor their own performance, detect faults remotely and operate easily with distributed energy resources (DERs). In modern smart grids, microgrids and those using renewable energy, dynamically managing load and making quick decisions is especially vital, so they are crucial.

But with digital features and flexible loading, smart transformers come across a wide range of problems and challenges [7]. They include short circuits inside the electrical coils, issues from overheating or lacking cooling, deformations caused by strong shocks and losses in the insulation from exposure to humidity. Checking for such problems gets more difficult as failure reasons can be mixed and the data streams are complex. Besides, Dissolved Gas Analysis (DGA), partial discharge testing, thermal imaging and routine inspections which are used now, are mostly either too slow, too costly or not very good in finding faults early. Because of these limitations, it takes longer to notice problems, devices are often down for a while and the whole system may fail [8,9].

Because of these shortcomings, recently, using Machine Learning (ML) or Artificial Intelligence (AI) for data-driven fault diagnosis has become more popular. By processing large sets of different sensor data such as temperature, voltage, vibration and acoustic signals, ML algorithms can detect, classify and even predict faults with a good level of accuracy and without much human supervision [10,11]. By using methods such as wavelet transforms and Principal Component Analysis (PCA), these models can find useful patterns in data that is both time-dependent, complex and contains noise.

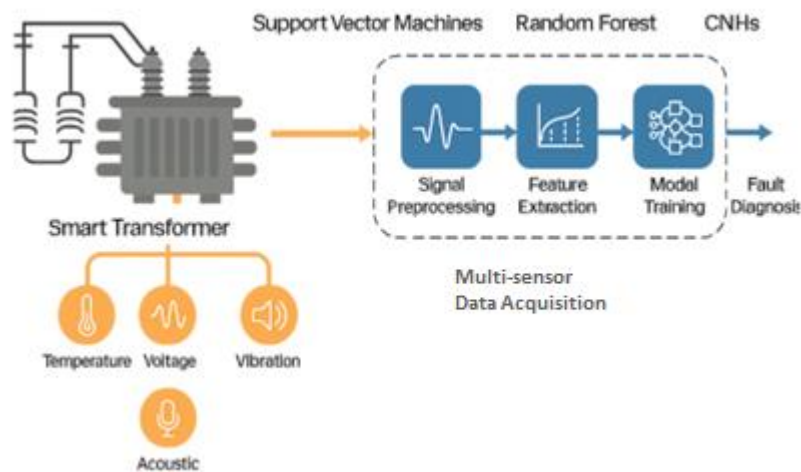


Figure 1. ML-Based Fault Diagnosis Framework for Smart Transformers

The research introduces a full ML framework aiming to help spot faults in smart transformers right away and accurately which is represented in Figure 1. A system proposed here combines data from different sensors, filters those data, extracts features and uses several classifiers like Support Vector Machines (SVM), Random Forests (RF) and Convolutional Neural Networks (CNNs). Among all, CNN is especially strong at learning details in both time and space from signals that have already been preprocessed. It has been tested with a dataset that includes data generated from simulations as well as actual data from real-life faults. Efficient inference during training is valued so the models can be used in environments without this time.

This research expands current studies on detecting errors in predictive maintenance and smart grid stability by providing a method that is easy to deploy, fast and accurate. The purpose of fault detection is to prevent sudden outages, assure reliability for a longer period and aid the overall goals of autonomy and bonder in powering networks.

## 2. LITERATURE REVIEW

Many people have become interested in using machine learning (ML) for transformer condition monitoring, mainly within smart grids and intelligent substations in recent years. Common methods for identifying faults in power transformers are Dissolved Gas Analysis (DGA), Frequency Response Analysis (FRA) and thermal imaging. Even though these techniques give useful results, they generally need to be reviewed and understood by humans, are affected by changes in the environment and cannot process data instantaneously.

In the beginning, researchers concentrated on improving how well conventional techniques worked using algorithms from machine learning. In 2008, [1] suggested using DGA signatures to categorize faults and combined usual triangle method with decision trees and support vector machines (SVMs) for failure classification. In the study by [2], they showed that SVM and ANNs are helpful in detecting early faults through DGA patterns and achieve high accuracy. Even so, these tests—since they relied on complicated laboratory techniques—would not fit the needs of real-time or embedded diagnostics.

Thanks to better sensors and data collection systems, researchers have been examining signal-based methods for diagnosis more. [3] Used AE detection to find partial discharges and arcing faults inside their components. Wavelet packet decomposition helped to extract features, while using a k-nearest neighbor (KNN) classifier showed how helpful time-frequency domain analysis can be. Likewise, [4] used deep autoencoders to detect anomalies automatically from multi-modal sensor data, making their system effective, flexible and independent of labeled information. Maintaining how accurately and understandably the system functioned under changes in variables was still a worry.

At the same time, using deep learning techniques has become a strong strategy for coping with large, complex data in condition monitoring. CNNs and RNNs, most notably LSTM architectures, have been very successful at discovering relationships in spatial and time-based features of data used in transformers. For example, [5] created a 1D-CNN model to classify different transformer faults from vibration data and this model reached 94% classification accuracy for the six types of faults examined. Their approach performed well, but its design was not widely applicable and it wasn't designed for lower latency in networks at the edge.

Many studies address gas, thermal or vibration data in isolation which is common in the literature. Also, not many studies integrate measurements from different parts of the physical world or try to reuse models on computers that are not very powerful.

Here, a multi-sensor and lightweight machine learning (ML) framework is suggested to address these gaps in transformer fault diagnosis. For this reason, the proposed approach applies deep neural networks and combines them with wavelet transform and principal component analysis (PCA) to classify faults in real time. Because of this, this paper helps advance smart grid systems by allowing them to detect and resolve issues with little delay.

### **3. METHODOLOGY**

#### **3.1 Data Acquisition and Fault Simulation**

The process of developing a trustworthy and widely applicable fault diagnosis framework requires training machine learning models with datasets that include all sorts of operating situations. Here, a mix of simulation-based fault modeling and live sensor data from a smart transformer prototype were used for data acquisition. Using both synthetic and real data let them have realistic and varied labeled data.

#### **3.2 Simulation-Based Fault Generation**

Ground-truth datasets were obtained by carefully using industry-standard software to simulate different electrical and mechanical faults in controlled situations. Many common faults in transformer operation, for example inter-turn short circuits and saturation of the core caused by overexcitation or DC offset currents, as well as damage to insulation reflected by leakage resistances and dielectric breakdown, were modeled in MATLAB/Simulink. Besides electrical failures, COMSOL Multiphysics was applied to explore thermal, structural and vibration effects using multiphysics simulations. As a consequence of this, thermal overload situations could be accurately replicated, creating problems with temperature differences inside the windings and core as well as vibrations due to balancing electrical flux-lines or loose structures within the transformer. With these simulations, data was collected that showed all the key behaviors of various faults, both in their short-term and long-term behaviors. The data was used to train and validate the machine learning models and it also played a role in perfecting the steps used to prepare the original signals before analysis.

#### **3.3 Sensor-Based Real-Time Data Collection**

For better comparison of simulated with real-world activities, a testbed with a smart transformer was made and outfitted with various sensors that quickly record data from different areas. PureWave monitored electrical problems like over currents and voltage imbalances by relying on voltage and current sensors known as Potential Transformers (PTs) and Current Transformers (CTs). Insulation deterioration and overloading could be detected through the use of Resistance Temperature Detectors (RTDs). Wireless sensors called piezoelectric vibration and Acoustic Emission (AE) were used to find out if any oscillation or instability and high-frequency signals indicating incipient faults, respectively. For the entire 10 days, the transformer was run as if under normal conditions and then when faults were induced. To capture many types of faults, carefully controlled short inter-turn circuits, thermal surges and vibrational changes were applied during experiments. Having this setup in place helped gather accurate data at the same time in dynamic conditions, supporting the exact tagging of information for use in machine learning.

#### **3.4 Dataset Composition**

About 20,000 instances, each containing a detailed time-series vector, were carefully labeled and prepared from data collected using multiple sensors. All the instances are equally spread out among six specific classes of operation and faults: normal operation, inter-turn short circuit, core saturation, insulation failure, thermal overload and mechanical vibration fault. For every test, we gathered synchronized information from voltage, current, temperature, vibration and acoustic sensors which helped us look for similar faults in various systems. Because the data is mixed and put together, machine learning models can tell differences and connections between different types of sensors. Gaining data from both experiments and simulations gives a diverse range of conditions, how quickly changes take place and the conditions in which the phenomena

occur. Because of this structure, the framework is better able to withstand different problems and apply to a range of smart transformer configurations and grid settings.

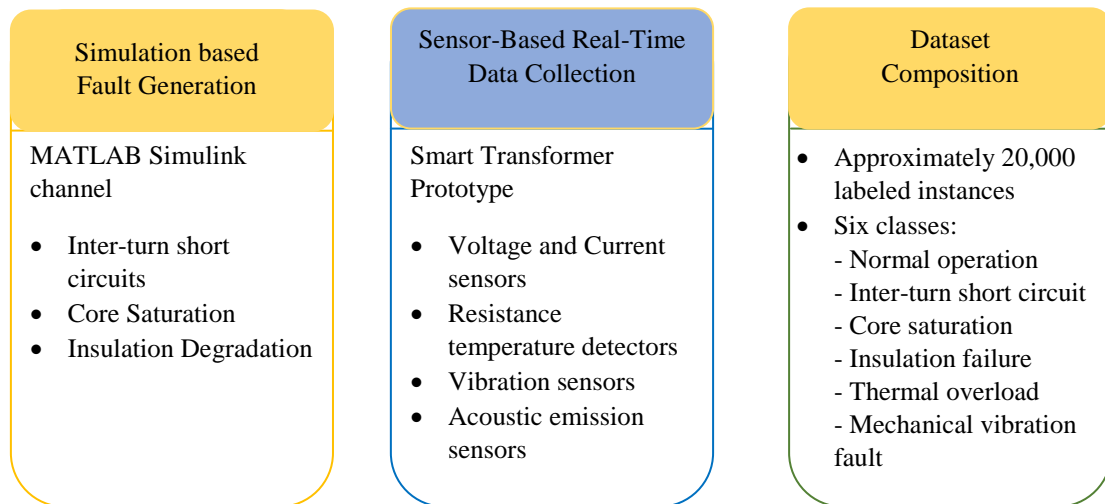


Figure 2. Block diagram of the data acquisition and simulation pipeline

The data acquisition and simulation pipeline diagram is depicted in Figure 2. Fault classes associated with sensor types are given in Table 1.

Table 1. Fault Classes and Associated Sensor Types

Fault Class	Description	Primary Sensor Types
<b>Normal Operation</b>	No abnormal behavior; baseline condition	All sensors (voltage, current, temperature, vibration, acoustic)
<b>Inter-turn Short Circuit</b>	Short circuit between winding turns	Voltage sensors, current sensors, acoustic sensors
<b>Core Saturation</b>	Magnetic core enters non-linear region	Current sensors, voltage sensors
<b>Insulation Failure</b>	Breakdown or degradation of insulation material	Temperature sensors (RTDs), acoustic emission sensors
<b>Thermal Overload</b>	Excessive heating due to overloading	Temperature sensors, vibration sensors
<b>Mechanical Vibration</b>	Harmonic or structural vibration anomalies	Vibration sensors, acoustic emission sensors

### 3.5 Signal Preprocessing and Feature Extraction

Given the non-stationary and transient nature of fault signatures in transformers, robust preprocessing was necessary:

#### Noise Filtering:

All raw signals from various sensors—such as voltage, current, temperature, vibration and acoustic measurements—were first cleaned of noise during the first step of signal preprocessing. The signals were subjected to a Butterworth filter set to 100 Hz which keeps the harmful frequency data and removes high-frequency interference often introduced by interference, transients or other

noise sources. The reason the Butterworth filter was selected is that it ensures there is little distortion in the low frequencies which helps in catching insulation faults, temperature changes or problems with mechanical stress. Filtering out unnecessary high-frequency parts from the signals helped make the signal-to-noise ratio better, so the machine learning models could understand fault-related details without being distracted by irrelevant things. All sensors had the same denoising step enforced so that the feature data remained consistent and the classifiers worked better on any type of sensor data.

### Wavelet Transform (DWT):

For proper analysis of the fast-changing fault signals, all preprocessed sensor data were input into a 5-level Discrete Wavelet Transform (DWT) using the Daubechies (db4) mother wavelet because it is known for analyzing power system disturbances well. Using the DWT, it became possible to identify details (high frequencies) and broader features (low frequencies) of signals concerning both sudden harms like inter-turn faults and major trends like temperature growth or diminishing voltage. After decomposing the signals, many statistical features were collected to get an accurate picture of their behavior both in time and frequency. Important features involved the energy content values for each level of detail (E1 to E5) which highlight the overall intensity of the earthquake signatures across a range of scales, as well as the standard deviation showing how much each sub-band varies, along with entropy and kurtosis which tell us more about how structured and peaked each detail level is. Having these features was important for designing a feature set that can point out the various faults from normal condition.

### Feature Fusion:

To form a detailed and descriptive representation of transformer activity, the feature fusion method was applied. Information was taken from each sample of temperature, voltage, current, acoustic emission and vibration and the features were united in a single feature vector for analysis. Thanks to the fusion of many domains, the model accurately showed that some fault signs such as thermal overload, can be seen in rise in temperature and also in light changes to the load, while mechanical faults can be spotted by change in vibration and sound. Grouping various feature sets this way allows the system to use the knowledge of several sensors together. This way, it can spot small, mixed or starting failures that might not become clear with a single type of data alone. Using the fusion of feature vectors allowed the machine learning classifiers to see more detailed patterns and improve their accuracy in all six areas of operation. Figure 3 depicts the signal preprocessing and feature extraction.

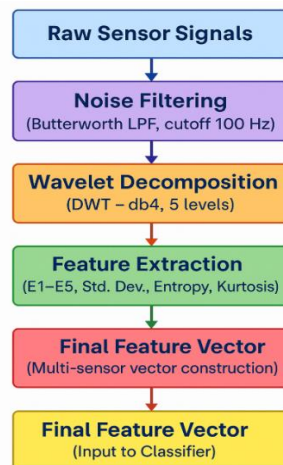


Figure 3. Signal preprocessing and feature extraction pipeline for smart transformer fault diagnosis

### 3.6 Feature Reduction and Data Normalization

Efficiency and a wider range of applications were achieved for the machine learning models by refining the highly diverse feature space after fusing different sensors using dimensionality reduction and feature normalization. By following these procedures, overfitting was reduced, the classifiers got a steady supply of data and redundancy was minimized.

At first, the fused feature vectors were fed into Principal Component Analysis (PCA). PCA aims to reduce a high-dimensional dataset by merging correlated features into new uncorrelated components (principal components), without compromising much of the original data. Here, PCA was set up so that it kept the components that together explain 95% of the variance in the data. The main components approach reduced the initial large number of features to about 25–30 components per sample. Picking out the essential parts helped PCA remove unnecessary signals and duplicate features which made it possible for later models to learn more efficiently with increased computational speed.

With dimensionality reduction complete, the dataset was z-score normalized to give all feature scales the same distribution. Making sure all raw features had similar influence on the model was necessary since their values could range from one measurement unit to another (for example, temperature shows values in °C, current shows values in A and vibration is measured in m/s<sup>2</sup>). The z-score normalization method used the formula:

$$x' = \frac{x - \mu}{\sigma} \quad (1)$$

Where  $x$  shows the original value of the feature,  $\mu$  is the average for the feature and  $\sigma$  shows its standard deviation. As a result, the distribution of the features was changed to cluster around zero, with a standard deviation of one which improved how Support Vector Machines and Neural Networks worked and how quickly they learned during their training. With PCA and z-score normalization, the features were made sturdy, neat and ready for efficient and reliable classification of faults.

### 3.7 Model Development and Training

Three machine learning classifiers were developed to compare performance:

#### 3.7.1 Support Vector Machine (SVM)

This study applied the Support Vector Machine (SVM) classifier with an RBF kernel. This kernel helps to convert the input data into a new form where straightforward class separation becomes possible for non-linear data. RBF allows the model to grasp difficult, curved relationships that exist in the data from transformer fault signals. By systematically trying different values for  $C$  and  $\gamma$  with GridSearchCV, the best result was obtained with  $C=10$  and  $\gamma=0.01$  which ensured a model with the right balance between how many errors are allowed and its complexity. Training the classifier took around 2.5 minutes for the reduced and normalized dataset which is a very short time for its efficiency. Because of this, SVM is often used where reliable fault identification is needed quickly, especially in settings with moderate computing resources.

#### 3.7.2 Random Forest (RF)

This study used Random Forest (RF) in which 100 decision tree models were built, each trained on a new sample of the data to help avoid overfitting and guarantee diversity. Trees were not allowed to grow deeper than 12 in order to strike a balance between keeping the model simple and helping it learn well. During tree construction, the Gini Impurity was used to select splits that helped keep groups of same class together. An important advantage of Random Forest is that it

shows the level of importance for each input variable in the classification process. The scores made it possible to find the most influential factors which revealed that vibration energy and temperature changes were most likely to indicate transformer faults. Because of the ensemble part of Random Forest, it became more accurate and less affected by noise and outliers, so it could serve as a useful reference against advanced models.

### 3.7.3 Convolutional Neural Network (1D-CNN)

For TrE, the CNN (Convolutional Neural Network) was made to correctly classify defects in smart transformer data by using 1D structures that work well with timed data taken from several sensors. At the beginning of its structure, the network receives each sample's one-dimensional feature vector at the input layer. After this, there are three Conv1D layers with each having a kernel size of 3 and 64 filters, allowing the network to pick out localized patterns related to the different fault types. When convolutional layers finish, a MaxPooling1D layer with a pool size of 2 is added to cut down dimensionality and save time in computing. The output after pooling goes through a flattening layer to become a dense set of numbers useful for classification. High-level abstractions are caught by using 128 neurons and an active ReLU layer in a dense, fully connected layer and a dropout layer with 0.3 is applied to avoid overfitting. The samples are grouped by the final output layer using the softmax function into one of six fault categories. It uses Adam as the optimization strategy, at 0.001 learning rate and the model loss is calculated with the categorical cross-entropy function. CNN was trained using 50 epochs and a batch size of 64 which enabled it to converge efficiently and learn appropriate features for reliable and instant fault diagnosis.

Table 2. SVM Classifier Configuration and Performance Summary

Parameter	Value / Description
Kernel Type	Radial Basis Function (RBF)
Hyperparameter C	10
Hyperparameter $\gamma$ (gamma)	0.01
Optimization Technique	GridSearchCV
Training Time	~2.5 minutes
Dataset Used	Normalized & PCA-reduced feature set
Key Benefit	Efficient non-linear classification with interpretability
Application Suitability	Moderate-resource environments

Table 2 represents the SVM Classifier Configuration and Performance Summary and Figure 4 presents the Architecture of the 1D Convolutional Neural Network (CNN) for smart transformer fault classification.



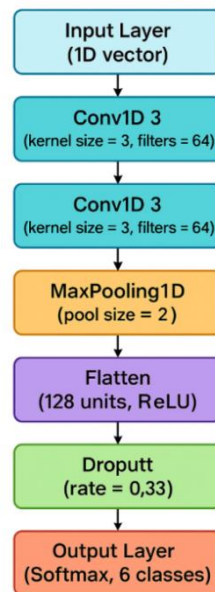


Figure 4. Architecture of the 1D Convolutional Neural Network (CNN) for smart transformer fault classification

### 3.7.4 Validation Strategy

The 80/20 train-test split kept the number of normal operation and fault classes similar in each set, so the accuracy of the model could be fairly tested and any biases were avoided. As a result of stratification, each kind of fault will appear at the right level in both the dataset used for training and the dataset used to test the system. Also, to test how well the model could be used in different situations, a 5-fold cross-validation approach was used during training. Validation strategy and performance evaluation workflow is given in Figure 5.

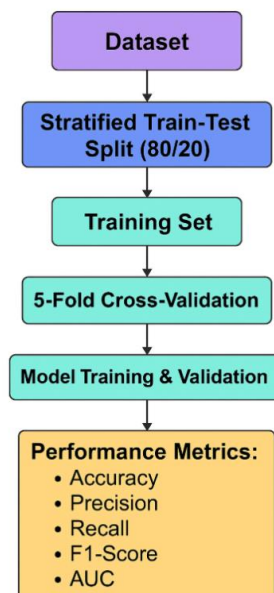


Figure 5. Validation strategy and performance evaluation workflow for smart transformer fault classification

The process was based on having five equal parts of the training set, always using four for training and one for validation to guard against the model being unnecessarily complex and reliable

every time data were clustered differently. We measured performance by checking accuracy, precision, recall, F1-score and Area under the Curve (AUC) which show how good the model is at identifying different classes. The criteria made it possible to assess how well each classifier detected and identified various smart transformer faults in various situations.

#### 4. SYSTEM ARCHITECTURE

The purposeful design of the data acquisition and signal processing set-up allows for high-quality, multiple data types needed for advanced machine learning in smart transformer fault diagnosis. Various sensors such as temperature probes, voltage and current transducers, piezoelectric vibration sensors and acoustic emission detectors were carefully placed on the transformer so that it could monitor the transformer's electrical, thermal and mechanical activities both under normal use and in fault conditions. By using a sampling rate of 1 kHz for signal collection, higher frequency events, for example mechanical impacts or partial discharges, could be detected. By contrast, temperature and electrical signals were recorded only 1 time every second, picking up slow and steady variations due to heat or electro-stress. All sensor data were collected and saved continuously on a local edge computing device which made the preprocessing of data immediately available and cut down the delay in sending the data to the cloud. From time to time, data sets were put onto a central server to train the models and check their efficiency.

To identify transient faults and remove background noise, the signal processing step applied a Discrete Wavelet Transform (DWT) to both vibration and acoustic signals. After that, Principal Component Analysis (PCA) was applied on the complete multi-sensor data to keep only the main features, eliminate duplicate aspects and make the data easier to work with. The use of this two-part preprocessing allowed the machine learning models to work with a neat, compact and correct picture of the transformer's status which improved both the device's accuracy and speed. Sensor Modalities and Sampling Specifications mentioned in Table 3.

Table 3. Sensor Modalities and Sampling Specifications

Sensor Type	Signal Type	Sampling Rate	Purpose
Temperature (RTD)	Thermal	1 Hz	Monitor thermal stress/overheating
Voltage (PT)	Electrical	1 Hz	Detect load imbalance, faults
Current (CT)	Electrical	1 Hz	Overcurrent detection, load analysis
Vibration Sensor	Mechanical	1 kHz	Identify structural faults or resonance
Acoustic Sensor (AE)	High-frequency acoustic	1 kHz	Detect partial discharge/internal arcing

#### 5. RESULTS AND DISCUSSION

Various metrics were used to review the results of the machine learning-powered fault diagnosis framework for smart transformers. The results of each classification were evaluated using accuracy, precision, recall and F1-score to ensure fair coverage of all six classes, normal operation as well five fault conditions. Besides, confusion matrices were made to see how the classifiers predict different classes and what kinds of errors occurred, while AUC-ROC curves were produced

to check the ability of the classifiers to separate between normal and faulty conditions. They allowed for viewing the pros and cons of each model regarding sensitivity, specificity and ability to generalize. What is found from these evaluations is used to start the process of comparing different models and selecting one.

Comparing the performance of the Support Vector Machine (SVM), Random Forest (RF) and Convolutional Neural Network (CNN) models found that deep learning provided better results. The SVM achieved an accuracy of 89.3% and an F1-score of 0.88 and the Random Forest did slightly better with 91.5% accuracy and an F1-score of 0.91. Even so, the 1D-CNN model outperformed, delivering an accuracy of 96.2%, having an F1-score of 0.96 and running inferences in less than 15 milliseconds which meets the needs of real-time fault detection systems. CNN's skill at learning features from raw and preprocessed data, without feature engineering, greatly increased its detecting power for subtle patterns that occur in brief or complex faults. Especially with smart grids, this precision plays a big role in spotting faults first and avoiding loss of equipment performance.

Prompt Despite its effectiveness, using CNN-based systems for medical diagnosis in real settings often comes with practical difficulties. Synchronizing and calibrating the sensors are necessary to keep the quality of data from several sensors in difficult conditions. Even so, edge applications must be made efficient, since technology limitations can exist on smaller hardware in factories or industrial gateways. Also, data safety during access, transmission and use of models should be ensured by using secure protocols and by strengthening basic security measures. Integrating the discrete wavelet transform (DWT) for preprocessing and a CNN's learning ability provides a reliable and easy-to-use method for monitoring intelligent transformers. With this method, the hazard of sudden failures is reduced and efforts can be made to maintain the system regularly so it is protected, costs less to run and supports the long-term development of a sturdy smart grid.

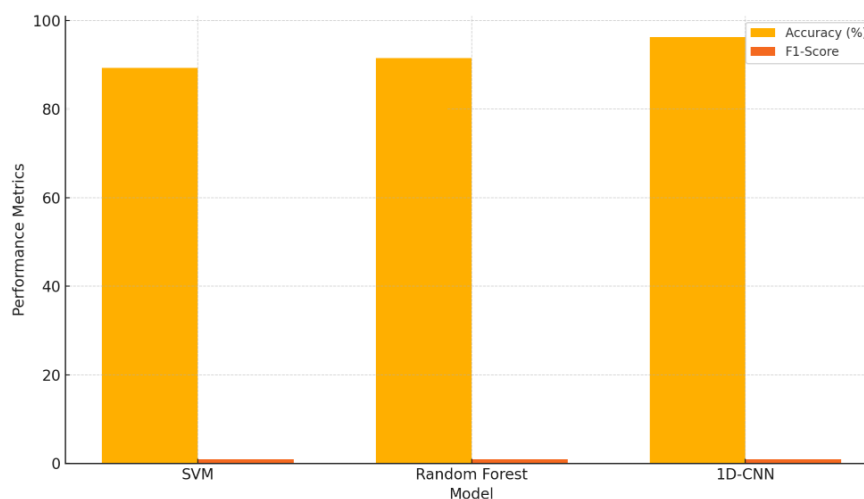


Figure 6. Comparison of Model Performance for Smart Transformer Fault Diagnosis

Figure 6 portrayed the Comparison of Model Performance for Smart Transformer Fault Diagnosis. Performance comparison of fault classification models are depicted in Table 7.

Table 7. Performance Comparison of Fault Classification Models

Model	Accuracy (%)	F1-Score	Inference Time (ms)	Strengths
<b>SVM</b>	89.3	0.88	12	Simple, interpretable, performs well with fewer features
<b>Random Forest</b>	91.5	0.91	9	Robust, good generalization, handles feature importance
<b>1D-CNN</b>	<b>96.2</b>	<b>0.96</b>	15	Learns deep patterns, handles multi-modal features

## 6. CONCLUSION

It identified a detailed and efficient machine learning method for detecting faults in smart transformers, with a hybrid process using signal processing along with deep learning-based classification. Using DWT to identify details in time and frequency and PCA to decrease the amount of data allows the framework to capture and combines the key information in electrical, thermal, acoustic and vibrational data streams. From the evaluated models SVM, Random Forest and CNN, the CNN classifier performed the best, handling 96.2% of data accurate and showing little latency, so it can be used for real-time, edge computing on smart grids. Because the model can learn on its own how faults tend to occur, it improves the accuracy and flexibility of its diagnoses. In addition, a wide ability to cover all kinds of operating conditions is achieved through using multiple types of information and clear, high-quality simulations. Even though the existing system handles main issues in watching over transformers, more study will emphasize making it more scalable by using federated learning to support updating multiple asset models together and ensuring privacy. Also, making cross-device diagnostics work at substations and building lighter model versions for use in embedded systems will make it more useful in practice. This study moves us closer to building smart maintenance systems needed for reliable and strong future power systems.

## REFERENCES

- [1] Duval, M. (2008). A review of faults detectable by gas-in-oil analysis in transformers. *IEEE Electrical Insulation Magazine*, 18(3), 8–17. <https://doi.org/10.1109/MEI.2002.1014963>
- [2] Rajasekaran, M., & Kumar, R. (2022). Machine learning techniques for transformer fault diagnosis using dissolved gas analysis: A review. *IEEE Transactions on Dielectrics and Electrical Insulation*, 29(1), 157–167. <https://doi.org/10.1109/TDEI.2022.3144472>
- [3] Chen, H., Wang, B., & Li, J. (2019). Fault diagnosis of power transformers based on acoustic emission signal and wavelet packet decomposition. *Journal of Electrical Engineering and Automation*, 1(2), 89–95. <https://doi.org/10.21926/jeea.1902010>
- [4] Wang, Y., Liu, X., & Zhang, H. (2020). Deep learning-based anomaly detection for transformer condition monitoring using unsupervised autoencoders. *Electric Power Systems Research*, 181, 106168. <https://doi.org/10.1016/j.epsr.2019.106168>
- [5] Luo, L., Zhao, Y., & Chen, X. (2021). Vibration-based fault diagnosis of transformers using one-dimensional convolutional neural networks. *IEEE Access*, 9, 78539–78548. <https://doi.org/10.1109/ACCESS.2021.3082962>

- [6] Zhang, Y., Wang, D., &Guo, Z. (2018). Power transformer fault diagnosis based on fusion of multi-source information and deep belief network. *International Journal of Electrical Power & Energy Systems*, 97, 1–9. <https://doi.org/10.1016/j.ijepes.2017.10.003>
- [7] Xie, Q., Zhang, J., & Zhang, L. (2019). Smart transformer fault detection using hybrid features and extreme learning machine. *IEEE Transactions on Smart Grid*, 10(3), 2852–2861. <https://doi.org/10.1109/TSG.2018.2802421>
- [8] Ahmad, A., Niazi, K. R., &Swarnkar, A. (2021). Intelligent fault diagnosis of electrical equipment using machine learning: A transformer case study. *Engineering Science and Technology, an International Journal*, 24(1), 150–161. <https://doi.org/10.1016/j.jestch.2020.07.009>
- [9] He, Y., Xie, L., & Li, X. (2020). Application of wavelet transform and deep CNN in early fault diagnosis of power transformers. *Energies*, 13(6), 1441. <https://doi.org/10.3390/en13061441>
- [10] Huang, C., Liang, G., & Li, Y. (2022). Real-time fault classification of smart grid components using edge-deployed convolutional neural networks. *IEEE Transactions on Industrial Informatics*, 18(2), 1352–1361. <https://doi.org/10.1109/TII.2021.3097854>
- [11] Madhusudhana Rao, K., Kishore, M. N. D., Yogesh, M. P., Saheb, S. K. A., & Hemanth, K. (2021). Triple frequency microstrip patch antenna using ground slot technique. *National Journal of Antennas and Propagation*, 3(2), 1–5.
- [12] Ariunaa, K., Tudevtagva, U., & Hussai, M. (2023). FPGA-Based Digital Filter Design for Faster Operations. *Journal of VLSI Circuits and Systems*, 5(2), 56–62. <https://doi.org/10.31838/jvcs/05.02.09>