©2025 Published in 11th International Symposium on Intelligent Technologies in Engineering and Science 16-18 May 2025 (ISITES2025 Diyarbakır - Turkey)

High impedance fault detection with DWT based feature extraction and machine learning in low voltage distribution systems

Nur Bağnu Polat and İlker Dursun
Faculty of Technology, Department of Electrical and Electronic Engineering Sakarya University of Applied Sciences, Turkey
Faculty of Technology, Department of Electrical and Electronic Engineering Sakarya University of Applied Sciences, Turkey

Abstract

In this study, a method for the detection of high impedance faults (HIF) in distribution systems is developed. The method uses three-phase current measurements from a 900 meter long low voltage line. For feature extraction, the current signals are processed using the Discrete Wavelet Transform (DWT) method using basic features including root mean square (RMS), mean, standard deviation, maximum value, energy and entropy. Artificial Neural Network (ANN) and Support Vector Machine (SVM) classifier models are used for fault detection. An accuracy of over 99% was achieved with the DWT+ANN approach. Approximately 98% accuracy was achieved with the DWT+SVM approach. It is demonstrated that DWT-based feature extraction combined with machine learning classifiers achieves successful results for HIF detection in low voltage networks.

Key words: Artificial Neural Network (ANN), Discrete Wavelet Transform (DWT), Support Vector Machine (SVM), high impedance faults (HIF)

1. Introduction

In electrical power systems, faults are generally classified as open-circuit faults and short-circuit faults. Among short-circuit faults, the most frequently encountered type is the single-phase-to-ground fault, which is categorized under high impedance faults (HIF) [1]
HIFs occur as a result of broken conductors making direct or indirect contact with the ground, and since the fault current is close to the nominal current, it becomes difficult for existing protection equipment to detect these faults [2].
Protection systems in networks are capable of detecting only 17.5% of high impedance faults, which can lead to significant risks of injury, fatality, and property damage [3]. Moreover, the absence of visible electrical arcs in passive fault types further complicates the detection of such faults [4]. Although it has been indicated in the literature that fault detection can be performed by utilizing network imbalances, it is emphasized that conventional methods are insufficient for detecting HIFs [5].

*Nur Bağnu Polat: Address: Faculty of Technology, Department of Elektrical-Electronic Engineering Sakarya University of Applied Sciences, 54187, Sakarya TURKEY. E-mail address: nurbagnupolat@subu.edu.tr, Phone: +0264 616 0000-18-13

Table 1. Methods and Performance for High Impedance Fault (HIF) Detection

Ref.	Applied Methods	Test System	Objective and Performance
[6]	Normalized RMS current-voltage signal + Artificial Neural Network (ANN)	IEEE 9-bus system	Developed an ANN-based method using normalized RMS signals, achieving 100% fault detection accuracy and over 96% fault location accuracy.
[7]	Discrete Wavelet Transform (DWT) + ANN	IEEE 34-bus system	Applied 5-level DWT using the db3 wavelet family, extracted statistical features from current signals, and classified them using ANN with 99.5% accuracy.
[8]	DWT + CNN, RBFNN, SVM, NARX, PNN, FFNN, ANFIS	IEEE 5-bus system	Features extracted via DWT were classified using different AI models, and the RBFNN model demonstrated superior performance compared to others.
[9]	DWT + Random Forest (RF), Logistic Regression (LR), Multilayer Perceptron (MLP), Naive Bayes (NB)	IEEE 13-bus system + PV integration	Random Forest algorithm achieved higher accuracy than other methods for HIF detection.
[10]	K-Nearest + MLP + ANN, K-Nearest + KNN	IEEE 13-bus system	HIF classification was performed using K-Nearest-based approaches combined with ANN and KNN algorithms.
[11]	Empirical Wavelet Transform (EWT) + Teager Energy Operator (TEO) + ANN	4-bus system	High fault detection and classification accuracy were achieved using a scalogram and 2D-CNN model.
[12]	Continuous Wavelet Transform (CWT) + Scalogram + 2D-CNN	3-bus system	Sonuçlar, 2D-CNN modelinin üstün arıza tespit doğruluğu ve sınıflandırma yeteneklerini göstermektedir
[13]	Wavelet Optimum Feature Extraction + Frequency Domain Analysis (WONC-FD)		A training-free method was developed using DWT for time-frequency analysis and classified faults based on minimum Mean Square Error (MSE).
[14]	Voltage-Current (VI) + ANFIS, VI + SVM	IEEE 12-bus system	VI signals were used to develop ANFIS and SVM-based models, and ANFIS demonstrated superior performance compared to traditional methods.
[15]	DWT + SVM + Chaotic Initialization-Stable State Search (CI-SSS)		Features were extracted using DWT, and SVM was optimized with CI-SSS, achieving 98.82% classification accuracy.
[16]	Diferansiyel indeks	IEEE 37-bus / IEEE 9-bus systems	The DI method, using positive sequence currents and voltages, was able to detect faults within two and a half cycles in both simulation and RTDS real-time platform.

[17]	DWT (db4 wavelet family)	IEEE 16-bus system	DWT was applied using the db4 wavelet; asymmetric wavelets were preferred due to the irregular nature of HIF signals.
[18]	Residual Current + SVM	8-bus system	Nearly 100% accuracy was achieved for LG and LLG faults, while classification accuracy was relatively lower for LL and LLL faults.
[19]	S-Transform (ST)	6-bus system	Using the ST method, a fault detection accuracy of 98.48% and a non-fault event detection rate of 98.1% were achieved.

In this study, high impedance fault (HIF) scenarios are prepared on a low voltage (400 V) distribution line with a length of 900 meters. Discrete Wavelet Transform (DWT) was used to extract features of the signal. These features are given to ANN and SVM to determine whether the fault is present or not. The obtained results and success rates are presented.

2. Materials and Method

2.1. Discrete Wavelet Transform

Discrete Wavelet Transform (DWT) is a method used to analyze non-stationary signals by capturing both time and frequency domain information of a signal. With this feature, DWT stands out more than Fourier transform, which works only with frequency components, or Hilbert-Huang transform, which focuses directly on instantaneous frequency information.
DWT subjects the input signal to discrete filtering and sampling processes to decompose it into different resolution levels based on the selected mother wavelet. Thanks to this structure, transient events and sudden changes are detected more effectively. Although DWT works similarly to the continuous wavelet transform, it is a numerical version that applies discretization at the time and scale levels, and its basic mathematical expression is given in equation 1.

$$DWT_{(m,n)} = \frac{1}{\sqrt{2^m}} \sum_k f(k)\psi\left(\frac{n - k2^m}{2^m}\right) \tag{1}$$

Here, m represents the scale, n denotes the position, and ψ is the mother wavelet function.
Among various wavelet families, Daubechies (db), Symlets (sym), and Coiflets (coif) are widely used. In this study, the db4 wavelet was selected because of its strong performance with signals that have a high sampling rate and transient behavior.

2.2. Support Vector Machines

Among supervised learning algorithms, Support Vector Machines (SVM) have proven effective for both classification and regression tasks. SVM aims to find a hyperplane that separates different classes with the maximum margin. This hyperplane is expressed as:

$$w^T x + b \tag{2}$$

where w is the weight vector, x is the input feature vector, and b is the bias term.
Linear SVM is used if the dataset can be linearly decomposed. While Linear SVM provides simple and efficient classification by directly creating a hyperplane in the original feature space without the need for kernel functions, if the dataset cannot be linearly separated, the Kernel SVM method is applied. In Kernel SVM, the data is transferred to a higher dimensional feature space using kernel functions, making it possible to create a separating hyperplane. Commonly used kernel types include polynomial, radial basis function (RBF), sigmoid and user-defined kernel functions. The choice of kernel is determined by the nature and complexity of the data.

Table 1. Types of Svm

LİNEER SVM	**KERNEL SVM**
A hyperplane is directly selected if the data is linearly separable.	Polynomial Kernel
	RBF (Radial Basis Function/Gaussian) Kernel
	Sigmoid Kernel
	Custom (user-defined) Kernel

2.3. Artificial Neural Network

Artificial Neural Networks (ANN) are a computational method inspired by the biological nervous system. Each neuron calculates the weighted sum and bias term of the inputs it receives and applies them to an activation function. In this way, complex and non-linear relationships between inputs and outputs can be learned.

When calculating the activation function, each neuron calculates the sum of the weights of the inputs and the bias term. it learns the relationships between inputs and outputs by taking the updated weights at each cycle. The output of a neuron is expressed as follows:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{3}$$

x_i represents the input data, w_i represents the weights, b is the bias term, and f is the activation function.
The aim of training is to minimize the prediction error with the calculated weights. While parameters such as the number of neurons and the choice of parameters affect the performance of the model, it is also very important that the data set is sufficient and balanced.

2.4. Application of the Proposed Method

In this study, a low voltage (400 V) distribution line is modeled to detect high impedance faults (HIF). The system consists of a 900 meter long feeder and a 100 kVA distribution transformer. Three unbalanced loads are connected at the end of the feeder. Unbalanced is chosen to match the real grid parameters because in a real grid the system is not balanced.Each load is designed to have approximately 6500 W active power and -2150 VAR capacitive reactive power. Thus, the total

active power of the system is about 19.6 kW and the total capacitive reactive power is about 6.45 kVAR.
Matlab/Simulink 2019b program was used for the grid design and the design is presented in Figure 1.

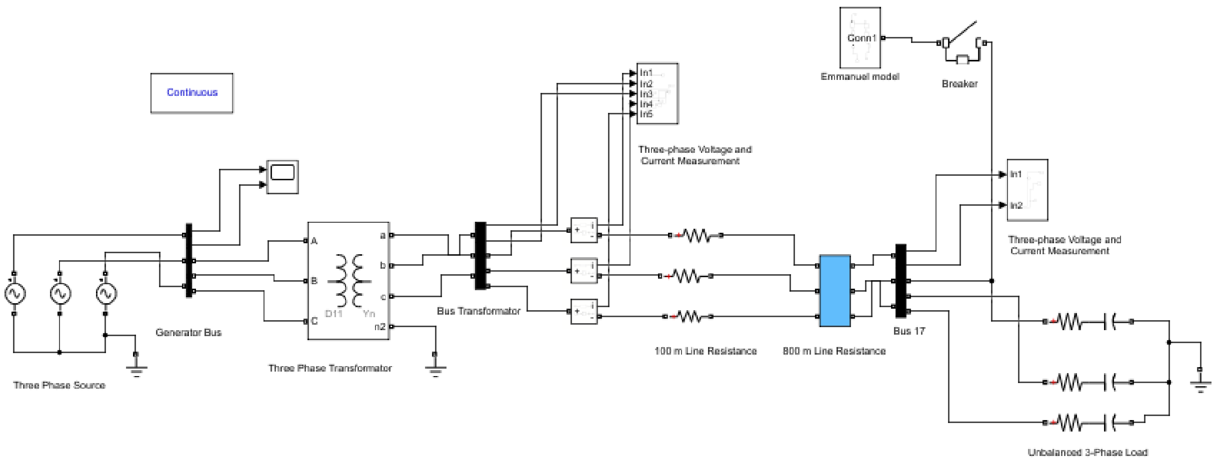

Figure 1. Three-Phase Unbalanced Load and High Impedance Fault Simulation Setup

High impedance faults are designed at 100 meters at the first output of the transformer and every 50 meters thereafter.

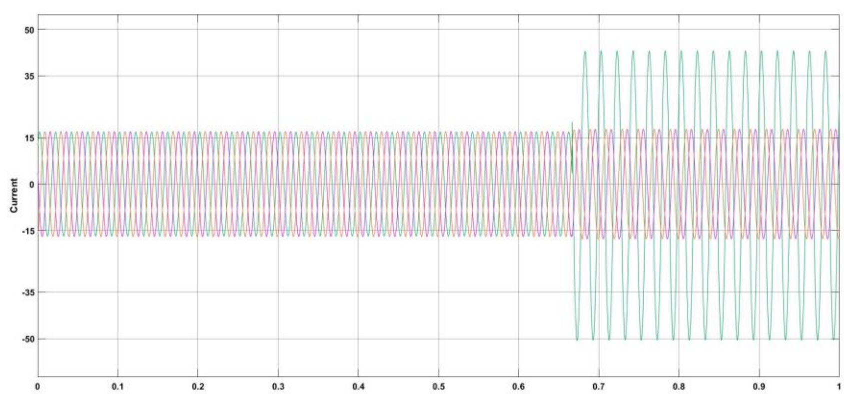

Figure 2. Phase currents when HIF occurs at 900 meters

Discrete wavelet transform (DWT) method was used to extract features from three-phase current signals. db4 (Daubechies 4) was selected as the main wave and 5-level decomposition was applied. Six features, namely MS, mean, standard deviation, maximum value, energy and entropy, are used for feature extraction with DWT and these features are made for each phase. The extracted features were trained with ANN and SVM for classification. The performance of both classification models was evaluated based on accuracy, confusion matrix, ROC curve and AUC.

A single-layer structure with 10 nodes was designed for ANN and scaled conjugate gradient was used for training the algorithm. The feature data obtained with DWT was divided into three parts as 70% training, 15% validation and 15% testing.

For SVM, radial basis function (RBF) kernel was preferred and 80% of the features obtained from DWT were divided into training and 20% testing.

The results obtained from the designed system showed that both ANN and SVM models have high performance in detecting high impedance faults.

3. Results

In this section, the results of DWT + ANN and DWT + SVM models developed for high impedance fault detection are presented.

3.1. DWT + SVM Method Results

In this study, the features extracted using DWT were applied to high impedance fault (HIF) detection with the support vector machine (SVM) model.
The confusion matrix is given in Figure 3 and allows us to analyze the classification success of the model in detail. When the matrix belonging to the DWT + SVM method is examined: the model correctly classified 111 out of 114 data that were faultless (97.36% accuracy), and only 3 faultless data were incorrectly classified as faulty. All 70 faulty data were correctly predicted as faulty, meaning 100% detection success was achieved for faulty data.

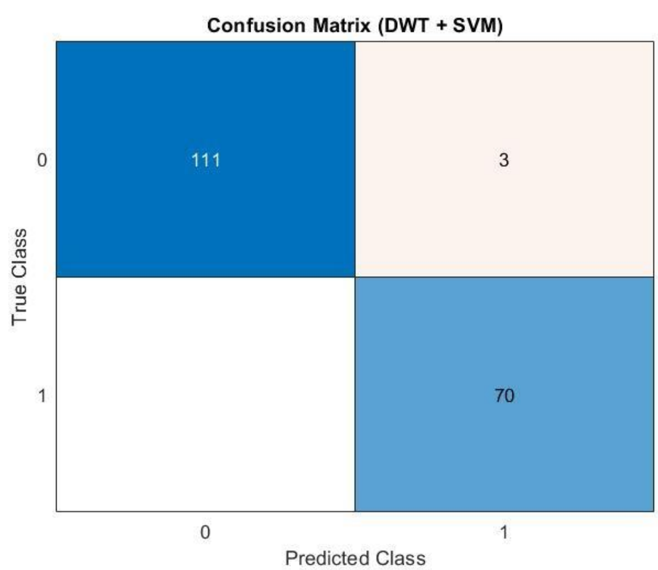

Figure 3. Confusion Matrix for DWT + SVM

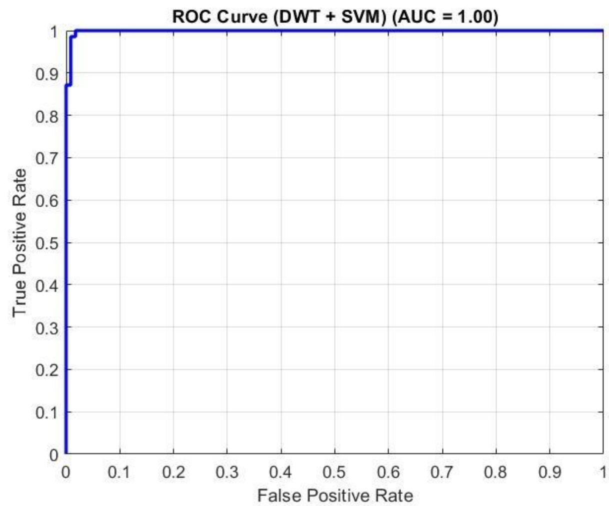

Figure 4. ROC Curve for DWT + SVM

When the ROC curve in Figure 4 is examined, it can be observed that the true positive rate quickly approaches 1 while the false positive rate remains at a minimum. This situation reveals that the model classifies faulty and non-faulty cases with high accuracy. The ROC curve is consistent with the confusion matrix. Both results support each other and confirm the classification reliability of the model.

3.2. DWT + ANN Method Results

In this study, the features extracted by DWT were used to classify high impedance faults (HIF) using an artificial neural network (ANN) model. The results obtained are presented in Figures 5, 6, 7.

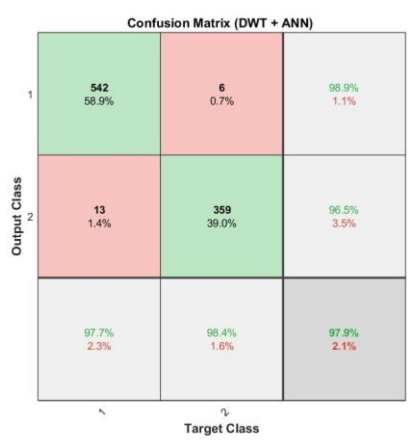

Figure 5. Confusion Matrix for DWT + ANN

When the confusion matrix in Figure 5 is examined carefully, the success rate was found to be 98.90% by correctly classifying 542 out of 548 non-faulty data. Only 6 non-faulty data were classified as faulty and incorrectly classified as faulty. Similarly, the success rate was found to be 96.5% by correctly classifying 359 out of 372 correctly classified as faulty. Only 13 faulty data were classified as faultless. Considering these results, the ANN model is highly successful in classifying both faulty and non-faulty data.

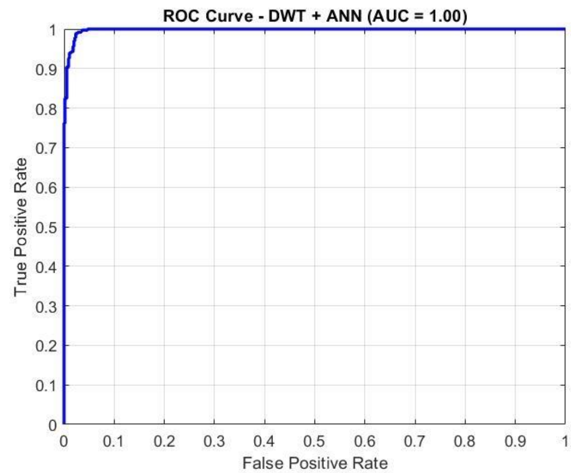

Figure 6. ROC Curve for DWT + ANN

When the ROC curve in Figure 6 is examined, it can be observed that the true positive rate quickly approaches 1 while the false positive rate remains at a minimum. This situation reveals that the model classifies faulty and non-faulty cases with high accuracy. The ROC curve is consistent with the confusion matrix. Both results support each other and confirm the classification reliability of the model.

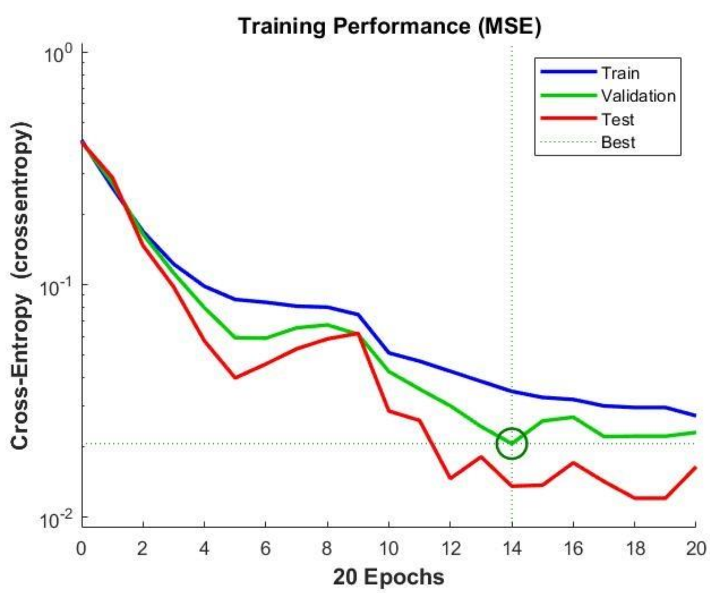

Figure 7. Training Performance for DWT + ANN

In Figure 7, the training, validation and test curves exhibit a consistent decrease, which represents effective learning. The lowest validation error was obtained at the 14th epoch. This situation shows the optimum model performance obtained at this point. The proximity of the three curves to each other shows that it has good generalization ability and perform well.

4. Discussion

In this study, the dataset processed using with Discrete Wavelet Transform (DWT) based feature extraction method for the detection of high impedance faults in low voltage distribution systems was evaluated using Artificial Neural Network (ANN) and Support Vector Machine (SVM) classifiers. The time-frequency based features extracted via DWT enabled both models to distinguish healthy and faulty situations with high accuracy [7], [17]. The correlation coefficient (R = 0.96632) obtained for the ANN model and the AUC value (1.00) calculated for the SVM model highlight the strong classification performance [18].

It was observed that the classification algorithms such as ANN and SVM, which are widely adopted in the literature, also demostrated effective performance on a low voltage, long distance and unbalanced loaded system in this study. In the future, it is recommended that these models be tested under different load profiles, variable sampling rates and other fault types and their robustness be evaluated.

5. Conclusion

In this study, Discrete Wavelet Transform (DDT) based feature extraction was performed for the detection of high impedance faults on a distribution system with low voltage and unbalanced loads. The obtained features were evaluated with Artificial Neural Network (ANN) and Support Vector Machine (SVM) classifiers and the successful results obtained by both models revealed the suitability of the methods used for high impedance fault detection.

Thanks to the multiple fault scenarios created with different Emmanuel model parameters, it was observed that the proposed method can work compatible with suddenly changing fault resistances and real-time fault conditions.

The findings show that both ANN and SVM models provide reliable, flexible and practical solutions in high impedance fault detection.

The conceptual development of this study was influenced by the approaches and findings explored in Project No. 122E675, supported by TÜBİTAK under the 3501 Career Development Program. The study, which was built on the balanced load model developed in the author's master's thesis, was extended to include unbalanced load scenarios and restructured using DWT-based feature extraction together with ANN and SVM models.

The authors would like to express their gratitude to Project No. 122E675 for its conceptual contribution.

References

[1] Chen, J., Phung, T., Blackburn, T., Ambikairajah, E., & Zhang, D. (2016). Detection of high impedance faults using current transformers for sensing and identification based on features extracted using wavelet transform. *IET Generation, Transmission and Distribution*, *10*(12), 2990–2998. https://doi.org/10.1049/iet-gtd.2016.0021

[2] Lavanya, S., Prabakaran, S., & Kumar, N. A. (2022). Behavioral dynamics of high impedance fault under different line parameters. *energy*, *9*, 12.

[3] Wang, S., & Dehghanian, P. (2020). On the Use of Artificial Intelligence for High Impedance Fault Detection and Electrical Safety. *IEEE Transactions on Industry Applications*, *56*(6), 7208–7216. https://doi.org/10.1109/TIA.2020.3017698

[4] Mahanty, R., & Gupta, P. (2004). Voltage stability analysis in unbalanced power systems by optimal power flow. *IEE Proceedings-Generation, Transmission and …*, *151*(3), 201–212. https://doi.org/10.1049/ip-gtd

[5] Güllüdereli, M. (2007). *Enerji iletim sistemlerinde arıza analizinin otomasyonel çözümleri*. Sakarya Üniversitesi.

[6] Mondal, S., & Pradhan, R. (2024, November). Fault Classification for High Impedance Faults Using Back Propagation Based Artificial Neural Network Technique. In *2024 International Conference on Sustainable Power & Energy (ICSPE)* (pp. 1-6). IEEE

[7] Attar, M. S., & Miveh, M. R. (2025). High-Impedance Fault Detection in Distribution Networks Based on Support Vector Machine and Wavelet Transform Approach (Case Study: Markazi Province of Iran). *Energy Science & Engineering*, *13*(3), 1171-1183.

[8] Cano, A., Arévalo, P., Benavides, D., & Jurado, F. (2024). Integrating discrete wavelet transform with neural networks and machine learning for fault detection in microgrids. *International Journal of Electrical Power & Energy Systems*, *155*, 109616.

[9] Mahzan, N. N., Othman, M. L., Wahab, N. I. A., Veerasamy, V., Salim, N. A., Abidin, A. A. Z., & Islam, S. Z. (2025). Performance Analysis of Intelligent Classifiers for High Impedance Fault Detection in a PV-Integrated IEEE-13 Bus System Analyse des performances des classificateurs intelligents pour la détection des défauts à haute impédance dans un système de bus IEEE-13 intégré au système PV. *IEEE Canadian Journal of Electrical and Computer Engineering*.

[10] Guttimari, T. Y. (2024, August). K-Nearest Neighbor Based High Impedance Fault Detection for Radial Feeder. In *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)* (pp. 1-8). IEEE.

[11] Farkhani, J. S., Çelik, Ö., Ma, K., Bak, C. L., & Chen, Z. (2024). Fault Detection, Classification, and Location Based on Empirical Wavelet Transform-Teager Energy Operator and ANN for Hybrid Transmission Lines in VSC-HVDC Systems. *Journal of Modern Power Systems and Clean Energy*.

[12] Nayak, P., Das, S. R., Mallick, R. K., Mishra, S., Althobaiti, A., Mohammad, A., & Aymen, F. (2024). 2D-convolutional neural network based fault detection and classification of transmission lines using scalogram images. *Heliyon*, *10*(19).

[13] Sakovich, N., Aksenov, D., Pleshakova, E., & Gataullin, S. (2025). Wavelet-Based Optimization and Numerical

Computing for Fault Detection Method—Signal Fault Localization and Classification Algorithm. *Algorithms*, *18*(4), 217.

[14] Kurmaiah, A., & Vaithilingam, C. (2025). Optimization of Fault Identification and Location Using Adaptive Neuro-Fuzzy Inference System and Support Vector Machine for an AC Microgrid. *IEEE Access*.

[15] Kumar, V. R., & Jeyanthy, P. A. (2025). Fault Classification and Detection in Transmission Lines by Hybrid Algorithm Associated Support Vector Machine. *Transactions on Emerging Telecommunications Technologies*, *36*(3), e70034.

[16] Villarreal, R., Chamorro-Solano, S., Vega-Sampayo, Y., Espejo, C. A., Cantillo, S., Gaviria, L., ... & Montoya, C. (2025). A New Approach to Electrical Fault Detection in Urban Structures Using Dynamic Programming and Optimized Support Vector Machines. *Sensors*, *25*(7), 2215.

[17] Samal, S., Samantaray, S. R., & Sharma, N. K. (2025). A New Differential Index-Based Fault Detection Scheme for Microgrids. *IEEE Transactions on Industry Applications*.

[18] Shalby, E. M., Abdelaziz, A. Y., Ahmed, E. S., & Abd-Elhamed Rashad, B. (2025). A comprehensive guide to selecting suitable wavelet decomposition level and functions in discrete wavelet transform for fault detection in distribution networks. *Scientific Reports*, *15*(1), 1160.

[19] Janiabadi, A. A., & Hasheminejad, S. (2025). A new algorithm for the identification of high impedance faults in distribution systems utilizing S transform. *Ain Shams Engineering Journal*, *16*(4), 103334.