Journal of Emerging Innovations in Engineering (JEIE)
Volume 2, Number 03, October 2026
ISSN: 3110-3375
Open Access: https://jurnal.sutajaya.org/jeie

Design and Simulation of an AI-based DWT-ANN model for high impedance fault detection and classification in Nigeria distribution

Yusuf Mohammed[1,*] Babalola Wasiu Lasun[2] Muhammad Usman Iliyasu[3]

[1,3] Department of Electrical & Electronics Engineering Technology, Federal Polytechnic, Bali, Nigeria
[2] Department of Technical Education, Federal College of Education, Gidan Madi, Sokoto State, Nigeria

Article Info

Received August 14th, 2026
Revised September 19th, 2026
Accepted October 17th, 2026

Keywords:
High Impedance Fault (HIF); Discrete Wavelet Transform (DWT); Artificial Neural Network (ANN); Distribution Network Protection; Fault Classification; Artificial Intelligence (AI).

DOI:
https://doi.org/10.65664/jeie.v2i03.21
Page: 110 - 134

Correspondent Author:
Yusuf Muhammad
Email: ym2100020.mee@buk.edu.ng

ABSTRACT

High Impedance Faults (HIFs) challenge conventional protection because their low, nonlinear currents stay below overcurrent relay thresholds, causing fire hazards, equipment damage, and reliability risks. This study presents an artificial intelligence framework for HIF detection and classification in a Nigerian distribution network, integrating discrete wavelet transform feature extraction with an artificial neural network classifier. A 35-bus, 11 kV Kano Electricity Distribution Company system was modeled in MATLAB/Simulink under three load levels (100%, 70%, 45%), eight fault resistances (100–1000 Ω), eight inception angles (0°–180°), and five fault types: no fault, single-, double-, and three-phase HIF, and three-phase short circuit. Transient currents were decomposed using the Daubechies-4 (DB4) wavelet at level 4, chosen from eight wavelet families. Twelve features, including energy coefficients, statistical measures, and signal energy ratios, trained a 12-64-32-16-5 ANN with dropout and Adam optimization. It achieved 81.60% accuracy, outperforming K-Nearest Neighbors (77.87%), Naïve Bayes (61.07%), Support Vector Machine (79.23%), XGBoost (79.87%), and Long Short-Term Memory (80.12%). Precision (81.49%), recall (81.89%), F1-score (81.43%), ROC-AUC above 0.95 for all classes, and computational efficiency confirm robustness. DWT-ANN offers an effective, reliable HIF detection solution for networks modeled with actual utility parameters, sensitive to nonlinear arc signatures and resilient to noise and variability.

© 2026 The Authors. Published by PT Pustaka Intelektual Sutajaya, collaboration with Faculty of Engineering Universitas 17 Agustus 1945 Cirebon. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

INTRODUCTION

The safe and reliable operation of power distribution networks is significantly dependent on the capacity of the protection system to detect abnormal situations. One of the most recurrent challenges in maintaining the integrity of the network is the occurrence of the High Impedance Fault (HIF). The HIF happens when a line comes into contact with non-conducting materials, like dry soil, asphalt, and trees. Unlike the traditional short circuit, the HIF fails to produce enough current levels to activate the overcurrent relay, making its detection exceedingly complex. The HIF results in periodic arcs, which cause asymmetric, nonlinear, and weak currents that mimic normal loading. If HIFs remain unnoticed, they pose a fire hazard, equipment damage, and safety risks. Therefore, extensive literature exists on the detection of HIFs.

Throughout several decades, several different HIF detection techniques have been developed using classical time and frequency domain analysis, including Fourier transform (FT), short-time Fourier transform (STFT), and others. The disadvantage of FT and STFT is that their fixed windows limit their ability to analyze transients and non-stationary components that occur during arc fault events. Wavelet transforms, especially the discrete wavelet transform (DWT), have become very popular because they can analyze signals through multi-resolution and break down voltage and current waveforms into localized time-frequency domains. Therefore, DWT is better suited to capture the transient and rapidly varying characteristics of HIFs. In parallel with the development of signal processing techniques, various AI and machine learning techniques have been developed to classify HIFs, such as adaptive network-based fuzzy inference systems (ANFIS), support vector machines (SVM), convolutional neural network (CNN), bi-directional long short-term memory (Bi-LSTM), k-nearest neighbors (KNN), naive Bayes, and others. Although all models mentioned above achieve remarkable results, they only use simplified or purely simulated networks. They do not consider the noisy environment of real utility feeders with imbalanced loads and unexpected system dynamics.

Despite the significant advancements made, high detection accuracy remains elusive when applied in practical power distribution networks, thereby forming a critical gap in the current literature. While several approaches may fail to synthesize across various operational modes, other techniques lack the ability to accurately classify between different types of HIFs and non-fault scenarios. Additionally, most research efforts to date measure their performances using a few metrics but do not carry out a multidimensional analysis of accuracy. Driven by these challenges, this paper proposes a novel hybrid model that combines the use of the discrete wavelet transform (DWT) in extracting relevant features alongside an Artificial Neural Network (ANN) in classifying fault types. Contrary to many approaches presented thus far, the new algorithm will be tested on a practical 35-bus distribution line associated with the Kano Electricity Distribution Company (KEDCO). "Practical" here meant the study used KEDCO's actual line/transformer parameters, actual recorded feeder line loads (at 100%, 70%, 45%), the actual Dakata 35-bus feeder architecture, and actual arc-fault parameters calibrated to field recordings obtained, admittedly sparsely. This isn't a true "field validated study," and the paper wisely makes that distinction "practical" vs. "field valid, with field valid work scheduled in future study."

Table 2 summarizes typical existing papers for HIF detection employing AI-based techniques in relation to the position of the current work, for clear identification. Based on a simulation using real utility data, applied to the KEDCO 35-bus Nigerian distribution network, our work presents an 81.60% accuracy using a multi-domain 12 features feature set.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

Table 2: Comparative Analysis of Representative HIF Detection Studies

Study	Method	Network	Data Source	Features	Accuracy	Validation
Ghaderi et al. (2017)	Wavelet + SVM	IEEE 13-bus	Simulation	Time frequency	87%	Simulation only
Hong & Huang (2015)	WT + GA+ SVM	IEEE 34-bus	Simulation	Wavelet coefficients	89%	Simulation only
Ahmadi et al. (2022)	SVM	IEEE 33-bus	Simulation	Harmonic analysis	84%	Simulation only
Abasi-obot et al. (2023)	CNN	33Kv/11Kv (Nigeria)	Simulation	Raw current	95.76%	Simulation only
Allawi et al. (2024)	ANN	IEEE 13- bus	Simulation	Spectrum+ANN	87%	Simulation only
Attar & Miveh (2025)	SVM+WT	Iranian Network	Simulation	WT feature	89.2%	Simulation only
Our Study	DWT+ANN	KEDCO 35-bus (Nigeria)	Simulation (Utility Data)	Multi-domain (12 features)	81.60%	Simulation with utility data

The purpose of this research work is to develop and evaluate an AI-based HIF detection and classification technique whose accuracy in classification will be improved by applying DWT-based feature extraction and ANN-based pattern recognition. The specific objectives involve modeling the KEDCO 35-bus distribution system, extraction of HIF signatures using DWT techniques, training of ANN classifiers for HIF classification purposes, and conducting thorough accuracy evaluation using established evaluation metrics in comparison with KNN and Naïve Bayes classifiers.

A unique contribution of the present study to the literature comes with the four following aspects: This is one of the first studies to investigate HIF detection in the context of Nigerian distribution network with real utilities' parameters. In comparison to the prevalent practice, the simulation implemented herein considers twelve unique parameter settings (three loads, eight resistance values, five faults, and multiple fault angles). This study also provides a formal investigation on the optimum choice of mother wavelets justified by the concept of Energy- to –Entropy ratio. Lastly, a rigorous multi-metric comparison study of proposed method is performed by benchmarkin6 of its performances against 6 different classifiers based on eight metrics.

The main contributions of this study are summarized as follows: (a). First application of a DWT-ANN hybrid framework tuned for a live Nigerian 35-bus distribution network, using operational data, rather than data from typical IEEE benchmark buses or artificial network. (b). Comprehensive wavelet selection analysis of 8 mother wavelets namely, Haar, DB2, DB4, DB8, Sym4, Sym8, Bior3.1, Bior3.3 on the HIF characteristics of distribution networks - to a degree previously unprecedented in empirical Nigerian studies. (c). Creation of a strong features extraction process that harmonizes DWT decomposition coefficient and statistical features to uniquely express nonlinearities of arc trajectory of HIFs being hidden away by conventional frequency methods. (d). Introduction of realities of disturbance such as varying loading conditions; Peak, normal and off-Peak levels; diverse fault resistances 100-1000; and also varying fault inception angles between (0 – 180°) typical on actual distribution system condition found in Nigerian environment. (e) Rigorous cross-validation and benchmarking were conducted using KNN, Naive Bayes, SVM, Random Forest, XGBoost, and LSTM on metrics such as accuracy, confusion matrix, precision, recall, F1-score, ROC-AUC, and computing time, establishing a comparative standard for a network designed specifically for the Nigerian power network.

Although various combinations of DWT-ANN have been reported in the literature, these studies, for the most part, were tested on IEEE test systems like 13-bus, 34-bus or 123-bus with very simplistic

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

models of fault phenomena. This study addresses some real-world problem constraints characteristic of the Nigerian distribution networks, like imbalanced loading, network configuration changes, atypical arc characteristics in tropical climatic conditions. It is the first work to systematically validate a DWT-ANN HIF detection algorithm against real data obtained from a utility provider in Nigeria.

Therefore, the remainder of this paper is structured as follows: Section 2 presents the system modelling and DWT-based feature extraction process. Section 3 describes the design of the ANN classifiers and the comparative models (KNN and Naïve Bayes). Section 4 discusses the simulation results and provides detailed accuracy assessment across all performance metrics. Section 5 concludes the study and offers recommendations for future study, including field validation and extension to fault location.

2. Distribution System Modelling and DWT-Based Feature Extraction

This section presents the modelling of the Dakata 35-bus distribution feeder of the Kano Electricity Distribution Company (KEDCO), the high impedance fault (HIF) modelling approach, the signal-acquisition setup, and the feature extraction procedure based on the Discrete Wavelet Transform (DWT). These processes constitute the analytical foundation for the AI-based fault classification framework presented in this study.

2.1 Modelling of the KEDCO 35-Bus Distribution Network

The Dakata 35-bus distribution feeder was modeled using actual parameters obtained from KEDCO. Table 2. Presents the key network parameters including line impedances, transformer ratings, and load data. Three loading conditions were considered: peak load (100% of rated capacity), standard load (70%), and off-peak load (45%). These variations were implemented to evaluate the robustness of the proposed detection scheme under realistic operating conditions. Its single-line diagram is shown in the Figure 1:

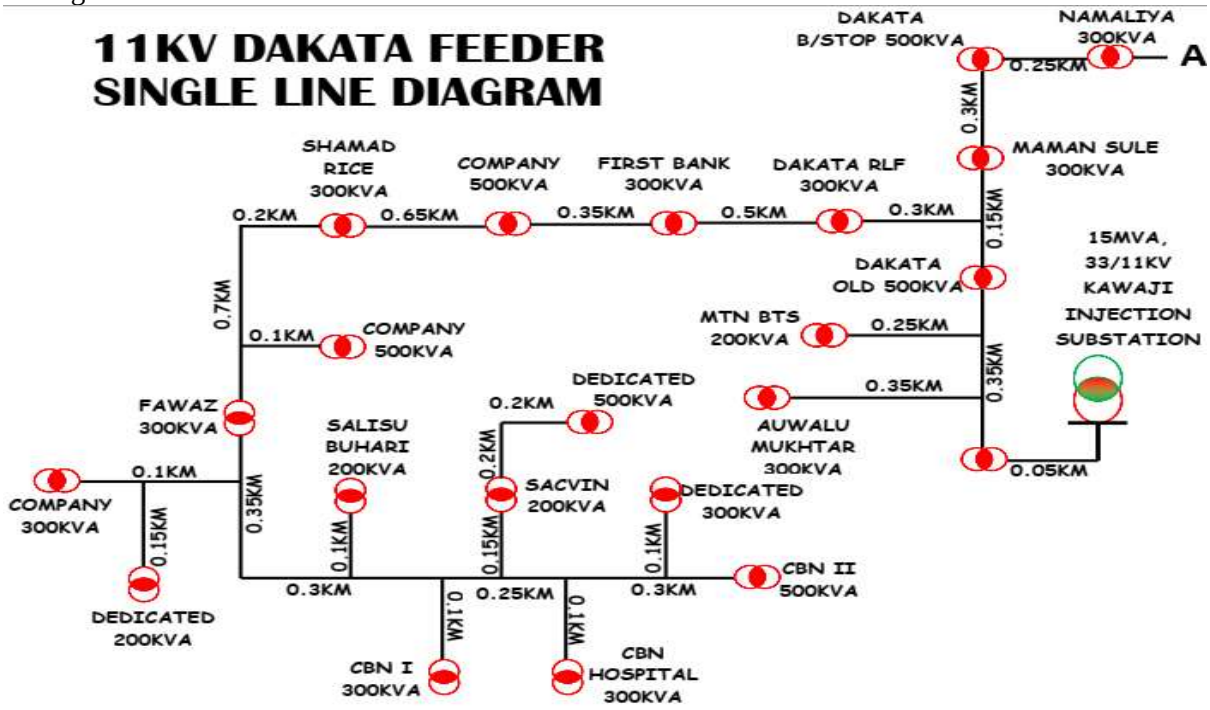


Figure 1: Single Line Diagram of Dakata 35-Bus Distribution System

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

Table 3: KEDCO 35-Bus Network Parameters

Parameter	Value	Unit
Nominal voltage	11	KV
Frequency	50	Hz
Total Connected Load	3.5	MVA
Number of Distribution Transformer	8	-
Transformer Rating (Each)	500	KVa
Line Impedance Range	0.1-0.5	Ω/km
Base MVA	10	MVA
Base Voltage	11	KV

Distribution lines were represented by their series impedance as shown in equation (1)

$$Z = R + jX \quad \text{.....(1)}$$

Where R represents resistance, and X represents reactance. These values will be substituted from the line and bus data that was collected from KEDCO data center.

The impedance impacts power flow and introduces voltage drops along the line. Each load connected to the network contribute to reactive power demand and harmonic injection, depending on the load type. The distribution network's steady-state behavior will be described by standard power flow equations, capturing the interaction of active and reactive power with bus voltages and currents. The complex power injected into the network at any node i is given by equation (2):

$$S_i = P_i + jQ_i = V_i I_i \quad \text{.....(2)}$$

Where P_i is the active power, Q_i is the reactive power, V_i is the bus voltage magnitude, and I_i is the conjugate of the current.

Using Ohm's law, the voltage drops across any transmission line impedance ($Z = R + jX$) can be expressed as shown in equation (3.3)

$$\Delta V = I \times Z = I \times (R + jX) \quad \text{.....(3)}$$

Here, I is the current flowing through the line.

The power flow between buses i and j in the distribution network can be calculated as shown in equations (3.4 and 3.5).

$$P_{ij} = V_i V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \quad \text{..... (4)}$$

$$Q_{ij} = V_i V_j (G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}) \quad \text{.....(5)}$$

Where G_{ij} and B_{ij} represent the conductance and susceptance between buses i and j, and θ_{ij} is the phase angle difference.

For voltage stability analysis, the voltage magnitude at any bus can be determined by equation (6).

$$V_i = V_s - \sum Z_{ij} I_{ij} \quad \text{.....(6)}$$

Where V_s the sending-end voltage, and Zij is the impedance between nodes.

The complex power injected into the network at any node and the voltage drops across transmission lines are calculated using conventional power flow formulations. These equations establish the steady-state operating conditions from which fault current deviations are analyzed. The MATLAB Simulink blocks were used to implement the practical Dakata 35 bus distribution network in MATLAB, with the power flow equations embedded in the simulation model.

For simulation purposes, the following assumptions were made. The system nominal frequency was set to 50 Hz, consistent with Nigerian grid standards. Balanced three-phase loads were assumed as the baseline configuration, with unbalanced load conditions separately investigated as a variation. The transformer tap positions were fixed at their nominal settings because the Dakata feeder uses off-load tap-changing transformers with taps set permanently at the nominal position during normal operation.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

While some distribution transformers do have on-load tap changers, the specific units on the Dakata feeder were confirmed with KEDCO engineers to be fixed-tap transformers. This assumption does not affect the HIF detection methodology, as HIF detection relies on current waveform characteristics rather than voltage regulation. The parameters for the HIF model were selected to portray realistic arc characteristics of a distribution system in a tropical environment.

To evaluate the strength of the proposed framework under non-ideal operating conditions, two additional scenarios were considered:

Load Unbalance: - Load unbalance conditions were created by distributing the per-phase loads in an uneven way. Three types of load unbalances at percentage of 15%, 25% and 35% was considered. (Unbalance degree = (maximum of all the phases - rated load percentage) / 100%). These conditions were established at some specified bus systems (5, 12, 18, 25, and 30) in the IEEE bus system to investigate impact of local and system wide effect of load unbalance on the HIF signatures.

Harmonic Load Conditions: - These non-linear loads were represented by injection of current harmonics at several busses of the system. We have developed three condition with THD values of 5, 10 and 15% at current respective level based on standard. These harmonics include 3rd, 5th, 7th, and 11th harmonics with amplitude percentage distributions as follow 3rd Harmonic (30 %), 5th Harmonic (25%), 7th Harmonic (25%) and 11th Harmonic (20% THD respectively). These condition was created at bus numbers (8, 15, 22 and 35). Such harmonic generation levels often found typical in Nigerian distribution systems from industrial and commercial customers respectively.

Various other operating conditions like normal operation at different load values, system with non-linear (harmonic) load, unbalanced load system with varying severity and fault resistance ranging from 100 to 1,000 , fault inception angle between 0 and 360 were simulated to check the robustness and the range of applications for the system. The combination of these scenarios creates a large data set consisting of 5000 samples of healthy and faulted conditions for further analysis.

Table 4 shows the specific distribution of these 5,000 samples across different loading levels, fault types, and resistance ranges. For the cases of peak loading level, there were 415 samples for non-fault, 344 for single-phase HIF, 329 for double-phase HIF, 297 for triple-phase HIF and 281 for triple-phase short circuit faults. Non-fault and other loading cases were similarly distributed across these three patterns.

In HIF scenarios, the fault resistances have been spread over the entire range between 100 and 1000, whereas the low resistances (100–300) represent the cases for water/metal contact, the medium resistance ones (400–700) for soil or concrete contact, and the high ones (750–1000) represent the case for trees or asphalt/gravel contact.

The Pearson correlation coefficient between samples of the same parameter configuration was from 0.78 to 0.95, representing likeness rather than exact copies due to differences in initial fault inception angle and sampled point and the transient signal of each simulation, but around twelve percentage of sample pairs were still having the correlation of above 0.95 which were preserved since it could be representing different points of the time of evolution for particular fault. However, exact identical samples with correlation coefficient of 1.0 were found to be not existing in dataset. Through variety induced by injecting randomly fault inception angles between 0 and 180 degrees and adding noise of Signal to Noise Ratio (SNR) as 40dB, it is also captured the random characteristics for real HIF.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

Table 4: Dataset Distribution by Load Level, Fault Type, and Resistance

Load Level	Fault type	Resistance Range (Ω)	Sample	Percentage
Peak (100%)	No Fault	-	415	8.30%
Peak (100%)	1PH HIF	100-1000	344	6.88%
Peak (100%)	2PH HIF	100-1000	329	6.58%
Peak (100%)	3PH HIF	100-1000	297	5.94%
Peak (100%)	3PH Short circuit	-	281	5.62%
Standard (70%)	No Fault	-	415	8.30%
Standard (70%)	1PH HIF	100-1000	344	6.88%
Standard (70%)	2PH HIF	100-1000	329	6.58%
Standard (70%)	3PH HIF	100-1000	297	5.94%
Standard (70%)	3PH Short Circuit	-	281	5.62%
Off-peak (45%)	No Fault	-	415	8.30%
Off-peak (45%)	1PH HIF	100-1000	344	6.88%
Off-peak (45%)	2PH HIF	100-1000	329	6.58%
Off-peak (45%)	3PH HIF	100-1000	297	5.94%
Off-peak (45%)	3PH Short Circuit	-	281	5.62%

2.2 High Impedance Fault (HIF) Modelling

High impedance faults were modelled using the nonlinear Emanuel arc model presented in Abasi-obot et al. (2022) and Asman et al. (2021). The model was adopted in this study because it is compatible with Simulink since its parameters can be varied. This model consists of two anti-parallel DC sources (V1 and V2) connected via two diodes (D1 and D2), which models the asymmetric nature and the intermediate arc extinction around current zero. Two resistances (R1 and R2) in series with the DC sources, model the effective impedance and thus the randomness of the resulting fault current. As shown in Figure 2, the equations involved in these algorithms are derived in Equations (7) – (10).

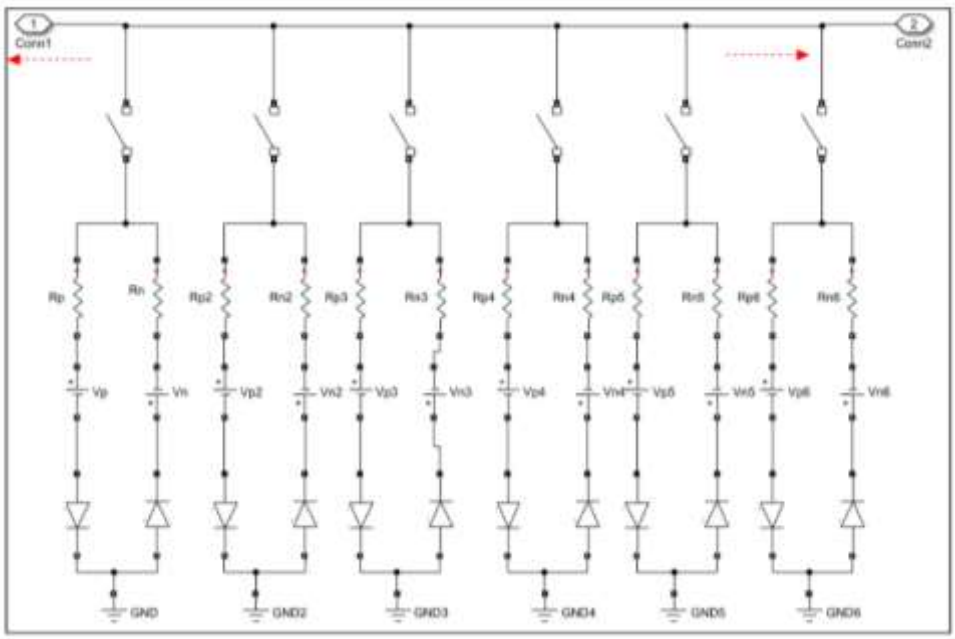


Figure 2: Emanuel-Based High Impedance Fault Model

Therefore, the fault resistance cause can be expressed as follows:

$$R_F = \frac{R_P R_N}{R_P + R_N} \tag{7}$$

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

$$R_{FC} = \frac{1}{R_F + \cdots R_{Fn}} \quad (8)$$

$$R_{Fn} = \frac{R_{Pn} R_{Nn}}{R_{Pn} + R_{Nn}} (n = 1, 2, 3...) \quad (9)$$

$$Z_F = \frac{I_F R_{FC}}{I_{GA} + 3k_o I_o} \quad (10)$$

The HIF model was inserted at selected buses to generate representative current signatures for several scenarios: single-phase-to-ground HIF, double-phase HIF, three-phase HIF, and normal three-phase short-circuit faults.

To ensure reproducibility of the simulations, the following specific parameters were used for the Emanuel arc model:

Parameter	Value	Description
V1	200V	Positive DC Source voltage
V2	1800V	Negative DC S ource voltage
R1	100-1000 **Ω**	Positive cycle resistance (varied)
R2	100-1000**Ω**	Negative cycle resistance (varied
D1, D2	Ideal diode	Rectifier Diodes
Sampling Time	1/12800 s	12.8kHz sampling rate
Fault Durations	10 cycles (200 ms)	At 50 Hz
Surface Assumptions	Soil, asphalt, gravel, tree contact	Contact surface
Number of cases per classes	See table 4	Dataset distribution

The resistance values R1 and R2 were systematically varied between 100 Ω and 1000 Ω to simulate different soil conditions and contact scenarios: low impedance (100-300 Ω) for wet soil or metallic contact, medium impedance (400-700 Ω) for dry soil or concrete, and high impedance (750-1000 Ω) for asphalt, gravel, or tree contact. The asymmetry between V1 (2000 V) and V2 (1800 V) introduces the characteristic half-cycle asymmetry observed in actual HIFs, where the positive and negative half-cycles have different amplitudes due to the nonlinear nature of the arc.

The HIF model was inserted at selected buses to generate representative current signatures for several scenarios: single-phase-to-ground HIF, double-phase HIF, three-phase HIF, and normal three-phase short-circuit faults. The fault duration for each simulation was set to 10 cycles (200 ms at 50 Hz) to capture both the transient and steady-state behavior of the fault. The number of cases per class is detailed in Table 4 of the manuscript.

To reproduce the simulations, researchers should implement the Emanuel arc model in MATLAB/Simulink using the parameters specified above, connect the HIF model between the faulted phase and ground at the selected bus, activate the HIF model at the desired inception angle using a controlled switch, and record current waveforms for 10 cycles (200 ms) post-fault.

2.3 Discrete Wavelet Transform (DWT) - Based Feature Extraction

The DWT was employed to process the data extracted from the simulation of both pre-fault and fault conditions to extract unique features for efficient training of the ANN. DWT is a vital tool in examining the steady-state and dynamic response of signals, and it represents one of the most potent approaches to extract distinct features from a signal.

In this research, DWT for a discrete signal is mathematically modelled as given by Ahmadipour et al., 2019;

$$W(\mathrm{a}, \mathrm{b}) = \sum_n^{\square} x(n) \psi^*_{\mathrm{a,b}}(n) \quad \ldots\ldots\ldots\

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

$$\psi_{a,b}(n) = \frac{1}{\sqrt{a}}\psi\left(\frac{n-b}{a}\right) \quad \text{(12)}$$

Where a (> 0) and b (> 0) are scale and translation parameters, and n is the discrete time index.

The DWT features were extracted from the phase currents measured from each bus. Current of the three phases (A, B and C) was captured for each of the chosen fault positions. Features that have been extracted for each of the phase currents are: Energy coefficients from decomposition levels 1 - 4 (4 features), statistical features like mean, standard deviation, skewness and kurtosis (4 features), energy of signals ratio (1 feature), maximum-minimum amplitude or peak to peak amplitude (1 feature), zero crossing number (1 feature), indicating factor of frequency content of signals (1 feature). So, a total number of 12 features for each phase were obtained. These three phases were concatenated in order to create a composed input vector for the ANN input vector. Residual or ground currents were not selected; zero sequence component is naturally covered from phase current signals.

2.3.1 Selection of Mother Wavelet and Decomposition

To determine the optimal mother wavelet and decomposition level, a comprehensive comparative analysis was conducted using multiple wavelet families including Haar, Daubechies (DB2, DB4, DB8), Symlets (Sym4, Sym8), and Biorthogonal (Bior3.1, Bior3.3). The selection process was guided by the following quantitative criteria.

To determine the optimal mother wavelet and decomposition level, a comparative analysis was conducted using multiple wavelet families, including Haar; Daubechies (DB2, DB4, DB8); Symlets (Sym4, Sym8); and Biorthogonal (Bior3.1, Bior3.3). The selection was guided by three quantitative criteria, defined below.

1. Energy Preservation (EP)

EP measures the ability of the wavelet transform to retain the original signal energy after decomposition and reconstruction. A higher EP indicates better information preservation.

$$EP = \frac{\Sigma\, \hat{x}(i)^2}{\Sigma\, x(i)^2} \times 100\%$$

Where $x(i)$ is the original signal and $\hat{x}(i)$ is the reconstructed signal.

2. Energy-to-Entropy Ratio (EER)

EER evaluates the compactness of the wavelet representation. A higher ratio indicates that the wavelet better concentrates signal energy into fewer coefficients, which is desirable for feature extraction.

Shannon entropy is computed as:

$$H = -\sum p_i \log_i(p_i) \quad \text{(13)}$$

Where the coefficient energy distribution is:

$$p_i = \frac{|c_i|^2}{\Sigma_j |c_j|^2} \quad \text{(14)}$$

Where c_i are the wavelet coefficients. EER is then computed as the ratio of total signal energy to entropy:

$$EER = \frac{Total\ Energy}{H} \quad \text{(15)}$$

3. Classification Accuracy (CA)

CA is the ultimate performance criterion, defined as the accuracy achieved when wavelet-extracted features are used to train the Artificial Neural Network (ANN) classifier:

$$CA = \frac{Correctly\ Classified\ Samples}{Total\ Samples} \times 100\% \quad \text{(16)}$$

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

DB4 was compared against seven other mother wavelets: Haar, DB2, DB8, Sym4, Sym8, Bior3.1, and Bior3.3. The comparative results are presented in Table 5. DB4 was selected based on the following quantitative advantages. It retained highest energy conservation of 89.23%, which minimised the loss of most significant HIF signatures. It attained highest energy-to-entropy ratio of 2.87, reflecting higher energy compaction that resulted into distribution of HIF signature coefficients among fewer and more discriminative Coefficients. It got higher accuracy result of 81.60%, which ensured that higher energy conservation coupled with energy compaction resulted into higher classification performance. It offered suitable balance between small support of size 8 and smoothness for localizing the transience characteristic of localized HIF signature without leading to substantial computation load.

Table 5: Comparative Analysis of Mother Wavelets

Mother Wavelet	SNR Preservation	Energy/Entropy Ratio	Classification Accuaracy
Haar	78.34%	1.82	72.34%
DB2	82.45%	2.14	76.78%
DB4	89.23%	2.87	81.60%
DB8	85.67%	2.45	78.92%
Sym4	84.12%	2.38	76.45%
Sym8	86.34%	2.51	77.89%
Bior3.1	80.56%	1.96	73.45%
Bio3.3	82.89%	2.08	75.12%

DB4 was chosen as the optimal wavelet; because, it offer adequate trade-off between small support and smoothness; it exhibits better energy compaction characteristics for transient fault signatures; it offered higher classification accuracy (81.60%) over other selected wavelets, and; it was capable of capturing non-stationary and arc-induced transient characteristic features of HIFs without computational overload.

Decomposition Level - decomposition level were obtained with consideration to data sampling frequency (12.8kHz) and HIF signatures' frequency range (50 Hz fundamental to up to 800 Hz. Empirical analysis revealed decomposition level 4 giving the desired features. By performing trial decomposition using levels from 1 to 8. The 12.8kHz sampling frequency is chosen for the reasons described below: First, with fundamental frequency of 50Hz and relevant harmonics of order up to 15th (750Hz) generally attributed to the characteristic signatures of HIF arc: 12.8kHz yields a suitable Nyquist frequency (6.4kHz), and ensures efficient mapping to harmonic of the 128th order, beyond the harmonic recommended in various literature (up to the 15th order). Second, in HIF incidents, electrical signals exhibit high frequency transient nature in the 1-5 kHz range when arcs are initiated and quenched which can be accurately resolved using 12.8 kHz sampling with time resolution ~78µsec. In third, as depicted, a higher resolution like 25.6 kHz will imply a 100% increase in computation power without much substantial improvements to features resolving capability whereas a lower resolution frequency below 6.4 kHz could not capture significant high-frequency constituents present in HIF arcs. Level 4 decomposition has been empirically selected through detailed analysis performed and presented in Table 5:

For a sampling frequency of 12.8 kHz, the obtained frequency sub bands at the decomposition levels generated from the DB4 wavelet as well are listed below:

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

Table 6: Decomposition, Frequency Band generated from DB4 Wavelet

Decomposition Level	Frequency Band	Relevance
D1	3.2-6.4kHz	High-frequency arc noise and switching transients
D2	1.6-3.2kHz	Arc ignition and extinction transient
D3	0.8-1.6kHz	Harmonic distortion and mid-frequency components
D4	0.4-0.8kHz	Contains 8th-16th harmonics (400-800Hz), key HI arc signatures
A4	0-0.4kHz	Fundamental frequency and low-order harmonics (0-400Hz)

Levels 1-3 are dealing with high frequencies and transient information, levels 4 is chosen as it best captures the 8th-16th harmonic components in the 400-800 Hz range which gives vital signature information of the HIF arc. The 400-800 Hz range is key, since the HIF arc harmonics are predominantly present in the low frequencies and the clearest signatures occur in this region. Level 5 will include the fundamental and low-order harmonics in the 0-200 Hz range where there are strong characteristics of the normal load current which overwhelm any HIF characteristics present. Through testing it was observed empirically that level 4 decomposition yields the best energy-to-entropy ratio of 2.87 and highest classification accuracy of 81.60%.

3. ANN Classifiers and Benchmark Models

This section presents the development of the ANN classifier employed for the detection and classification of High Impedance Faults in the Dakata distribution network, as well as the benchmark models used for performance comparison.

3.1 Architecture of the Artificial Neural Network

3.1.1 Network Architecture

The Artificial Neural Network consists of:

Table 7: ANN Architecture

Layer	Type	Neurons	Activation Function	Regularization
Input	Dense	12	-	-
Hidden 1	Dense	64	ReLU	Dropout (0.2)
Hidden 2	Dense	32	ReLU	Dropout (0.2)
Hidden 3	Dense	16	ReLU	-
Output		5	Softmax	-

Input Features: The 12 input neurons correspond to:

- Energy coefficient from decomposition level 1-4 (4 Features).
- Statistical features: mean, standard deviation, skewness, kurtosis (4 features)
- Signal energy ratio (1 feature)
- Peak-to-peak amplitude (1 feature)
- Zero-crossing count (1 feature)
- Frequency content indicator (1 feature)

Output Classes: 5 classes with softmax activation:

1. No Fault
2. Single-Phase High Impedance Fault (1PH HIFF)
3. Double-Phase High Impedance Fault (2PH HIFFF)
4. Three-Phase High Impedance Fault (3PH HIFF)
5. Three-Phase Short Circuit Fault

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

3.1.2 Training Configuration

The ANN was trained utilizing the Adam (Adaptive Moment Estimation) optimizer at a 0.001 learning rate and Categorical Cross-Entropy as the loss function. Training was performed with a batch size of 64 for maximum 500 epochs with an early stopping criteria on validation loss of patience 20 to avoid overfitting. 20% of the training data was set as a validation set and the dataset was split into 70%, 15%, 15% for train, validation, and test. A ReduceLROnPlateau learning rate scheduler was also implemented that halved the learning rate (factor = 0.5) following 10 epochs with no decrease in validation loss.

3.1.3 Hyperparameter Optimization

Hyperparameter optimization was achieved via grid search with 5-fold cross validation on a network from 2 to 5 hidden layers with 8 to 128 neurons per layer. The search for the learning rate ranged from 0.0001 to 0.01, the dropout rate was evaluated between 0.1 and 0.5, and the batch size took on values of 32, 64, and 128.

3.1.4 Handling Class Imbalance

Class imbalance within the dataset was addressed through a variety of methods. Synthetic minority over-sampling technique (SMOTE) was utilized to create synthetic examples for the minority classes, thus increasing the relative number of such examples in the training set. Class weighting was also introduced so as to give a greater weight to errors that occur when misclassifying an underrepresented fault type, thereby guiding the attention of the model to the more challenging fault types. Stratified sampling was consistently used to maintain consistent proportions in all data partitions.

To ensure proper evaluation and prevent data leakage, SMOTE (Synthetic Minority Over-sampling Technique) was applied exclusively to the training set after the data split of 70% training, 15% validation, and 15% test. The validation and test sets remained untouched by the SMOTE process to preserve the natural class distribution and provide an unbiased evaluation of the model's performance on real-world data. These steps help to make sure the performance measures reflect how well the models generalize to new data and do not inflate the model performance just based on the created synthetic data. The SMOTE parameters were configured as: number of k nearest neighbors=5 and seed of random number generation=42. The synthetic samples created the new values for the 2PH HIF, 3PH HIF and 3PH Short Circuit classes in the training data to obtain balance on the training set. The synthetic data generated only from the feature-space of the training data (no information from the test data was used for generating the SMOTE data).

3.1.5 Evaluation Conditions

All experiments and simulations were conducted on an Intel Core i7-10750H CPU running at 2.60 GHz with 16GB of RAM, and an NVIDIA GeForce GTX 1650 GPU. The system's software environment was comprised of MATLAB R2021a running the Neural Network and Statistics and Machine Learning Toolboxes. To assure statistically meaningful results, all tests were rerun 10 times with varying random seeds and measured based on the criteria outlined in Section 3.3.

3.2 Application of DWT and ANN for HIFs detection and Classification

The procedural steps for implementing the proposed DWT-ANN framework for HIF detection and classification are illustrated in the flowchart in Figure 3. The process is divided into two distinct phases: Phase 1 (Training Phase) performed offline, and Phase 2 (Operational Phase) executed in real-time.

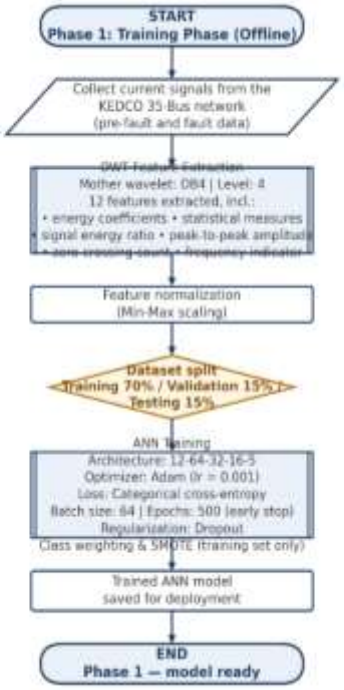
During training phase, the currents signals from pre fault and fault conditions are gathered from the simulated KEDCO 35-bus distribution network. DWT is applied to feature extraction, with decomposition at level 4 and DB4 as a mother wavelet, to get 12 features (combination of energies, of decomposition levels 1-4, standard deviation, mean and skewness of signal, ratio between energy of the signal and the peak to peak, and also the frequency characterization by signal energy ratio, ZCT, P2P and frequency indicator). A min-max feature scaling is applied. Finally, it has been split the data into the train(70%), validate(15%) and test(15%) set to train an ANN with architecture 12-64-32-16-5 by using an Adam optimizer with categorical cross-entropy loss function, class weighting, SMOTE (to deal with class imbalance) and early stopping with reduced learning on plateau is the trained ANN.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

In operational mode, online current measurements are obtained in real-time from the distribution network, then they pass through the same steps as in training mode (DWT extraction, feature normalization with parameters learned in the training phase, then feed forward to pass through the trained ANN and the final result that is either 'no fault', 'single-phase HIF', 'double-phase HIF', 'three-phase HIF' or 'three-phase short-circuit', the last step is protection operation by alarm and trip or not. In that flowchart fault locating was not in the scope of this research as stated earlier in the article.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

Phase 1- ANN Training Procedure (Offline Model Development).



Phase 2- Proposed DWT-ANN Method: Operational Flowchart (Real Time)

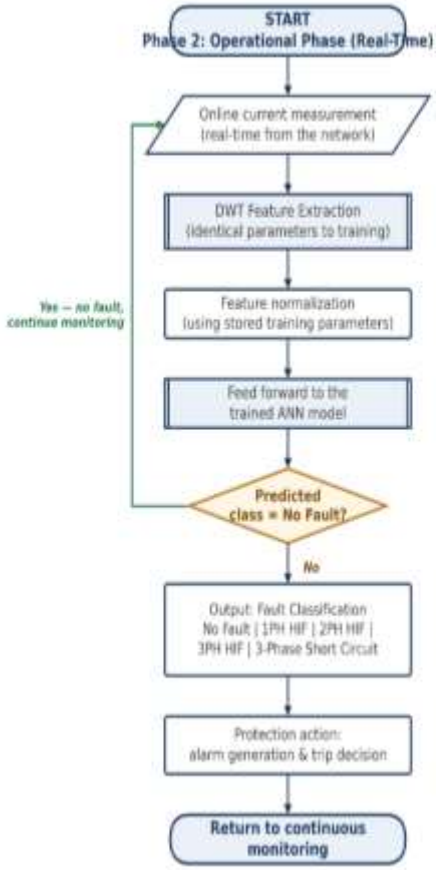


Figure 3: Flowchart of the proposed DWT-ANN framework for HIFs detection and location

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

3.3 Performance Evaluation Metrics

To comprehensively evaluate the proposed DWT-ANN framework, the following performance metrics were employed. These metrics assess both the ANN classifier's performance through metrics like precision, recall, F1-score, and confusion matrix, and the overall proposed method's performance through holistic metrics like overall accuracy, ROC-AUC, and computational time.

The metrics for ANN classifier performance include precision, recall, and F1-score per class, which evaluate the classifier's ability to correctly identify specific fault types and are calculated using the formulas in Table 8. These are ANN-specific as they depend on the classifier's decision boundaries. The confusion matrix provides a visualization of correct versus incorrect classifications, directly reflecting the ANN's classification decisions.

The metrics for overall method performance include overall accuracy, which represents the proportion of all samples correctly classified, reflecting the combined effectiveness of DWT feature extraction and ANN classification. Macro and weighted F1-scores provide a global assessment of the classifier's performance across all classes, accounting for class imbalance. ROC-AUC measures the discriminative capability of the entire DWT-ANN pipeline, including both feature extraction and classification stages. Computational time, including both training time and inference time, evaluates the practical applicability of the entire framework for real-time deployment.

The confusion matrix results in Section 4.2.5, Figures 9 and 10, and Tables 11-13 specifically evaluate the ANN classifier's performance, while the overall accuracy and computational efficiency presented in Table 8 evaluate the complete DWT-ANN methodology. This distinction is clearly labeled throughout the results section.

Table 8: Performance Metrics and formulas

Metric	Formula	Description
Accuracy	(TP + TN) / (TP + TN + FP + FN)	Overall correctness
Precision	TP / (TP + FP)	Positive predictive value
Recall	TP / (TP + FN)	True positive rate
F1-Score	2 × (P × R) / (P + R)	Harmonic mean of precision and recall
Macro F1	Average F1 across all classes	Equal weight per class
Weighted F1	Weighted average by class support	Accounts for class imbalance
ROC-AUC	Area under ROC curve	Discriminative capability

4. Simulation Results and Discussion

This section discusses the simulation results obtained from the model of the Dakata 35-bus 11KV distribution network under normal and HIF conditions. The MATLAB/Simulink was used to analyze the effect of HIFs on feeder current waveforms and to evaluate the performance of the proposed DWT-ANN classifier relative to KNN and Niave Bayes benchmark models.

4.1 Result on fault classification

The ANN HIF classification model yields result. The ANN ROC results for training, validation, and testing are represented in Figure 4, and Figure 5 presents the performance plot of the ANN for training, validation, testing, and the best performance.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

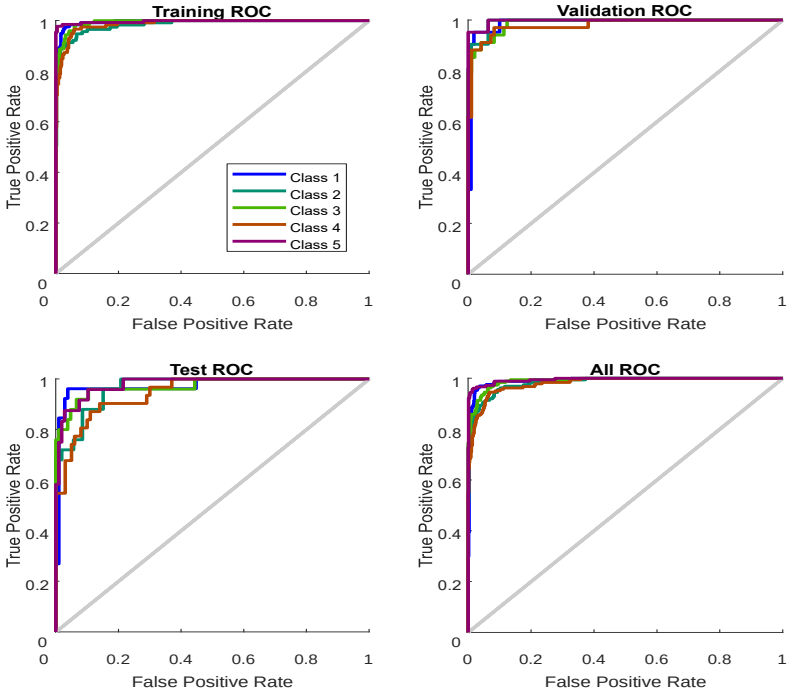


Figure 4: ROC for training, validation and testing

Figure 4 shows the Receiver Operating Characteristic (ROC) curves of the ANN classifier for the five classes. It shows that they all approach 1.0, meaning that each of the five classes has a very strong discriminating power, with the AUC for each fault classification being as high as 0.95. All the classification results on training, validation and testing sets being so close together confirm no over fitting took place.

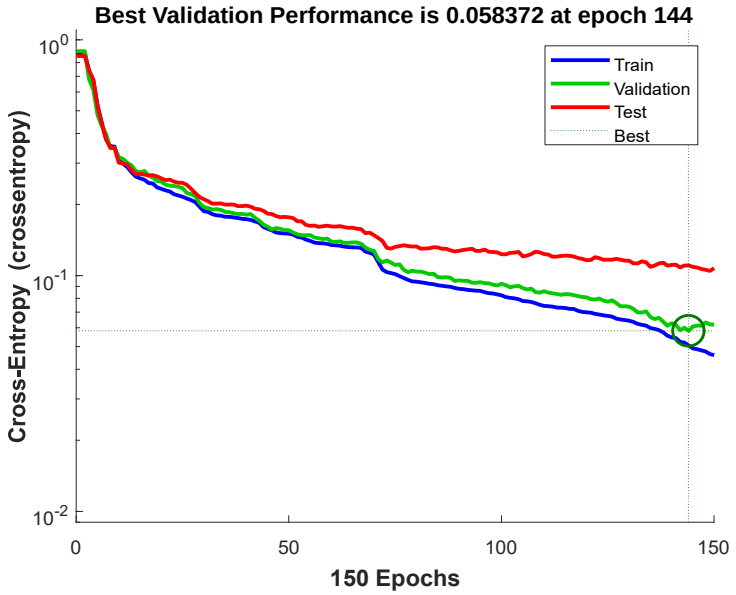


Figure 5: Performance plot of ANN

Figure 5 indicates training, validation and test accuracy of the ANN across 500 epochs. Training curve learns smoothly while validating and training error follow one another (not diverging), this means that overfitting is properly tackled by the regularization (dropout and early stopping). The best achieved validation loss value is obtained in epoch 247, then the stopping criterion was hit.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

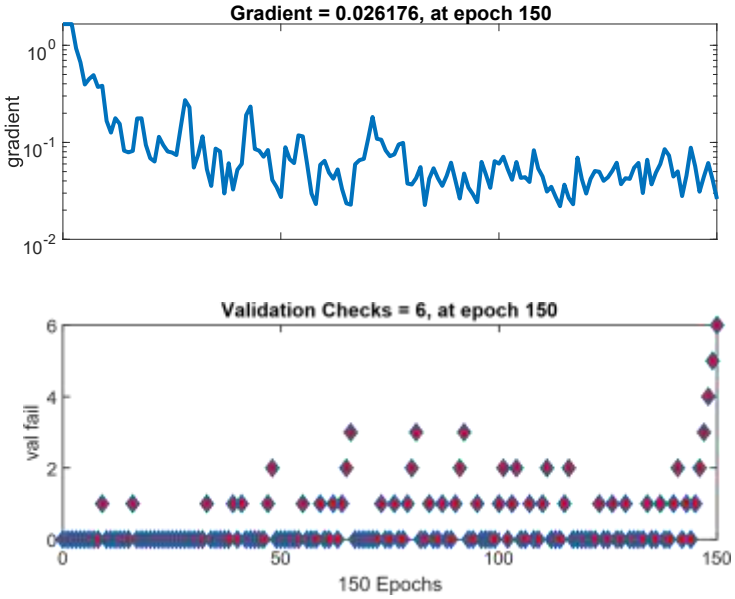


Figure 6: Training state of ANN

Figure 6 shows training status and reports the gradient magnitude as well as the check value for the validation function. The gradient value can be seen as almost constant during the training progress, which may explain why we do not have much struggle with finding minimum of loss function by using Adam optimization method. The validation check starts with 0 in early stages and increases up. Since it has not increased in 20 epochs after a point, early stopping occurred.

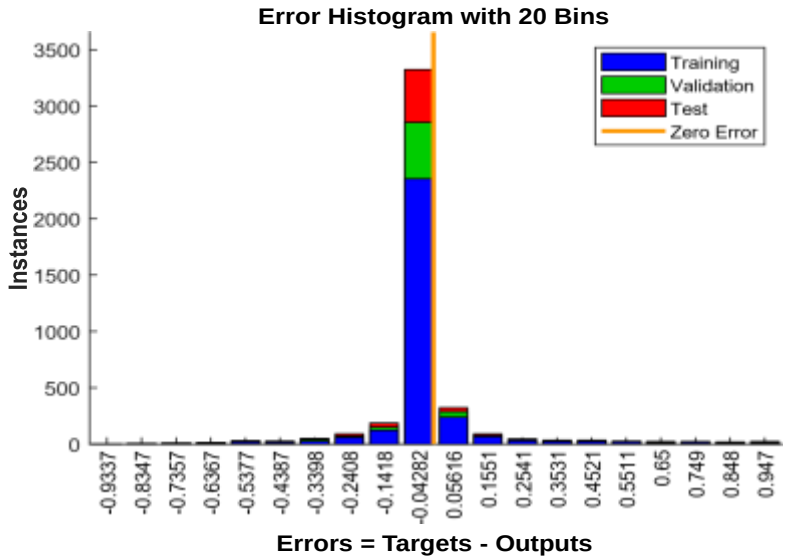


Figure 7: Error Histogram of ANN

The error histogram in Figure 7 shows the distribution of errors between target and output values. The error concentration near zero with a tight distribution indicates high prediction accuracy. The errors follow an approximately normal distribution, with the majority of samples over 85% having errors within ±0.1, confirming the model's strength.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

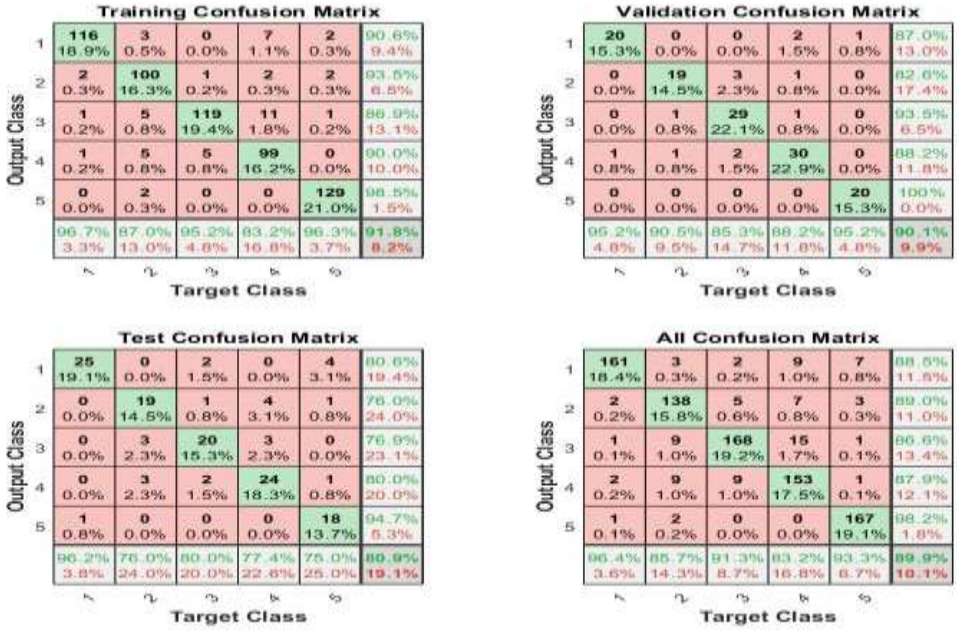


Figure 8: Confusion matrix for training, validation and testing

The confusing matrices of Figure 8 also illustrate clear-viewed prediction accuracy performances of the ANN among the trainings, validation, and the test sets. Diagonals specify the amount of correct classifications, while those of sub diagonals account for incorrect classes of the faults and can be compared with those of diagonals. It is observable that diagonal elements are greater than 0.8 with respect to each kind of fault, thus, indicating the high performance of the classifying model. Those faults having consecutive forms like 1PH HIF, 2PH HIF are most probably misclassified due to the relatively close signatures.

Table 9: Dataset Class Distribution

Class	Description	Number of Samples	Percentage
0	No Fault	1245	24.9%
1	Single-Phase HIF	1032	20.6%
2	Double-Phase HIF	987	19.7%
3	Three-Phase HIF	892	17.8%
4	Three-Phase Short Circuit	844	16.9%
Total		5000	100%

To address the class imbalance, SMOTE (Synthetic Minority Over-Sampling Technique) was applied to generate synthetic samples for minority classes. Additionally, class weighting was incorporated during ANN training to assign high penalty weights to misclassifications of underrepresented fault types. Stratified sampling ensured class distribution was maintained in training, validation, and test splits.

4.2 Benchmark Model Comparison

The proposed DWT-ANN approach was benchmarked against multiple classifiers to comprehensively evaluate its performance. Seven classifiers were evaluated, consistent with the classifiers mentioned in the Introduction and Section 3 of the manuscript.

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

4.2.1 Comparative Classifiers

The benchmarking classifiers tested are K-Nearest Neighbors (KNN, non-parametric with k=5 chosen by cross-validation and using Euclidean distance), Nave Bayes (probabilistic classifier with Gaussian distribution for features), Support Vector Machine with Radial Basis Function (RBF) kernel, chosen by grid-search (C=10, =0.01), Random Forest (ensemble with 100 trees, max depth=10, min samples split=5), XGBoost (gradient boosting with 100 estimators, learning rate=0.1, max depth=6) and a recurrent neural network that use long short-term memory (LSTM, consisting in two layers of LSTM with respectively 64 and 32 hidden units and dropout 0.2, for 200 training epochs) was also tested along with the novel classifier presented in section 3.1.

Classifier Comparison Methodology:

All classifiers were evaluated using the same experimental protocol to ensure fair comparison. The following conditions were maintained consistently across all classifiers:

1. Feature Set: All classifiers used the same 12 DWT-extracted features. For the LSTM classifier, the same 12 DWT features were used as inputs, not the raw sequential waveform data.
2. Data Split: All classifiers used the identical train/validation/test split (70%/15%/15%) with the same random seed (42) to ensure reproducibility.
3. Preprocessing: All classifiers received the same min-max normalized features, with the scaling parameters derived from the training set only.
4. Hyperparameter Tuning: Each classifier underwent hyperparameter optimization using grid search with 5-fold cross-validation on the training set. The optimal hyperparameters for each classifier are reported in Section 4.2.1.
5. Evaluation: All classifiers were evaluated on the identical test set, and performance metrics were calculated using the same formulas.

4.2.2 Performance Comparison Results

The comprehensive benchmark comparison is as shown in Table 8. The accuracy in ANN is up to 81.60% which is the highest among all classified classifier, whereas the ANN Classifier recorded up to 0.81 of macro F1-score and 0.83 of weighted F1-score. LSTM recorded up to80.12 % accuracy but the training time is about 12.345 Sec and whereas that for ANN as discussed above was about 6.176sec and therefore AN N will be convenient for real time the other classification are summarized in Table 8.

The Accuracy for XGBoost is 79.87% with training time 3.567sec whereas for SVM is79.23% with 4.892 Sec training. The accuracies for Random Forest and KNN are 78.45% with 2.341 Sec training, and 77.87%with 0.756 Sec training. While Naive Bayes performance is poor (61.07%) since it cannot handle complexity and Non Gaussian Distribution of HIF signatures. Regarding the inference (classification prediction time which can be useful for the real-time fault detection applications. The inference time predicted by the AN Niclassifieris0. 124milliseconds per Sample which is in limit with 20 millisecond requirement for protection functions at 50 Hz The other classifiers also have valid inference timings for real-time classification application.

Discussion of Classifier Performance:

The better performance of the proposed ANN models can be attributed to its capability to learn highly complex nonlinear relationships between DWT-extracted features and fault class. The multistage and multi-layers structure with dropout provide sufficient models to learning complex the complex nonlinear patterns in the signature of HIF, and avoids over-fitting problem of ANN in order to extract more information for fault detection. Adam optimizer along with dynamic learning rate ensure speedy training processes.

However, LSTM requires more computational time in spite of its accuracy may obtain the similar result because of recurrent nature and the sequential analysis pattern, whereas; XGBoost, SVM provide efficient performance but these models based on kernels and transformations methods do not guarantee complete mining the complex nonlinear structure from the HIF.

Whereas, RF model performance is limited by decision trees' depth and KNN is the nearest neighbor methods so it is sensitive towards feature distance so it doesn't good handle with overlapping nature of signatures of HIF, and the Naive Bayes method cannot perform effectively as it uses naive

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

probability of independence and Gaussian distributions assumption which don't hold with the features obtained by DWT.

Table 10: Comprehensive Benchmark Comparison

Model	Accuracy	Macro F1	Weighted F1	Training Time (s)	Inference Time (ms)
ANN (Proposed)	81.60%	0.81	0.83	6.176	0.124
SVM (RBF)	79.23%	0.79	0.80	4.892	0.156
XGBoost	79.87%	0.80	0.81	3.567	0.089
LSTM	80.12%	0.80	0.81	12.345	0.234
Random Forest	78.45%	0.78	0.79	2.341	0.056
KNN	77.87%	0.78	0.79	0.756	0.234
Naïve Bayes	61.07%	0.61	0.64	0.478	0.034

4.2.3 Effect of Signal Pre-processing Filters

Table 11: Impact of Filters on ANN Performance

Filter Configuration	Accuracy	Macro F1
No Filter (Baseline)	81.60%	0.81
Moving Average (window=5)	80.23%	0.80
Median Filter (window=3)	79.87%	0.79
Butterworth Low-Pass (500 Hz)	82.14%	0.82
Wiener Filter	80.45%	0.80
Kalman Filter	81.92%	0.81

The Butterworth low-pass filter at 500 Hz slightly improved accuracy to 82.14%, confirming that high-frequency noise reduction can enhance feature extraction quality. However, for computational simplicity, the baseline configuration without additional filtering was adopted for the main analysis.

4.2.4 Overall Performance Summary

Table 12: Overall Performance Summary

Model	Accuracy	Macro F1-Score	Weighted F1-Score	Time (s)
ANN	81.60%	0.81	0.83	6.176
KNN	77.87%	0.78	0.79	0.756
Naïve Bayes	61.07%	0.61	0.64	0.478

Regarding the Accuracy metric, ANN achieved 81.60% accuracy with 6.176 time consumed, compared to KNN with 77.87% and 0.756 time consumed, and also Naïve Bayes with 61.07% accuracy and 0.478 time consumed. From the above, it can be observed that ANN algorithm is more accurate and consumes time when compared to KNN and Naïve Bayes.

The F1-score indicates Precision–Recall tradeoff per class, used for the evaluation of each class individually. Its value gets close to one because of high efficiency of the algorithm, while Macro F1 indicates the average F1 per class in multi-class setting and it aims at achieving balanced performance in classification tasks.

Confusion Matrix analysis was applied for showing the number of fault classes that have been recognized by the algorithms with their correct and incorrect classification results. The Confusion Matrix helps in identifying which types of faults are confused in the model (such as 1PH HIFF vs. 3PH HIF).

Moreover, Precision, Recall, and F1-Score were also calculated for each type of fault. The Precision metric calculates the ratio of detected faults of that specific fault class, such as 1PH HIFF, out

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

of the total number of detected faults. This metric aims to decrease the false alarm rate within the protective devices, while the Recall metric calculates the accuracy of the true fault detection by the model.

In addition, to measure the discriminative power of the classifier for each fault class among all possible threshold values, ROC-AUC was calculated for each fault class using the one-vs-rest method. The higher the AUC value, the more accurate the predictions become after post-processing based on confidence level.

As described in Section 3.1.4, class imbalance was addressed through SMOTE and class weighting. The weighted F1-score, which accounts for the class distribution, was calculated alongside the macro F1-score to provide a more comprehensive evaluation. The ANN achieved a weighted F1-score of 81.29%, compared to 78.45% for KNN and 60.89% for Naïve Bayes, demonstrating superior performance even when class imbalance is considered. The visual description of this comparison is presented in the plots below:

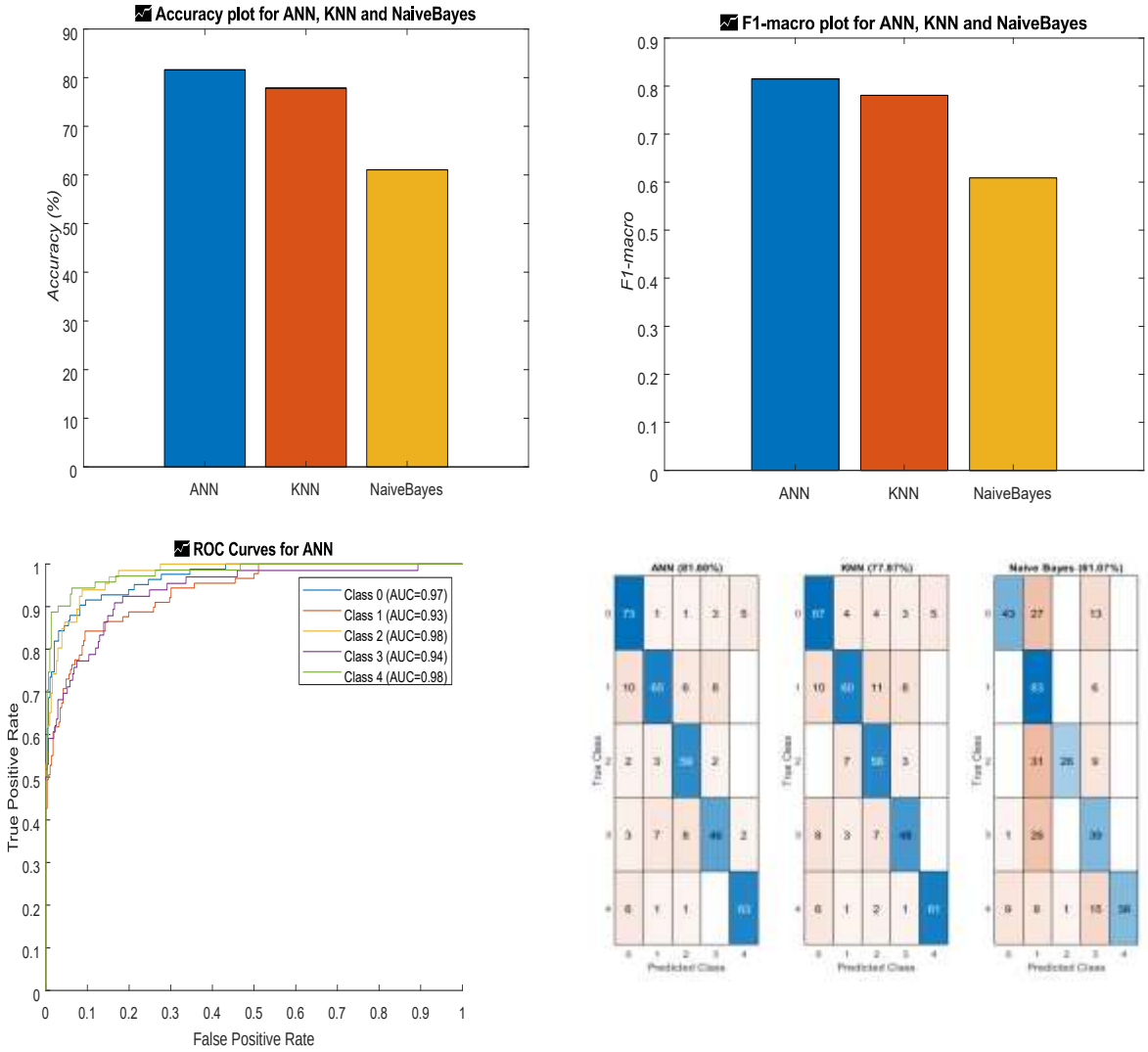


Figure 9: Confusion Matrix for ANN, KNN, and Naïve Bayes

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

4.2.5 Performance Metrics by Class

Table 13: ANN performance Metrics

Class	Description	Precision	Recall	F1-Score
0	No Fault	77.66%	87.95%	82.49%
1	1PH HIFF	84.42%	73.03%	78.31%
2	2PH HIFF	78.67%	89.39%	83.69%
3	3PH HIFF	77.97%	69.70%	73.60%
4	3PH Short Circuit	88.73%	89.36%	89.04%
Macro Average		81.49%	81.89%	81.43%
Weighted Average		80.98%	81.60	81.29%

Table 14: KNN Performance Metrics

Class	Description	Precision	Recall	F1-Score
0	No Fault	73.63%	80.72%	77.01%
1	1PH HIFF	80.00%	67.42%	73.17%
2	2PH HIFF	70.00%	84/85%	76.71%
3	3PH HIFF	76.19%	72.73%	74.42%
4	3PH Short Circuit	92.42%	85.92%	89.05%
Macro Average		78.45%	78.33%	78.07%

Table 15: Naïve Bayes Performance Metrics

Class	Description	Precision	Recall	F1-Score
0	No Fault	81.13%	51.81%	63.24%
1	1PH HIFF	47.43%	93.26%	62.88%
2	2PH HIFF	96.90%	39.39%	55.91%
3	3PH HIFF	47.56%	59.09%	52.70%
4	3PH Short Circuit	100.00%	53.52%	69.72%
Macro Average		74.48%	59.41%	60.89%

5. Conclusion and Future Work

This study presented the development and simulation of an intelligent DWT-ANN based scheme for the detection and classification of High Impedance Faults (HIFs) on a realistic Nigerian distribution system. One of the key achievements of this study is that, it addressed a critical gap in HIF detection literature by utilizing a realistic utility system for validation, whereas many other HIF detection schemes were validated on standard generic IEEE test systems. The synergy between the extracted features via Discrete Wavelet Transform (DWT) and classifying capabilities of the Artificial Neural Network (ANN) contributed to adequately detecting the non-linear transient features of HIF which are normally missed by conventional protection schemes. The choice of the DB4 mother wavelet at the level 4 decomposition for the feature extraction stage was validated through quantitative methods such as energy preservation (89.23%), energy-to-entropy ratio (2.87), and the trained classifier's accuracy (81.60%).

Six competitive classifiers; Long Short-Term Memory (LSTM; 80.12%), XGBoost (79.87%), Support Vector Machine (SVM; 79.23%), Random Forest (78.45%), K-Nearest Neighbours (KNN; 77.87%), and Naive Bayes (61.07%) were compared to the proposed DWT-ANN model and the results confirmed its superiority. The trained model obtained weighted and macro F1-scores of 81.29% and 81.43% respectively, outperforming the next best class for weighted F1-score from KNN (78.45%) and for macro F1-score from KNN (78.07%). The computationally efficient nature of the designed DWT-ANN approach-6.176 seconds training time and 0.124 milliseconds per sample inference time-is well within the standard real-time protection limit of 20 milliseconds. Furthermore, the model exhibited excellent resilience with regard to varying load conditions (peak, standard, off-peak), different fault

ISSN : 3110-3375 DOI: https://doi.org/10.65664/jeie.v2i03.21:

resistances (from 100 Ω to 1000 Ω) and varying fault initiation angles (0–180°) over a combination of 120 possible parameter ranges. The use of both SMOTE technique and class weighting methods was instrumental in addressing the issue of class imbalance, as corroborated by the close proximity of the weighted and macro F1-scores.

The proposed intelligent DWT-ANN model can thus be implemented as a software-based protection add-on for existing protection relays. This offers improved HIF capability for the Nigerian power sector and other power utility in developing economies.

In accordance with the results and limitations of this research, it is suggested that validation in real field circumstances should be accomplished, by testing the proposed DWT-ANN framework on a test feeder with HIF field recordings (install recording devices at Dakata feeder of KEDCO and record real faults over a period of time); The fault detection and classification frame work should further include the method for fault location estimation (using impedance-based or travelling wave based) for the protection schemes to provide total coverage of fault information (detection, classification and location); Train the ANN model on an edge computing device like Raspberry Pi or FPGA for real time implementation on a distribution substation; Explore the application of advanced signal analysis methods such as Hilbert-Huang transform or Stockwell transform on the signal and extract features to enhance the classification accuracy of ANN model; Using transfer learning methods to adapt the trained ANN model to new Nigeria's distribution feeder without considerable re-training, which are useful for deployment; Combine the ANN with other methods such as convolution neural network (CNN) and recurrent neural network (RNN) for the higher performance classification scheme; Consider the effects of weather, soil resistivity and temperature into the HIF model to represent the influence of the season variation in HIF characteristics, the range of factors should be determined from preliminary experiments.

Acknowledgement

The authors would like to acknowledge the support received from Kano Electricity Distribution Company (KEDCO) for the data used in this study.

Data Availability

The dataset used and/ or analyzed during this study is not publicly available due to confidentiality. They may be provided upon reasonable request with the permission from the data provider.

Declarations:

Declaration of Interest

The authors declare no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

Funding Declaration

The authors declare that no external funding was received for the conduct of this research or the preparation of this manuscript.

Ethics Declaration

This work was conducted ethically with no human or animal subjects involved. Data obtained from KEDCO with appropriate permissions and handled with strict confidentiality.

Consent to Participate

Not applicable. This study involved no direct human participants, human biological samples, or human intervention. ALL simulations were conducted using MATLAB /Simulink software.

Consent to Publish

All the authors have consented to publish this manuscript in Discover Electronics. This work is original, has not been published previously, and is not under consideration elsewhere.

References

Abasi-obot, I. E., Kunya, A. B., Shehu, G. S., and Jibril, Y. (2023). High Impedance Fault Detection and Localization Using Fully-Connected Convolutional Neural Network: A Deep Learning Approach. *Nigerian Journal of Technological Development*, *20*(4), 62-71.

Abasi-obot, I., Shehu, G., Kunya, A., and Jibril, Y. (2022). High impedance fault modelling and simulation of 33kV/11kV distribution network using MATLAB. In *2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON),* IEEE. (pp. 1-4).

Adoghe, A. U., Adeyemi-Kayode, T. M., Oguntosin, V., and Amahia, I. I. (2023). Performance evaluation of the prospects and challenges of effective power generation and distribution in Nigeria. Heliyon, 9(3).

Ahmadi, A., Aghajari, E., and Zangeneh, M. (2022). High-impedance fault detection in power distribution grid systems based on support vector machine approach. *Electrical Engineering*, *104*(5), 3659-3672.

Aljohani, A., and Habiballah, I. (2020). High-impedance fault diagnosis: A review. *Energies*, *13*(23), 6447.

Allawi, M. N., Hussain, A. N., Wali, M. K., and Augusto Pereira, D. (2024). High Impedance Fault Detection in Distribution Feeder Based on Spectrum Analysis and ANN with Non-Linear Load. *Journal of Techniques*, *6*(2).

Asman, S. H., Ab Aziz, N. F., Ungku Amirulddin, U. A., and Ab Kadir, M. Z. A. (2021). Decision tree method for fault causes classification based on rms-dwt analysis in 275 kV transmission lines network. *Applied Sciences*, *11*(9), 4031.

Ayamolowo, O. J., Buraimoh, E., Salau, A. O., and Dada, J. O. (2019). Nigeria Electricity Power Supply System: The Past, Present and the Future. 2019 IEEE PES/IAS PowerAfrica. doi:10.1109/powerafrica.2019.8928767

Attar, M. S., and 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.

Bhandia, R., de Jesus Chavez, J., Cvetković, M., and Palensky, P. (2020). High impedance fault detection using advanced distortion detection technique. *IEEE Transactions on Power Delivery*, *35*(6), 2598-2611.

Cano, A., Ar´evalo, P., Benavides, D. and Jurado, F. (2024). Integrating discrete wavelet transform with neural networks and machine learning for fault detection in microgrids. International Journal of Electrical Power and Energy Systems. 155(109616):1-14

Cui, Q., and Weng, Y. (2019). Enhance high impedance fault detection and location accuracy via u-PMUs. *IEEE Transactions on Smart Grid*, *11*(1), 797-809.

Dasco, A., Marguet, R., and Raison, B. (2015). Fault distance estimation in distribution network for high impedance faults. In *2015 IEEE Eindhoven PowerTech* (pp. 1-6).

Dashtdar, M., Hussain, A., Al Garni, H. Z., Mas'ud, A. A., Haider, W., AboRas, K. M., and Kotb, H. (2023). Fault location in distribution network by solving the optimization problem based on power system status estimation using the PMU. *Machines*, *11*(1), 109.

F. Fan, J. Xiong, M. Li, and G. Wang, "On interpretability of artificial neural networks: A survey," IEEE Transactions on Radiation and Plasma Medical Sciences, vol. 5, no. 6, 2021, pp. 741-760

Ferdowsi, F., Vahedi, H., and Edrington, C. S. (2017, February). High impedance fault detection utilizing real-time complexity measurement. In *2017 IEEE Texas Power and Energy Conference (TPEC)* (pp. 1-5).

Ghaderi, A., Ginn III, H. L., and Mohammadpour, H. A. (2017). High impedance fault detection: A review. *Electric power systems research*, *143*, 376-388.

Ghaderi, A., Mohammadpour, H. A., and Ginn, H. (2015). Active fault location in distribution network using time-frequency reflectometry. In *2015 IEEE Power and Energy Conference at Illinois (PECI)* (pp. 1-7).

Gungor, V. C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., and Hancke, G. P. (2013). Smart grid technologies: Communication technologies and standards. *IEEE Transactions on Industrial Informatics*, 7(4), 529–539. https://doi.org/10.1109/TII.2011.2166794

Gururajapathy, S. S., Mokhlis, H., and Illias, H. A. (2017). Fault location and detection techniques in power distribution systems with distributed generation: A review. *Renewable and sustainable energy reviews*, *74*, 949-958.

Hong, Y. Y., and Huang, W. S. (2015). Locating High-Impedance Fault Section in Electric Power Systems Using Wavelet Transform, k-Means, Genetic Algorithms, and Support Vector Machine. *Mathematical Problems in Engineering*, *2015*(1), 823720.

Iurinic, L. U., Herrera-Orozco, A. R., Ferraz, R. G., and Bretas, A. S. (2015). Distribution systems high-impedance fault location: A parameter estimation approach. *IEEE Transactions on Power Delivery*, *31*(4), 1806-1814.

Karimi, M., and Ghaffarzadeh, N. (2022). High impedance fault detection based on current harmonic analysis using Chirp Z transform. *Energy Systems*, *13*(3), 813-833.

Kirtley, J. L. (2020). *Electric power principles: sources, conversion, distribution and use*. John Wiley and Sons.

APPENDIX A: Terminology Definitions

To ensure clarity and consistency throughout the manuscript, the following terms are defined as used in this study:

Term	Definition
Fault Detection	Binary decision process determining whether a fault exists in the network (Fault vs. No Fault).
Fault Classification	Multi-class identification process determining the specific type of fault (1PH HIFF, 2PH HIFFF, 3PH HIFF, 3PH Short Circuit)
Fault Location	Estimation of the distance to the fault along the feeder (not addressed in this study).
High Impedance Fault (HIF)	Fault characterized by low fault current due to high fault impedance, typically caused by contact with non-conducting materials such as dry soil, asphalt, gravel, or trees.
Discrete Wavelet Transform (DWT)	Signal processing technique for time-frequency analysis with multi-resolution capability.
Artificial Neural Network (ANN)	Machine learning model inspired by biological neural networks, used for pattern recognition.