# High Impedance Fault Detection and Localization Using Fully-Connected Convolutional Neural Network: A Deep Learning Approach

I. E. Abasi-obot[1*], A. B. Kunya[2], G. S. Shehu[2], Y. Jibril[2]

[1]Department of Electrical and Electronic Engineering, Akwa Ibom State University, Ikot Akpaden, Akwa Ibom State, Nigeria.
Department of Electrical Engineering, Ahmadu Bello University, Zaria, Nigeria.

***ABSTRACT:*** The detection and localization of high impedance faults (HIF) in power systems are challenging due to the low fault current magnitude, which often falls below the detection threshold of conventional devices. HIF events introduce harmonics into the network, posing risks to the safety of connected equipment, including the potential for igniting fire which endangers lives and properties. In this study, Emanuel's HIF model was used to generate HIF signatures resembling real HIF events. Model parameters were adjusted to mimic various contact surface impedances. Two datasets were created: 'no-fault' data, simulating the network without HIF, and 'fault' data, incorporating HIF waveforms by simulating single and multiple lines with the HIF model. The faulted line was divided into five segments along the 33 kV line to capture fault signatures at different locations. The generated data, including current waveforms and magnitudes, were processed and divided into an 80:20 ratio for training, validation, and testing using a deep fully connected Convolutional Neural Network for HIF detection and location. The results showed an impressive accuracy rate of 99.44% and 99.78% for detection and location respectively, representing a significant advancement in HIF detection and location, and offering practical applications for enhancing power line safety.

***KEYWORDS:*** *Distribution Network, High Impedance Fault, Deep Learning, Convolutional Neural Network*

[Received Nov. 4, 2023; Revised Nov. 16, 2023; Accepted Nov. 20, 2023] Print ISSN: 0189-9546 | Online ISSN: 2437-2110

## I. INTRODUCTION

In distribution networks, certain fault situations occur where, instead of the expected high fault current, a significantly low fault current is observed. This happens when a live conductor comes in contact with a high-impedance path, object, or surface. This type of fault is regarded as a high impedance fault (HIF). They are not cleared by the over current protection and even the earth fault relays fail to clear HIF because the pickup current (detection threshold) of these devices is higher than the HIF current. The HIF current is estimated to be less than one tenth of normal load current (Aljohani & Habiballah, 2020; Ghaderi *et al*., 2017). Hence, HIF could remain on the network unnoticed until damage is done (Vlahinić *et al.*, 2020). The sustained presence of this fault remains in unsafe condition endangering lives and properties due to the possibility of unsuspecting persons coming in contact with the live conductor (Farias *et al*., 2018; Hong & Huang, 2015). The devastating consequence of HIF includes electric shock, electrocution, wildfire, system malfunctioning and damages (Lavanya *et al.*, 2022; Ozansoy & Zayegh, 2023). HIF presents itself in two ways: It could be shunt or Series. For shunt, the conductor remains unbroken such as a live overhead conductor touching the ground via a poorly conducting surface such as tree branches (Kavaskar & Mohanty, 2019; Leão *et al*., 2022; Vieira *et al*., 2019). In the case of the series faults, they are designated by the presence of broken conductors. Series HIF can be subclassified as passive i.e., where no ground is involved, or active i.e., where there is contact to ground which could be either at source or load-side (Aljohani & Habiballah, 2020; Kavaskar & Mohanty, 2019; Vieira *et al*., 2019).

Identifying HIF in distribution networks is highly demanding, primarily because of the intricate, unpredictable, and irregular features that define these faults (Varghese *et al*., 2023). Typical characteristics of HIF include sporadic arcs, imbalances in current waveforms, the production of non-stationary and unstable currents, wide variations and unpredictability, high impedance value, generation of feeble signals, non-linearity, low current intensities, as well as the presence of distinctive characteristics like 'buildup' and 'shoulder' effects (Aljohani & Habiballah, 2020; Baharozu *et al*., 2023; Farias *et al*., 2018; Kavaskar & Mohanty, 2019; Lavanya *et al*., 2022). These characteristics are influenced by factors such as feeder configuration, weather condition, voltage levels, load type, network topology, system operation, ground surface material and surface humidity (Baharozu *et al*., 2023; Chaitanya *et al*., 2019; Ghaderi *et al*., 2017; Kujur & Biswal, 2017). These make the characteristic features of HIF events distinctively unique according to its origin.

The development of HIF detection and location techniques has since attracted concerted efforts in the power industry and

*Corresponding author: iniobongabasiobot@aksu.edu.ng

*doi:* http://dx.doi.org/10.4314/njtd.v18i4.2143

research institutions. This dates back to 1977 when the Electric Power Research Institute (EPRI) in California issued a request for proposals for a HIF detector, resulting in the creation of the first HIF model which utilized a spark gap phenomenon (Aucoin & Russell, 1982; Wang *et al*., 2017). Subsequently, the Institute of Electrical Electronics Engineers (IEEE), Power System Relaying Committee (PSRC) proposed an additional HIF model (Gautam & Brahma, 2012), and a computer model for HIF was introduced to compute the effects of harmonics generated by HIF (Gajjar, 1991).

The three primary methods for modeling HIF are:

1. Transient Analysis of Control Systems (TACS) Controlled Switch: Based on the arcing phenomenon, this approach models arc conduction, re-ignition, and extinction. It offers the advantage of adjusting the phase difference between the applied voltage and fault current (Wai & Yibin, 1998).
2. Dynamic Arc Model: This model is based on the energy balance within the arc column, originated from the Kizilcay arc model and derived from control theory principles (Kizilcay & Pniok, 1991; Zhang *et al*., 2016).
3. Emmanuel's HIF Model: This model employs two anti-parallel DC sources connected via two diodes, with variable resistors modeling effective impedance and non-linearity in fault current. It also accounts for the asymmetric nature and intermediate arc extinction around current zero (Yu & Khan, 1994). This method was selected for implementation in the study due to its ability to generate fault features resembling real-life HIF events.

The development of HIF models has significantly evolved, progressing from a simple one-resistor model to the diverse range of models available today. These models have greatly contributed to the understanding of HIF features and their application in solving HIF-related issues thereby giving rise to the various techniques being employed. HIF detection techniques are categorized in various ways depending on the approach adopted. The detection of HIF can be broadly grouped into two distinct groups based on the type of data utilized for the analysis. The use of data derived from controlled and staged faults, while the other hinges on data generated through computational simulations (Gautam & Brahma, 2012). The approach of utilizing the results of deliberately staged faults (also known as mechanical approach) is limited in that they are system-specific, hence, making them non-transferable to other systems or networks (Gautam & Brahma, 2012; Soheili *et al*., 2016). Furthermore, the approach of staged faults necessitates the use of specialized equipment and a skilled workforce and is intrinsically associated with the inherent risk of service disruption. Conversely, the techniques developed utilizing simulation data (often also categorized as electrical method) can be subjected to extensive testing, rendering them more adaptable to a wide array of systems or networks with minimal or no necessitated modifications (Soheili *et al*., 2016). Nevertheless, the simulated data may not be as inherently specific as may be required (Gautam & Brahma, 2012).

By methodological approach, HIF detection can be broadly categorized into two: classical and heuristic (Chakraborty & Das, 2018; Gomes & Ozansoy, 2021). In classical approaches, various techniques have been developed, primarily grounded in frequency domain analysis. These techniques encompass the examination of sequence components, odd harmonic components, even harmonic components, inter-harmonic components, and burst noise signals, all aimed at discerning patterns indicative of HIF. These methods take into account the presence of randomness and flickers within fault current waveforms. Popular feature extraction tools in this domain include the Fractal theorem, energy variance criterion, and Kalman filtering. Additionally, there are time domain-based techniques, such as the ratio ground relay-based protection scheme. Wavelet transform-based methodologies have also been proposed for HIF detection within this classical framework (Chakraborty & Das, 2018; Moloi & Davidson, 2022). Conversely, the heuristic approach to HIF detection involves methods such as knowledge-based expert systems and artificial neural networks. Notably, many of the techniques mentioned thus far rely on substation current measurements. However, it is important to acknowledge a limitation in the current-based approach, which is that the magnitude of fault current during HIF events tends to be exceptionally low (Chakraborty & Das, 2018; Moloi & Davidson, 2022).

To effectively detect HIF, a range of features beyond current magnitude is considered. These features encompass current harmonics content, waveforms, voltage, and arc properties. The frequency domain approach examines low and high-frequency components in HIF waveforms, primarily resulting from electric arcs. Features extracted for this technique include the various levels of harmonic components such as low-order harmonics, sub-harmonics, low-frequency spectrum high-frequency spectra (Adewole, 2012; Althi *et al*., 2022; Ghaderi *et al*., 2017). Nevertheless, a significant challenge lies in distinguishing HIF from similar events, as the HIF frequency spectrum closely resembles that of capacitor switching and other common system operating conditions.

Furthermore, the applicability of this approach is limited in traditional distribution networks due to the integration of smart devices (Aljohani & Habiballah, 2020; Ghaderi *et al*., 2017). Hybrid schemes, which combine time and frequency domain approaches, offer benefits such as improved feature interpretation speed and accuracy, as well as time and frequency localization. Nevertheless, hybrid schemes are characterized with disadvantages, including increased computational demands and the inability to differentiate non-linear loads from HIF, which elevates system noise and limits their practicality in smart networks (Ghaderi *et al*., 2017; Gomes & Ozansoy, 2021).

Heuristic schemes, commonly referred to as artificial intelligence techniques, are recognized for their speed, efficiency, accuracy, predictive capabilities, noise reduction, and compatibility with smart networks. Various AI techniques, including artificial neural networks (ANN), Support Vector Machine (SVM), decision trees, Adaptive Neuro-Fuzzy Inference System (ANFIS), and Genetic Algorithms (GA), are employed for HIF detection and location. However, these AI methods necessitate substantial training datasets, precise measurement, effective feature extraction tools, and thorough domain analysis. They are also characterized by a trial-and-error approach, which can introduce subjectivity and consume time (Naidu *et al*., 2020; Sedighizadeh *et al*., 2010; Soheili &

Sadeh, 2017). In summary, the detection of High Impedance Faults involves a multifaceted approach, incorporating the analysis of various features. Frequency Domain methods focus on the spectral characteristics of HIF waveforms but struggle to distinguish them from similar events. Hybrid schemes combine time and frequency domain techniques, while Heuristic schemes leverage AI methods for rapid and accurate detection, though they require careful training and domain analysis.

Also, most of the works done so far centred on series HIF being that a break on the conductor creates significant fault signatures for a detection threshold. Thus, the fault profiles at discontinuity vary significantly. This paper proposes a deep learning-based technique using CNN. The choice of a deep learning technique for this study is inspired by its capabilities given the complex nature of HIF signatures. These includes distinct pattern recognition and feature extraction directly from datasets without the need for a feature extraction procedure (which has its own computational complexities such as losing some fault signatures during extraction). This study focuses on deploying the unique features of the convolution layer in CNN, to detect and locate HIF in a typical 33 kV/11 kV network to reduce its devastating consequences. The future for this field includes the application of CNN to solve HIF problems. On this premise, a deep learning technique based on CNN which is purely data driven, capable of learning from complex and multidimensional data directly without the need for a feature extraction procedure, offers translation invariance, functions accurately regardless of noise, understands patterns at different levels, learns the details and overview simultaneously, and is adaptable to new systems using transfer learning is chosen for this study. The remainder of this paper is structured as follows. Section 2 presents the methodology which captures HIF modelling, and the development (building, loading, and training) of the CNN. In section 3, the results are presented and discussed while section 4 concludes the study.

## II. METHODOLOGY

High Impedance Faults (HIFs) pose a significant challenge in power distribution systems due to their elusive and potentially hazardous nature. Detecting and localizing HIFs is crucial for maintaining the reliability and safety of electrical grids. Traditional fault detection methods often struggle with identifying HIFs, which are characterized by high resistance and limited fault current. In response to this challenge, this study presents a novel approach to HIF detection and localization by harnessing the power of deep learning, specifically through the utilization of a Fully-Connected Convolutional Neural Network (FC-CNN). The detection of fault and its location is formulated based on high impedance modelling followed by the system's description.

The methodology presented here aims to enhance the accuracy and efficiency of HIF detection and localization using FC-CNNs. This deep learning approach leverages the capabilities of neural networks in pattern recognition and feature extraction, enabling the identification of HIFs, even in cases where conventional methods might not be applicable.

### *A. Formulation of the Fault Detection and Location Model*

The Emanuel's HIF model adopted in this study was used to generate HIF signatures similar to those obtained in real-life HIF events. Fig. 1 shows the Emanuel's HIF model implemented using Simulink in MATLAB 2022a. Its parameters were selected and continuously varied to model the impedance of various contact surfaces. This model consists of two anti-parallel DC sources ($V_1$ and $V_2$) connected via two diodes ($D_1$ and $D_2$), and two resistances ($R_1$ and $R_2$) in series with the DC sources. Two data sets were obtained, the no-fault data obtained by simulating the network without the HIF model staged on the network and the fault data (HIF current waveforms and magnitude) obtained by simulating the network with the HIF model staged on single-lines and multiple lines. The HIF model parameters range: $V_1$ and $V_2$ were varied between 5 kV to 8 kV, $R_1$ and $R_2$ were within a range of 150 Ω to 40 kΩ.

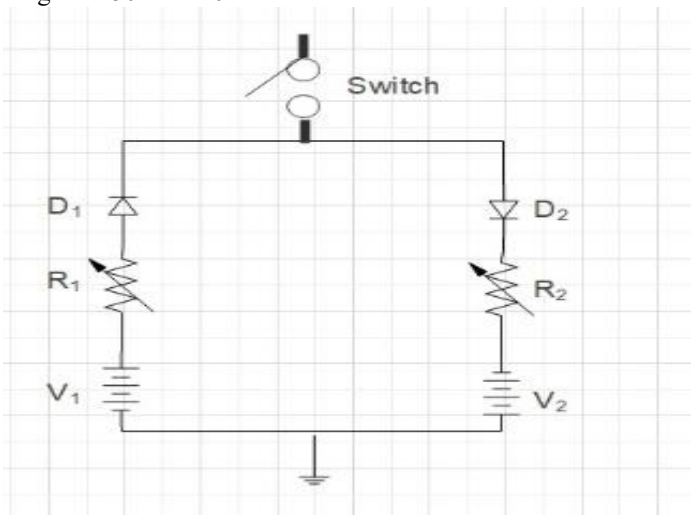


**Fig. 1: Implementation of HIF using SIMULINK**

To capture fault signatures at different points on the line data collection for the five locations considered in the study, the faulted 33 kV line was segmented into five points: 10%, 25%, 50%, 75% and 90% of the total length of the line (L). The faulted 33 kV line was simulated for single and multiple lines at each of the five points (locations) considered. This implies that one fault type (say Line B to ground fault) was simulated five times. This was repeated for all seven fault types: line A to Ground (A-G), line B to Ground (B-G), Line C to Ground (C-G), lines AB to Ground (AB-G), lines AC to G (AC-G), lines BC to Ground (BC-G) and ABC to line (ABC-Line). The generated data (current waveforms and magnitude) were exported to workspace and stored in cell arrays as time series data. This was shared in a ratio of 80:20 for training validation and testing respectively. Detailed description of the HIF model and the procedure for obtaining the training dataset used for training the deep CNN in this study had been previously published as contained in (Abasi-obot *et al*., 2022).

### *B. Modeling of High Impedance Fault*

This study presents an analytical model of HIF by faulting a transmission line as illustrated in Fig. 2.


**Fig. 2: Transmission line faulted due to HIF**

In Fig. 2, $V_s$ is the source voltage represented by an ideal alternating current source, $Z_s$ is the internal impedance of source $V_s$, $I_s$ is the source current, $Z_L$ is the load impedance and HIF is the high impedance fault incident on the line. In Fig. 3, the HIF is expressed in terms of its resistive, inductive, and capacitive (RLC) components, hence the expanded circuit. From Fig. 3, the mathematical expressions for the source voltage $V_s$, capacitor voltage $V_C$, fault voltage $V_F$, source current $I_s$, load current $I_L$, fault current $I_F$, current in the transmission line $I_T$, Resistance of the transmission line $R_s$, and inductance of the transmission line $L_T$ are derived as given in Eqns. (1) through (9).

$$V_s = I_s R_1 + L_1 \rho I_s + V_c \quad (1)$$
$$V_c = I_T R_T + L_T \rho I_T + V_F \quad (2)$$
$$V_F = I_L R_L + L_L \rho I_L \quad (3)$$
$$I_s = I_c + I_T \quad (4)$$
$$I_c = I_s - I_T \quad (5)$$
$$I_c = C \rho V_c \quad (6)$$

Where $R_1$ and $L_1$ are the resistive and inductive components of source impedance $Z_S$; $R_T$, $L_T$ and $I_T$ are the resistance, inductance and current of the transmission line due to HIF; $R_L$ and $L_L$ are the resistive and inductive components of the load impedance $Z_L$. $C_1$ is the capacitive component of the fault and GND1, GND2 GND3 GND4 are the grounds. $\rho V_c$ is the differential of the capacitor voltage, $V_c$.

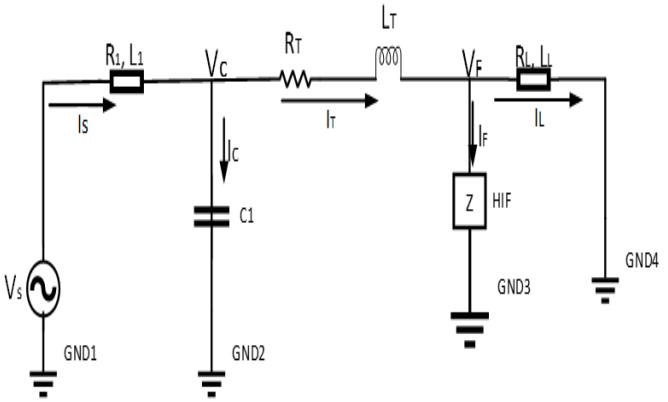


**Fig. 3: Faulted Line Showing RLC components of HIF.**

Observing Eqns. (5) and (6), they are equivalent as given in Eqn. (7)

$$C \rho V_c = I_s - I_T \quad (7)$$
$$I_T = I_L + I_F \quad (8)$$

Substituting Eqn. (8) in Eqn. (4),

$$I_s = I_c + I_L + I_F \quad (9)$$

Considering the arcing phenomenon at the fault point, fault voltage $V_F$ and fault current $I_F$ are determined with respect to the Mayr arc model. In (Sirojan *et al.*, 2018), the mathematical model of the electric arc behavior at fault point is reported. The Mayr arc model proposes that conductance G(t) varies exponentially with heat stored in the arc. This is given as (10).

$$\frac{1}{G_{(t)}} = \frac{dG_{(t)}}{dt} = \frac{1}{\tau_m}\left[\frac{u_{(t)} * i_{(t)}}{\rho_m} - 1\right] \quad (10)$$

where G = Arc conductance, $\tau_m$= arc time constant, $\rho_m$= cooling power of arc, u= arc voltage, $I$ =arc current. This analysis is aimed at obtaining the fault voltage $V_F$ and fault current $I_F$ in terms of the arc voltage *u* and arc current *i*, $V_F$ and $I_F$ are substituted for *u(t)* and *i*(t) respectively in (11).

$$\frac{1}{G_{(t)}} = \frac{dG_{(t)}}{dt} = \frac{1}{\tau_m}\left[\frac{V_F * I_F}{\rho_m} - 1\right] \quad (11)$$

Conductance:

$$G_{(t)} = \frac{I_F}{V_F} \quad (12)$$

And

$$\frac{1}{G_{(t)}} = \frac{V_F}{I_F} \quad (13)$$

Substituting Eqns. (12) and (13) into Eqn. (11), Eqn. (14) is obtained.

$$\frac{V_F}{I_F} = \rho \frac{I_F}{V_F} = \frac{1}{\tau_m}\left[\frac{V_F * I_F}{\rho_m} - 1\right] \quad (14)$$

Eqn. (14) is the arc model equation expressed in terms of fault current $I_F$ and fault voltage $V_F$. Resolving Eqns. (14), Eqn. (15) is obtained whose further resolution gives Eqn. (16). From Eqns. (7) and (9); Eqns. (17) and (18) are obtained, respectively.

$$\frac{V_F}{I_F}\left[\frac{V_F \rho I_F - I_F \rho V_F}{V_F^2}\right] = \frac{1}{\tau_m}\left[\frac{V_F * I_F}{\rho_m} - 1\right] \quad (15)$$
$$\frac{1}{I_F}\rho I_F - \frac{1}{V_F}\rho V_F = \frac{1}{\tau_m}\left[\frac{V_F * I_F}{\rho_m} - 1\right] \quad (16)$$
$$\rho V_c = \frac{I_s}{C} - \frac{I_T}{C} \quad (17)$$
$$\rho I_s = \rho I_c + \rho I_L + \rho I_F \quad (18)$$

From Eqns. (1), (2), (3) and (8), the differentials are expressed as given in Eqns. (19), (20), (21) and (22), respectively. The six dynamic systems describing the parameters of a network faulted by HIF and considering the arching phenomenon at fault point are derived as expressed in Eqns. (17), (19), (20), (21), (26) and (28) where $\rho V_c$, $\rho I_s$, $\rho I_L$, $\rho I_T$, $\rho I_F$, $\rho V_F$ are the differentials of the capacitor voltage $V_c$, source current $I_s$, load current $I_L$, Transmission line current $I_T$, fault current $I_F$ and fault voltage $V_F$ respectively.

$$\rho I_s = -\frac{R_1 I_s}{L_1} - \frac{V_c}{L_1} + \frac{V_s}{L_1} \quad (19)$$
$$\rho I_T = -\frac{R_T I_T}{L_T} - \frac{V_F}{L_T} + \frac{V_{sF}}{L_T} \quad (20)$$
$$\rho I_L = -\frac{R_1 I_s}{L_L} - \frac{V_F}{L_L} \quad (21)$$
$$\rho I_T = \rho I_L + \rho I_F \quad (22)$$

Thus,

$$\rho I_F = \rho I_T - \rho I_L \quad (23)$$

Substituting Eqns. (20) and (21) in (23), Eqn. (24) is obtained,

$$\rho I_F = -\frac{R_1 I_s}{L_L} - \frac{V_F}{L_L} + \frac{V_c}{L_T} - \left[-\frac{R_1 I_L}{L_L} + \frac{V_F}{L_L}\right] \quad (24)$$

$$\rho I_F = -\frac{R_1 I_s}{L_L} - \frac{V_F}{L_L} + \frac{V_c}{L_T} + \frac{R_L I_L}{L_L} - \frac{V_F}{L_L} \quad (25)$$

$$\rho I_F = \frac{R_L I_L}{L_L} - \frac{R_T I_T}{L_T} - \left[\frac{1}{L_T} + \frac{1}{L_L}\right] V_F + \frac{V_c}{L_T} \quad (26)$$

Substituting (26) into (16), (27) is obtain:

$$\frac{1}{I_F}\left[\frac{R_L I_L}{L_L} - \frac{R_T I_T}{L_T} - \left[\frac{1}{L_T} + \frac{1}{L_L}\right] V_F + \frac{V_c}{L_T}\right] - \frac{1}{V_F}\rho V_F = \frac{1}{\tau_m}\left[\frac{V_F * I_F}{\rho_m} - 1\right] \quad (27)$$

From (27) the differential of the fault voltage $\rho V_F$ can be expressed as in (28):

$$\rho V_F = \frac{V_F}{I_F}\left[\frac{R_L I_L}{L_L} - \frac{R_T I_T}{L_T} - \left[\frac{1}{L_T} + \frac{1}{L_L}\right] V_F + \frac{V_c}{L_T}\right] - \frac{V_F}{\tau_m}\left[\frac{V_F * I_F}{\rho_m} - 1\right] \quad (28)$$

### C. *Description of 33 kV/11 kV New-Haven Distribution Network*

This study was conducted on the 33 kV/11 kV New-Haven district network located in Enugu State, Nigeria. The 33 kV radial distribution network receives its power supply from the Okpai power station at 330 kV, which is then stepped down to 132 kV through two 330 kV/132 kV, 150 MVA transformers. The 132 kV supply from New-Haven is further reduced to 33 kV using a single 60 MVA 132/33 kV transformer. This 33 kV network is distributed through four 33 kV buses, which, in turn, supply power to thirteen buses, comprising seven 33 kV and six 11 kV buses. To facilitate voltage reduction, nine transformers are strategically installed at different points, with three rated at 15 MVA and six at 7.5 MVA.

These transformers serve to step down the voltages supplied to 14 load feeders, each catering to various loads of different magnitudes. In Figure 4, a simplified representation of the 33 kV/11 kV New-Haven to Government House line is depicted, highlighting the occurrence of High Impedance Faults (HIF) on lines AB-G. The faulted line(s) were deliberately manipulated to capture fault signatures for seven distinct fault types at five specific locations considered in this study.

This dataset was instrumental in creating the Convolutional Neural Network (CNN) training dataset, comprising fault current waveforms and their corresponding magnitudes. In total, this study generated a vast dataset, consisting of 1,400,635 rows and 108 columns of time series data, which was subsequently employed to train the developed CNN model.

### D. *Development of HIF Detection and Location Model using CNN*

The developed CNN model for implementation of HIF detection and location was designed using the application module of MATLAB 2022a. Fig. 5 shows the process flow for using the application Module of MATLAB 2022a to solve implement the fully connected CNN. The following layers were connected: 1-sequence input layer, 4-convolution layers, 4-rectified linear unit (ReLU) layers, 4-layer normalization layers, 1-global average, 1-global average pooling layer, 1-fully connected layer, 1-softmax layer, and 1-output classification layer. A total of seventeen layers were connected as shown in Fig. 6. The structure of the convolution layer consists of filters, responsible for identifying the features or patterns in the data that have been fed to the input layer. Thus, the convolution operation is basically passing the inputs through the filters in a sliding manner, each filter applies a unique operation to extract the features or patterns. To perform activation operation on the input the Rectified linear Unit (ReLU) is connected, which performs a nonlinear threshold operation where any input value less than zero is set to zero). The normalization layer is connected to normalize each input channel across a mini batch. This speeds up CNN's training and reduce the sensitivity to network initialization. Hence, the normalization layer is used between convolutional layers in this study. The global average pooling layer is connected after the fourth convolution set (ending with the fourth normalization layer) to perform down-sampling of the features

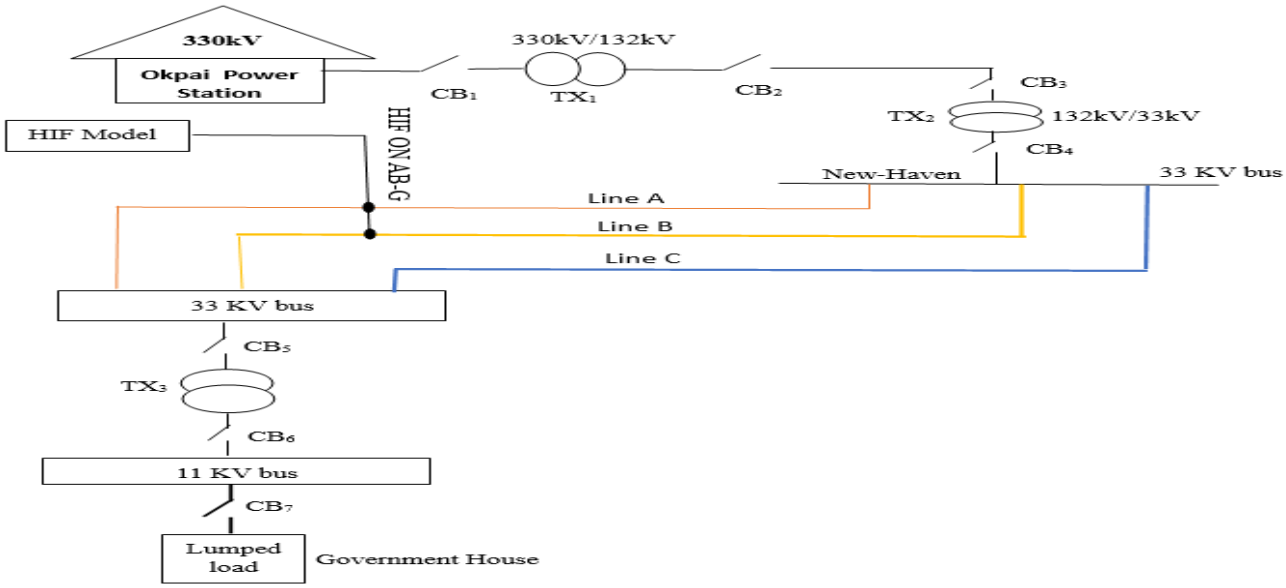


**Figure 4: HIF Incident on Lines AB-G of 33 kV/11 kV New-Haven to Government House Feeder**

by computing the mean of the height and width of its dimensions. The fully connected layer multiplies the input by a weight matrix and then adds a bias vector. The softmax layer is connected before the classification output layer of the deep learning network to perform a softmax operation (a complex arithmetic operation involving exponential functions and divisions).

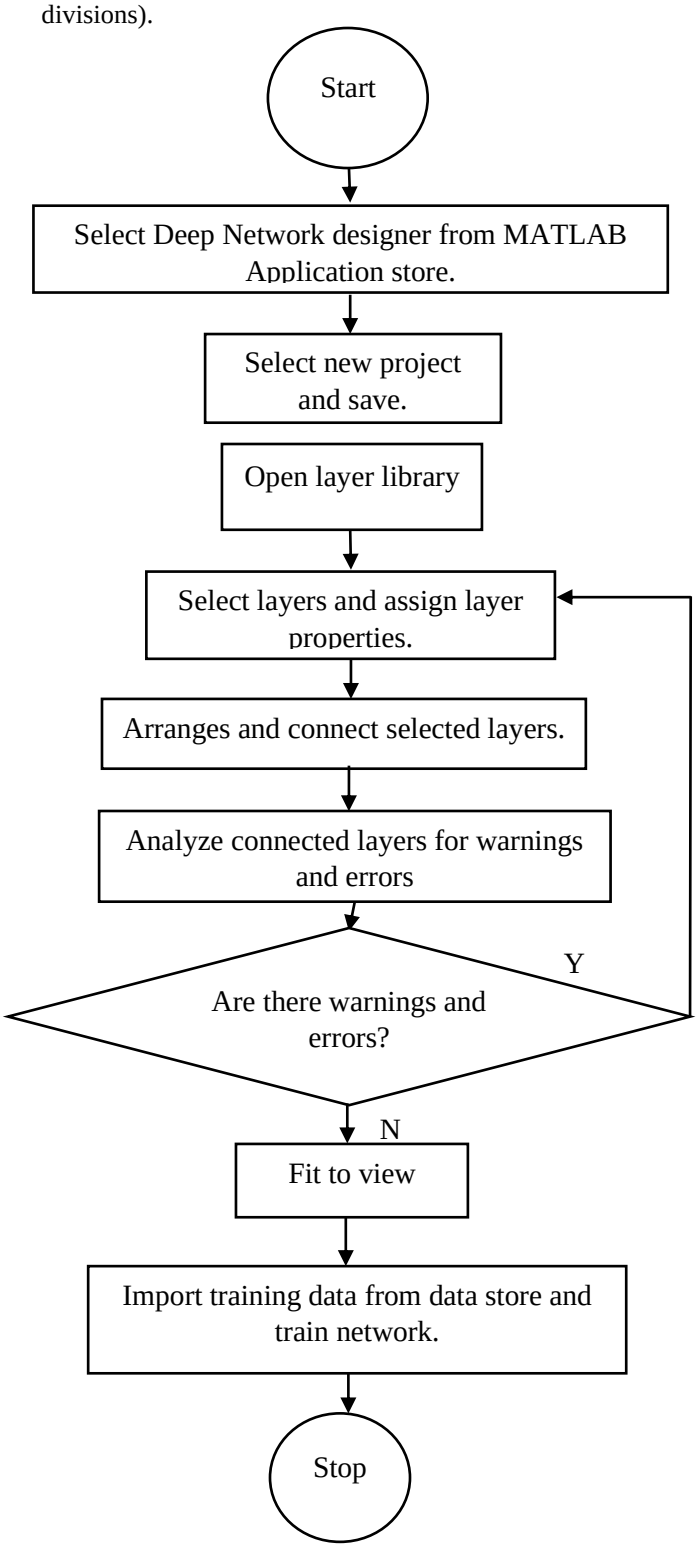


**Fig. 5: The process flow for the implementation of CNN using the application module of MATLAB 2022a**

Finally, the classification output layer is connected to compute the cross-entropy loss for multi-class classification problems with mutually exclusive classes. The CNN architecture was analyzed for errors and warnings, the identified errors were cleared by following the error prompt and reselecting the properties of the layers, and the network was re-analyzed until all errors and warnings were cleared. Fig. 5 shows the CNN developed in this study for HIF detection and location. The properties of the seventeen-layered CNN are as follows: the sequence input layer has an input size of 3, with a minimum length selected as 1 and normalization is selected as none while the normalization dimension is automatically selected. A filter size of 3 was chosen for the convolution layer, this is constant for all the convolution layers in the network, for the first convolution layer, the number of filters used is 32, this increases by a multiple of 32 for the subsequent convolution layer. Thus, the number of filters increases from 32 to 64, from 64 to 96 and from 96 to 128. The convolution layer padding is selected as causal. The stride value is set at 1. To speed up training and reduce the sensitivity to network initialization the epsilon is set to the default value of 0.00001. The fully connected layer multiplies the input by a weight matrix and then adds a bias vector. The output size is selected as 8-for the detection of HIF. This represents seven fault types: (line A to Ground (A-G), line B to Ground (B-G), Line C to Ground (C-G), lines AB to Ground (AB-G), lines AC to G (AC-G), lines BC to Ground (BC-G) and ABC to line (ABC-Line)) and one no-fault condition. This is presented as an 8-by-1 single-column matrix. An output size of 6 is chosen for the HIF location. This represents five possible fault locations, assigned as: (0.1*L, 0.25*L, 0.5*L, 0.75*L and 0.9*L), where: L is the length of line, and one no-fault condition, presented as a 6-by-1 single-column matrix. The developed CNN model was trained to detect and locate HIF, a training validation accuracy of 100% was attained after training for 13 minutes 7 seconds at a learning rate of 0.01 with a validation frequency of 5 iterations. The CNN was tested, results for four scenarios are presented and discussed in section 3.

## III. RESULTS AND DISCUSSION

In this section, results obtained from the testing and implementation of the developed CNN model was carried out based on four selected scenarios out of numerous possible combinations. For each of the lines, it was possible to simulate depending on the fault location along the transmission lines. This could be at 10% (near the receiver's end), 25%, 50%, 75%, 90% (near the sending end) of the transmission lines or no-fault scenario. In the first scenario (Scenario 1), HIF was implemented by selecting a fault occurring midway on the line AB-G i.e., at 50% of the length AB-G. The second and third scenarios selected were simulated for faults at 75% and 25% of the lines C-G and ABC, respectively. In the fourth scenario, HIF was implemented for fault near the receiving end of the line B-G. Table 1 gives a summary of the selected scenario showing the staged fault locations on the lines.

The four scenarios were implemented one after the other by simulating HIF on the selected lines and at the indicated locations. For each of the simulated scenario, performance evaluation in terms of the model's accuracy, precision, recall

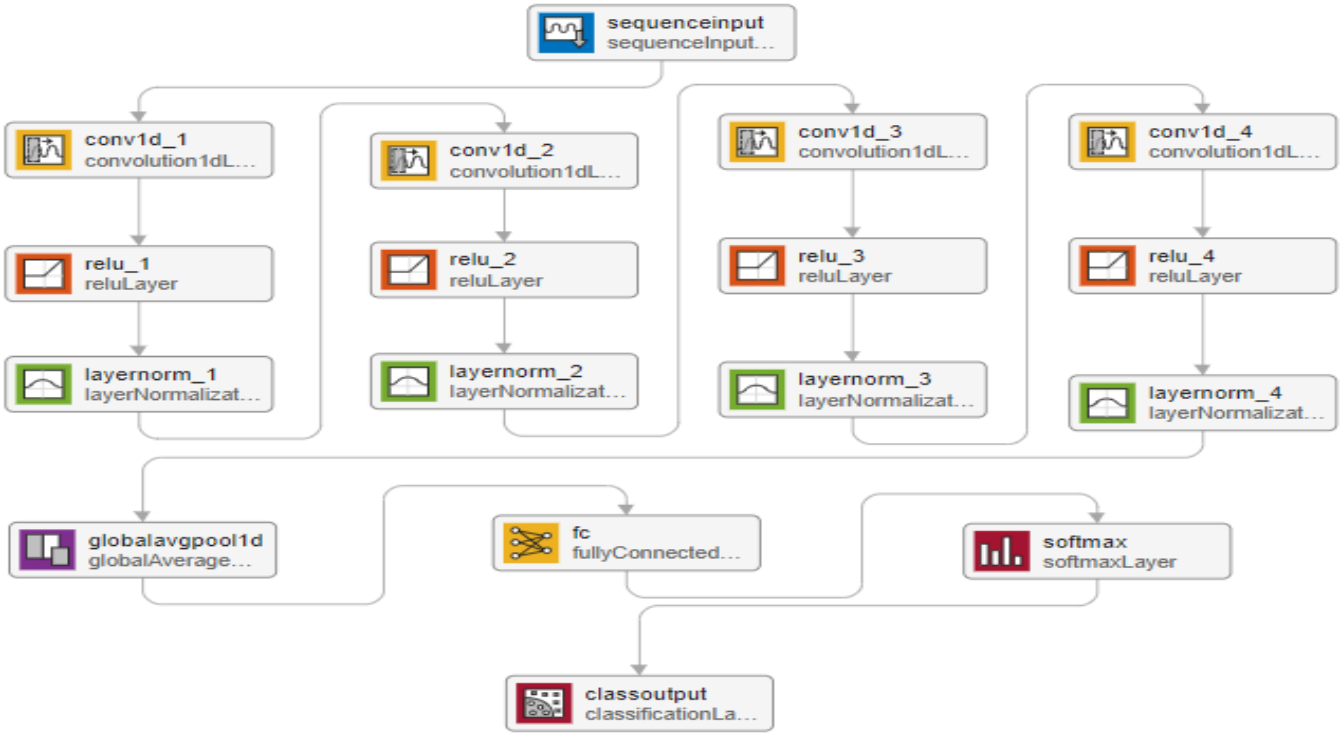


**Fig. 6: Developed Layers for the CNN Model for HIF Detection and Location on the selected 33 kV line**

and F1 score were recorded to give a direction on the effectiveness of the developed model in the determination of HIF.

**Table 1 Selected fault location for simulation**

| Scenario | Line | Selected Fault Location (% of Length) |
|---|---|---|
| 1 | AB-G | 50 |
| 2 | C-G | 75 |
| 3 | ABC | 25 |
| 4 | B-G | 10 (Near the receiving end) |

### *A. Scenario 1*

A High Impedance Fault was introduced to the case study network along lines AB-G at the midpoint of the line (scenario 1), as illustrated in Fig. 4. To detect and locate the staged fault, the generated data was processed using the developed CNN model. Table 2 shows the performance parameters for the detection and location. From table 2, highest detection accuracy obtained was 99.44% for the fault type AB-G while 99.78% was obtained as the location accuracy. The results indicate that the model accurately detected the staged fault condition. With a precision of 0.9957 and F1 score of 0.9960 for the detection, there exist an extremely low case of HIF not being detected.

**Table 2 Evaluation Parameters for Scenario 1**

| Evaluation parameters | Detection | Location |
|---|---|---|
| Accuracy (%) | 99.44 | 99.78 |
| Precision | 0.9957 | 0.9986 |
| Recall | 0.9963 | 0.9983 |
| F1 score | 0.9960 | 0.9984 |

### *B. Scenario 2*

In this scenario, the HIF was staged precisely at 75% length of line C-G. Implementing the developed model gives the obtained parameters as given in Table 3. The performance metrics according to the data in Table 3, indicate the highest accuracy achieved for fault detection was 97.66% for the C-G fault type, and 96.14% accuracy was achieved for fault location. These results indicate that the model effectively identified the intentionally created fault condition. Moreover, with a precision score of 0.9857, recall of 0.9809 and an F1 score of 0.9833 for the detection, the model can be comfortable deployed.

**Table 3 Evaluation Parameters for Scenario 2**

| Evaluation parameters | Detection | Location |
|---|---|---|
| Accuracy (%) | 97.66 | 96.14 |
| Precision | 0.9857 | 0.9688 |
| Recall | 0.9809 | 0.9762 |
| F1 score | 0.9833 | 0.9725 |

### *C. Scenario 3*

In this particular scenario, the HIF was deliberately placed at the precise location of 25% along the ABC line. The application of the developed model yielded the parameters outlined in Table 4. The performance metrics presented in Table 4 reveal that the highest level of accuracy achieved for fault detection was 98.48% for the C-G fault type, and there was a 96.16% accuracy in pinpointing the fault's location. These findings indicate that the model proficiently recognized the intentionally introduced fault condition. Furthermore, with a precision score of 0.9857, a recall of 0.9809, and an F1 score of 0.9833 for detection, the model can be confidently deployed.

### *D. Scenario 4*

This situation presents a different scenario where the fault is located close enough to the receiving end along the line B-

G. This scenario requires immediate response as fault must be detected and located before any significant damage is done. Implementing the developed CNN model gives the values of Table 5 for accuracy, precision, recall and F1 score as given. The location performance evaluation gives higher values than those of the detection algorithms.

**Table 4 Evaluation Parameters for Scenario 3**

| Evaluation parameters | Detection | Location |
|---|---|---|
| Accuracy (%) | 98.48 | 96.16 |
| Precision | 0.9900 | 0.9718 |
| Recall | 0.9883 | 0.9729 |
| F1 score | 0.9891 | 0.9723 |

**Table 5 Evaluation Parameters for Scenario 4**

| Evaluation parameters | Detection | Location |
|---|---|---|
| Accuracy (%) | 90.86 | 99.05 |
| Precision | 0.9424 | 0.9947 |
| Recall | 0.9272 | 0.9917 |
| F1 score | 0.9347 | 0.9932 |

*E. Summary of Findings*

From the eight fault types (as shown in the 8-by-8 confusion matrix of Fig. 7) presented in this study even though only four scenarios (selected randomly) were simulated, more scenarios could have been selected. Nevertheless, the four scenarios were just enough to demonstrate the performance of the developed model. The matrix of Fig. 7 encompasses the eight fault types with seven fault condition and one 'no-fault' condition. The true class and the predicted class for the fault types are depicted along the x-axis and y-axis respectively. The diagonal elements of the matrix reveal that each of the fault types can occur at any of the five possible locations, as indicated by the number 5. The 'no-fault' condition is represented by number 1, signifying a single event.

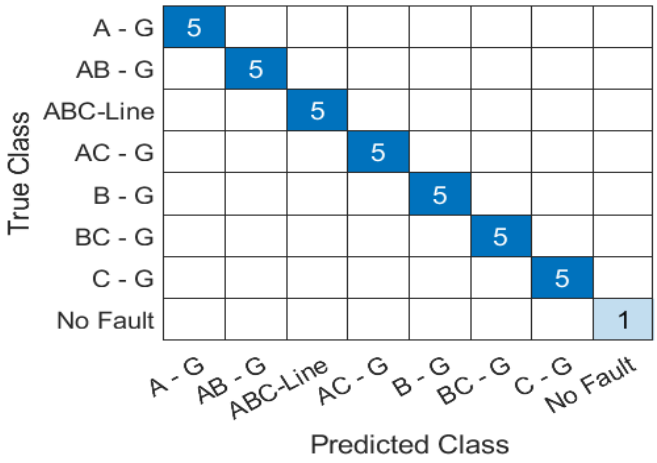


**Fig. 7: Confusion Matrix of the CNN model for HIF detection showing all HIF types considered in the study.**

This implies that each of the seven fault types presented could have occurred on any of the five locations considered in the study, else, it is a no-fault condition. The diagonal element in the confusion matrix captures these 7 fault types as represented by number 7, and the no-fault on the line is captures as a single incident represented by number 1. This is shown in Fig. 8.

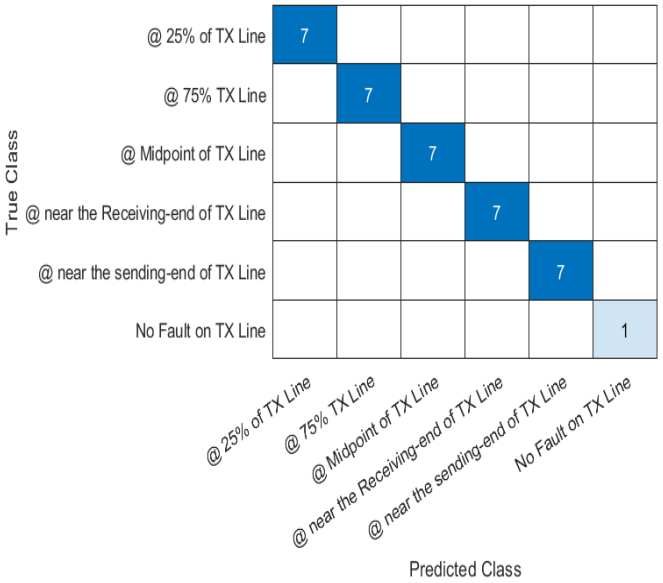


**Fig. 8: Confusion Matrix for HIF location showing all locations captured in the study.**

In summary, the implemented scenarios represent a true reflection of the other fault conditions. For each of the scenarios, a very high accuracy is achieved with high values of precision, recall and F1 indicating that the model performs excellently well. The summary of the obtained performance parameters is as given in Table 6. For all cases, the accuracy is far greater than 90% confirming the position of the study carried out by Rai *et al.* (2021) where the authors consider different scenarios with noises. This works presents a simpler and a more direct approach by adopting Emanuel's HIF model for generating real life fault cases. Nevertheless, the study of HIF detection and location requires that false cases are not significant given the cost implication of possible equipment damage where faults are not detected or shutdown of systems in cases of false alarms.

## IV. CONCLUSION

As a result of the devastating consequences experienced when HIF is not promptly detected and located, this study was embarked upon to develop a deep fully-connected CNN to detect and locate HIF. Due to the dearth of fault records, and because CNN is purely data driven, the first part of the work was to obtain HIF features by modelling and simulating HIF in the 33 kV/11 kV network in MATLAB. From this, the datasets required for training the CNN was generated and stored. The CNN was developed and trained using the deep network designer in MATLAB version 2022. The developed network was tested under different staged scenarios. The performance of the CNN proposed showed detection and location results up to 99.44% and 99.78% respectively. Comparatively, an improvement of 2.44% up to 3.96% was obtained when these results were compared with related studies on the basis of accuracy. This paper focused on shunt HIF, which has complex signatures compared to series HIF. Also, the detection and location of HIF was captured while most literatures only concentrated on detection. These shows a significant improvement in this area. The CNN developed in this study is

**Table 6 Summary of the results of the CNN Model for both Detection and Location for all fault types**

| | Detection | | | | Location | | | |
|---|---|---|---|---|---|---|---|---|
| Fault type | AB-G | C-G | ABC Line | B-G | AB-G | C-G | ABC Line | B-G |
| Accuracy (%) | 99.44 | 97.66 | 98.48 | 90.86 | 99.78 | 96.14 | 96.16 | 99.05 |
| Precision | 0.9957 | 0.9857 | 0.9900 | 0.9424 | 0.9986 | 0.9688 | 0.9718 | 0.9947 |
| Recall | 0.9963 | 0.9809 | 0.9883 | 0.9272 | 0.9983 | 0.9762 | 0.9729 | 0.9917 |
| F1 score | 0.9960 | 0.9833 | 0.9891 | 0.9347 | 0.9984 | 0.9725 | 0.9723 | 0.9932 |

adaptable to other networks and it's self-improving. It can be deployed as an additional functionality in relaying circuits of smart protective devices to detect and locate HIF on the power networks within a few power cycles with high precision.

## ACKNOWLEDGMENTS

The authors would like to acknowledge the support received from the Tertiary Education Trust Fund (TETFund) Nigeria, National Electricity Regulatory Commission (NERC), Nigeria, the Bureau of Education and Cultural Affairs (ECA) of the US Department of States and the Fulbright Scholarship Board for funding. Many thanks to Tennessee Technological University, Cookeville, Tennessee, USA.

## DECLARATIONS

Ethical Approval

This declaration is not applicable in this study as it's not a human or animal studies.

## COMPETING INTEREST

The authors would like to declare that there are no competing interests.

## AUTHOR CONTRIBUTIONS

The review of literatures on the subject matter and the research concept was formulated and spearheaded by **A. Kunya**. **I. Abasi-obot**, **G. Shehu:** was responsible for all the modelling and simulations in the study. **I. Abasi-obot:** also prepared the manuscripts while all other authors assessed and corrected the manuscript. The Team Chair, **Y. Jibril** vetted and approved the final manuscript for submission.

## FUNDING

This study has received funds from the following organizations: Tertiary Education Trust Fund (TETFund), Nigeria. National Electricity Regulatory Commission Nigeria (NERC), Fulbright Foreign Student Scholarship through the Bureau of Educational and Cultural affairs of the US Department of States. Fulbright Identity Number: PS00316173. All funds for this work were received by the corresponding author.

## DATA AVAILABILITY

The Single Line Diagram (SLD) of the case study network and the data generated during the current study for various impedance values and for different locations on the case study network are available from the corresponding author upon reasonable request.

## REFERENCES

**Abasi-obot, I.; G. Shehu; A. Kunya and Jibril, Y. (2022).** High Impedance Fault Modelling and Simulation of 33kV/11kV Distribution Network Using MATLAB. 2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON),

**Adewole, A. C. (2012).** *Investigation of methodologies for fault detection and diagnosis in electric power system protection* Cape Peninsula University of Technology].

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

**Althi, T. R.; E. Koley and S. Ghosh. (2022).** An Optimally Tuned Rotation Forest-Based Local Protection Scheme for Detecting High-Impedance Faults in Six-Phase Transmission Line During Nonlinear Loading. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, *46*(4), 1129-1147.

**Aucoin, B. M. and Russell, B. D. (1982).** Distribution high impedance fault detection utilizing high frequency current components. *IEEE Transactions on Power Apparatus and Systems*(6), 1596-1606.

**Baharozu, E.; S. Ilhan and G. Soykan. (2023).** High impedance fault localization: A comprehensive review. *Electric Power Systems Research*, *214*, 108892.

**Chaitanya, B. K.; A. Yadav and M. Pazoki. (2019).** An intelligent detection of high-impedance faults for distribution lines integrated with distributed generators. *IEEE Systems Journal*, *14*(1), 870-879.

**Chakraborty, S. and Das, S. (2018).** Application of smart meters in high impedance fault detection on distribution systems. *IEEE Transactions on Smart Grid*, *10*(3), 3465-3473.

**Farias, P. E.; A. P. De Morais; J. P. Rossini and, G. Cardoso Jr. (2018).** Non-linear high impedance fault distance estimation in power distribution systems: A continually online-trained neural network approach. *Electric Power Systems Research*, *157*, 20-28.

**Gajjar, J. (1991).** Efficient model for computing high-impedance fault generated harmonic propagation effects on radial power distribution feeders. 1991., IEEE International Sympoisum on Circuits and Systems,

**Gautam, S. and Brahma, S. M. (2012).** Detection of high impedance fault in power distribution systems using mathematical morphology. *IEEE Transactions on Power Systems*, *28*(2), 1226-1234.

**Ghaderi, A.; H. L. Ginn III and H. A. Mohammadpour. (2017).** High impedance fault detection: A review. *Electric Power Systems Research*, *143*, 376-388.

**Gomes, D. P. and Ozansoy, C. (2021).** High-impedance faults in power distribution systems: A narrative of the field's developments. *ISA transactions*, *118*, 15-34.

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

**Kavaskar, S. and Mohanty, N. K. (2019).** Detection of high impedance fault in distribution networks. *Ain Shams Engineering Journal*, *10*(1), 5-13.

**Kizilcay, M. and Pniok, T. (1991).** Digital simulation of fault arcs in power systems. *European Transactions on Electrical Power*, *1*(1), 55-60.

**Kujur, A. P. and Biswal, T. (2017).** Detection of high impedance fault in distribution system considering distributed generation. 2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA),

**Lavanya, S.; S. Prabakaran and N. A. Kumar. (2022).** Behavioral Dynamics of High Impedance Fault Under Different Line Parameters. *energy*, *9*, 12.

**Leão, A.; J. Vieira; W. Heringer; A. Sousa; R. Gadelha and M. Santos. (2022).** Experimental Investigation of High Impedance Faults in MV Overhead Distribution Networks due to Mango Tree Branches. *Simpósio Brasileiro de Sistemas Elétricos-SBSE*, *2*(1).

**Moloi, K. and Davidson, I. (2022).** High Impedance Fault Detection Protection Scheme for Power Distribution Systems. *Mathematics*, *10*(22), 4298.

**Naidu, K.; M. S. Ali; A. H. Abu Bakar; C. K. Tan; H. Arof and H. Mokhlis. (2020).** Optimized artificial neural network to improve the accuracy of estimated fault impedances and distances for underground distribution system. *Plos one*, *15*(1), e0227494.

**Ozansoy, C. R. and Zayegh, A. (2023).** Optimal Sampling Requirements for Robust and Fast Vegetation High Impedance Fault Detection. *IEEE Access*.

**Rai, K.; F. Hojatpanah; F. Badrkhani Ajaei and K. Grolinger. (2021).** Deep learning for high-impedance fault detection: Convolutional autoencoders. *Energies*, *14*(12), 3623.

**Sedighizadeh, M.; A. Rezazadeh and N. I. Elkalashy. (2010).** Approaches in high impedance fault detection a chronological review. *Advances in Electrical and Computer Engineering*, *10*(3), 114-128.

**Sirojan, T.; S. Lu; B. Phung; D. Zhang and E. Ambikairajah. (2018).** High Impedance Fault Detection by Convolutional Deep Neural Network. 2018 IEEE International Conference on High Voltage Engineering and Application (ICHVE),

**Soheili, A. and Sadeh, J. (2017).** Evidential reasoning based approach to high impedance fault detection in power distribution systems. *IET Generation, Transmission & Distribution*, *11*(5), 1325-1336.

**Soheili, A.; J. Sadeh; H .Lomei and K. Muttaqi. (2016).** A new high impedance fault detection scheme: Fourier based approach. 2016 IEEE International Conference on Power System Technology (POWERCON),

**Varghese, P.; M. Subathra; S. T. George; N. M. Kumar; E. S. Suviseshamuthu and S. Deb. (2023).** Application of signal processing techniques and intelligent classifiers for high-impedance fault detection in ensuring the reliable operation of power distribution systems. *Frontiers in Energy Research*, *11*, 1114230.

**Vieira, F. L.; P. H. Santos; J. M. Carvalho Filho; R. C. Leborgne and M. P. Leite. (2019).** A voltage-based approach for series high impedance fault detection and location in distribution systems using smart meters. *Energies*, *12*(15), 3022.

**Vlahinić, S.; D. Franković; B. Juriša and Z. Zbunjak. (2020).** Back up protection scheme for high impedance faults detection in transmission systems based on synchrophasor measurements. *IEEE Transactions on Smart Grid*, *12*(2), 1736-1746.

**Wai, D. C. T. and Yibin, X. (1998).** A novel technique for high impedance fault identification. *IEEE Transactions on Power Delivery*, *13*(3), 738-744.

**Wang, P.; B. Chen; H. Zhou; T. Cuihua and B. Sun. (2017)**. Fault location in resonant grounded network by adaptive control of neutral-to-earth complex impedance. *IEEE Transactions on Power Delivery*, *33*(2), 689-698.

**Yu, D. C. and Khan, S. H. (1994).** An adaptive high and low impedance fault detection method. *IEEE Transactions on Power Delivery*, *9*(4), 1812-1821.

**Zhang, W.; Y. Jing and X. Xiao. (2016).** Model-based general arcing fault detection in medium-voltage distribution lines. *IEEE Transactions on Power Delivery*, *31*(5), 2231-2241.