A New Voltage Based Fault Detection Technique for Distribution Network Connected to Photovoltaic Sources Using Variational Mode Decomposition Integrated Ensemble Bagged Trees Approach

Younis M. Nsaif 1,2, Molla Shahadat Hussain Lipu 3, Aini Hussain 1,4,* Afida Ayob 1,4, Yushaizad Yusof 1 and Muhammad Ammirrul A. M. Zainuri 1,*

1 Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
2 General Company of Electricity Production, Middle Region, Iraqi Ministry of Electricity, Al Rusafa, Baghdad 10045, Iraq
3 Department of Electrical and Electronic Engineering, Green University of Bangladesh, Dhaka 1207, Bangladesh
4 Centre for Automotive Research (CAR), Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
* Correspondence: draini@ukm.edu.my (A.H.); ammirrulatiqi@ukm.edu.my (M.A.A.M.Z.)

Abstract: The increasing integration of renewable sources into distributed networks results in multiple protection challenges that would be insufficient for conventional protection strategies to tackle because of the characteristics and functionality of distributed generation. These challenges include changes in fault current throughout various operating modes, different distribution network topologies, and high-impedance faults. Therefore, the protection and reliability of a photovoltaic distributed network relies heavily on accurate and adequate fault detection. The proposed strategy utilizes the Variational Mode Decomposition (VMD) and ensemble bagged trees method to tackle these problems in distributed networks. Primarily, VMD is used to extract intrinsic mode functions from zero-, positive-, and negative-sequence components of a three-phase voltage signal. Next, the acquired intrinsic mode functions are supplied into the ensemble bagged trees mechanism for detecting fault events in a distributed network. Under both radial and mesh-soft normally open-point (SNOP) topologies, the outcomes are investigated and compared in the customarily connected and the island modes. Compared to four machine learning mechanisms, including linear discriminant, linear support vector mechanism (SVM), cubic SVM and ensemble boosted tree, the ensemble bagged trees mechanism (EBTM) has superior accuracy. Furthermore, the suggested method relies mainly on local variables and has no communication latency requirements. Therefore, fault detection using the proposed strategy is reasonable. The simulation outcomes show that the proposed strategy provides 100 percent accurate symmetrical and asymmetrical fault diagnosis within 1.25 milliseconds. Moreover, this approach accurately identifies high- and low-impedance faults.

Keywords: distributed network; distributed generation; fault detection; protection scheme; protection strategy; high impedance fault; soft normally open point; variational mode decomposition; ensemble bagged trees method

1. Introduction

Distributed generation is increasingly being incorporated into the distribution grid because it improves the distribution network’s effectiveness, stability, and reliability. Distribution generation generates electricity by combining numerous small-scale sustainable energy sources [1]. Distribution generation enables customers to develop their own
energy and export surplus electricity into the distribution network (DN) [2]. In 2050, solar panels are predicted to provide around 20% of the total energy production. In terms of global installed capacity, photovoltaic generation ranks third among renewable energy sources [3]. Therefore, photovoltaic-based distribution generation injection into the electricity grid is an up-and-coming sustainable energy solution.

Additionally, distribution generation provides fault current dependent on generator size, type, location, and DN configuration [4]. The operational conditions of distribution grids have become more complex because of their flexible operation and quick response to consumer demand [5]. DN supports both operation modes: the typically connected and island. Island mode (ISM) contributes to the malfunctioning of DN protection [6]; due to the operating mode change in DNs, changes in the direction of the current and magnitude offer significant challenges to the fault diagnosis process. Additionally, inverter-interfaced distributed generators are particularly vulnerable to voltage fluctuations. In industrial systems, voltage dips are involved in 92 to 98 percent of all malfunctions in the DN [7]. Consequently, developing voltage-based protection technique operating in different operating modes are important for the safe operation of DN.

Identifying faults involving high impedance is a major challenge for distribution network protection designers. High-impedance fault (HIF) occurs when damaged electrical wires connect with high-impedance materials, such as concrete-base, grass-land, sand, or gravel [8]. Since the HIF has a relatively limited fault current compared to a low-impedance fault (LIF) [9], fuse, over-current relay, and distance relay are inadequate for fault detection. Energized conductors are accompanied by an arc when HIF occurs, even though the low HIF current will not destroy DN devices. Nevertheless, the arc poses a critical risk to human life and electrical equipment because it can cause fire and electrical shock [10]. To prevent network damages, DN fault diagnosis is vital. Due to these aspects, HIF needs to be identified in DNs efficiently and accurately.

SNOP could be used to improve DN reliability in network malfunction and transient disturbance [11]. Additionally, SNOP can maintain voltage regulation, reactive power compensation, and active-power flow management under normal operating conditions [12]. Furthermore, SNOP offers reliability by distributing power among adjacent feeders, and it delivers advantages for mesh networks [13]. The transition from a radial DN topology to a mesh SNOP can impact the fault detection process during the implementation of SNOP.

Recently, several methods for detecting faults in DNs have been developed. In such an effort, Y. Bansal and R. Sodhi used a Tellegen theorem and phasor measurement unit to diagnose the fault in DN [14]. However, the communication channel is the basis of this technology. In the meantime, Mishra and Rout proposed a differential protection mechanism based on the Hilbert-Huang transform and machine learning techniques (MLTs) [15]. The Hilbert-Huang transform involves feature extractions of the current signals and zero sequences. Mishra and Rout also found that the extreme learning technique outperformed other machine learning techniques such as the SVM and the Native Bayes classifier. In [16], Chaitanya et al. proposed a differential protection strategy by employing VMD and Hilbert transform. This technique requires a predefined threshold value to identify defects.

In contrast, the authors in [17] presented a protection relay that depends on voltage measurement to protect DN. The suggested relay algorithm utilises power and voltage sensitivity computations to diagnose faults in specified protective zones. However, a communication link is required for the techniques presented in [15–17]. Using the negative-sequence component of the impedance angle, Dubey and Jena [18] in 2020 suggested a differential protection scheme. This method does not validate symmetrical fault diagnosis and ISM. In addition to the cost, the precision of differential protection-based strategies is affected by loads in DN. Differential protection strategies fail when communication channel failure occurs. Consequently, the suggested method utilises only local data.

Many research papers have established reliable fault detection technologies. A strategy depending on wavelet-singular entropy and a fuzzy-inference system was presented by [19] to detect faults. This work extracted the positive-sequence component and three-
phase current signal features using wavelet singular entropy. These signal features represented the inputs for the fuzzy-inference system. Subsequently, the indexes of a fuzzy inference system are determined by using fuzzy sets and fuzzy rules. In order to identify and categorize faults in normal connected mode (NCM), the indices are transformed into perceptual variables. Based on wavelet-transform and SVM, Ahmadipour et al. introduced a fault detection method [3]. Wavelet transforms are used to identify prominent features in a voltage signal. Next, these prominent features are used for both training and testing the SVM to classify and detect faults. Nevertheless, this strategy did not consider the effects of three-phase faults, ISM, and HIF. On the contrary, Srinivasa Rao et al. presented a mechanism for fault diagnostics using genetic algorithm and neural networks [21]. The genetic algorithm method is employed to select the best features and the neural network is used for fault detection. The effects of ISM were not taken into account by both techniques mentioned in [20,21].

Several strategies for identifying HIFs in a distribution grid have been developed recently. The fuzzy logic methodology was used by Vyshnavi and Prasad to identify HIF in DN Prasad [9]. However, this strategy did not consider the impact of ISM and distribution generations. For HIF detection in smart grids, researchers in [22] implement a wavelet transform in combination with an extreme-learning machine. This protection strategy depends on extracting high-frequency components from three-phase current signals on both ends of the power line, which needs a highly dependable communication link. Whilst, Roy and Debnath presented a protection strategy in order to detect HIFs based on the wavelet transform for evaluating the power-spectral density [23]. Nonetheless, the threshold value is critical to the performance of this strategy. The detection time is affected by the threshold values. In the meantime, Manohar et al. suggested the least squares-Adaline technique and improved SVM to identify and categorize HIF in medium-voltage DN [24]. This strategy did not take into account the influence of SNOP operation. Xiao et al. proposed a neural network and decision tree approach for detecting HIF by employing the transient zero-sequence component of the current signal [25]. This strategy does not determine the effects of ISM and SNOP operation. Forouzesh et al. employed SVM to detect the faulty line in mesh DN using inter-harmonic injection [26]. However, this technique is verified only on ISM. Recently, the improved Hilbert-Huang transform and the ensemble bagged-trees approach were proposed by Nsaif et al. to detect and classify faults in DN [27]. This approach depends on a three-phase current signal. Their findings suggested that HIF can be efficiently and precisely identified by employing sophisticated algorithms.

A comparison between various existing fault detection approaches is tabulated in Table 1 in which there are a few existing approaches that are reasonable for DN. Four approaches are shown to require an extensive communication link. Only one technique is capable of detecting all types of faults (i.e., single-line to ground, double-line, double-line to ground, three-line, and three-line to ground). In addition, only three methods are capable of detecting the HIF. Furthermore, only three approaches can operate in both NCM and ISM. Several approaches do not consider the effect of the mesh topology of DN.
Table 1. Comparison between various existing fault detection approaches.

| Refs. | Existing Approaches for Fault Detection | Contribution/Key Findings | Research Gaps |
|-------|----------------------------------------|----------------------------|---------------|
| [17]  | Active power differential and voltage sensitivity computations | Detect and identify fault in both operation mode | Double-line to ground, and three-line to ground faults, required a communication link, and costly |
| [18]  | Negative-sequence component of the impedance angle, differential protection | Detect LIFs and HIFs | Three-line, three-line to ground faults, required a communication link, costly, ISM, and symmetrical fault |
| [28]  | Time-time transform, deep belief network, phasor measurement unit, and global positioning system | Diagnose the fault accurately and quickly. | Required a communication link, and costly |
| [23]  | Wavelet transform, power spectral density and a threshold value | Detect and identify HIFs | Three-line to ground fault, mesh topology of DN, and ISM |
| [29]  | CIGRE benchmark parameters and communication link | Detect LIFs | Three-line fault, mesh topology of DN, HIF, required a communication link, and costly |
| [30]  | Differential zero sequence component of current signal, and Kalman filter | Identify faulty feeder (single-line to ground) | Mesh topology of DN, and ISM |
| [25]  | Transient zero-sequence component, neural network (one dimensional variational prototyping encoder), and decision tree algorithm | Identify the HIFs (single-line to ground) and non-HIFs | ISM, and Mesh topology of DN |

In this paper, digital signal processing and MLT are presented to identify faults in low-voltage DN. Local measurements of the voltage signal are processed using signal extraction techniques to identify hidden features. MLT is applied to prevent an insufficient predefined threshold value, which has a substantial influence on the detection time and precision. Due to the variational mode decomposition (VMD) technique surpassing empirical-mode decomposition in non-stationary signals mathematically, VMD proposed extracting hidden features. Subsequently, a supervised MLT technique known as the ensemble bagged-trees method (EBTM) was utilised in order to improve detection accuracy. The major contributions of the presented strategy are highlighted as:

- Accurate detection of all types of faults, including symmetrical and asymmetrical faults in DN by utilising only local input variables. The suggested method provides a cost-effective technique for protecting DNs when compared to methods that rely on established communication channels.
- Consideration of the impact of operating modes changing from NCM to ISM. Protection strategies that operate primarily during NCM are inadequate during ISM due to the limited amplitude of the fault current. In contrast, the proposed strategy delivers a high accuracy throughout a diverse range of operation modes.
- Consideration of the influence of high impedance faults on the suggested strategy. Consequently, the HIF has a significantly lower fault current as compared with LIF. Traditional protective strategies are insufficient for identifying HIF. Nevertheless, the proposed strategy successfully identifies LIF, and HIF.
- Consideration of the influence of the mesh-SNOP in the proposed strategy. The proposed technique protects DN effectively during both radial and mesh-SNOP topologies, despite the fact that the deployment of SNOP has the potential to influence the fault detection process by altering the DN topology.
- Development of a new voltage-based protection strategy by using VMD, and EBTM. VMD is useful for addressing non-stationary signals. Moreover, the suggested EBTM technique is contrasted with four conventional machine learning algorithms that are trained and evaluated by employing the same dataset as EBTM. EBTM can identify faults with a high degree of accuracy.

This paper includes five main sections. The methodology is presented in Section 2. Subsequently, a brief summary of the system description and distribution network
models is provided in Section 3. Results and discussion are presented in Section 4. Finally, section 5 illustrates the conclusions.

2. Methodology

In this portion, the presented mechanism is described in detail. The VMD is employed to extract voltage signal characteristics. In addition, fault identification is accomplished by the implementation of EBTM.

2.1. Variational Mode Decomposition

VMD is deployed to extract the features in the time-frequency domain, which has several benefits, including being self–adaptive, not having an impact of mode mixing, and being insensitive to noise. Decomposing multi-component signals, identifying side bands, extracting intra-wave features, and handling with noise robustness are all aspects where the VMD outperforms the empirical-mode decomposition [31]. Consequently, VMD was utilised rather than empirical-mode decomposition to extract signal features for the purpose of this paper.

VMD is a distinct, non-recursive, fully intrinsic, and adaptive signal processing approach that breaks a signal into sub-signals, called intrinsic-mode functions (IMFs) [32]. The VMD algorithm divides a signal \( Y(t) \) into a specific number of IMFs. Every IMF signal has a band-limited bandwidth. Differences in sparsity features are used to distinguish between the modes, while the input signal is generally fully reproduced. The IMFs are constructed in the form of a sinusoidal waveform function.

\[
Y_k(t) = \sum_{k=1}^{K} M_k(t) \cos \theta_k(t)
\]  

The number of modes, denoted by \( K \). \( Y_k(t) \) are IMFs, \( M_k(t) \) are positive envelope of IMFs, and \( \theta_k(t) \) are represent the non-decreasing phase. The reproduction of an original signal by deconstructing modes has a specific sparsity characteristic. Particularly, the preponderance of each mode rotates around the fundamental frequency [31]. In fact, the IMFs are the re-production of varying amplitudes of substantial disruptions in the original signal caused by the fault. The original signal can be reconstructed by combining all of the IMFs. The bandwidth and centre frequency of every IMF can be calculated by continually evolving the best solution to a variational problem using a steady optimization method. The decomposition problem for any signal can be stated as:

\[
L(M, \alpha, \lambda) = \alpha \sum_k \left[ \frac{d}{dt} \left( \delta(t) + \frac{j}{\pi t} \right) * M_k(t) \right] e^{-j \omega_k t} + \left\| f(t) - \sum_k M_k(t) \right\|_2^2
\]

where \( \omega_k \) denotes the central angular frequency, \( \delta \) represents the dirac-distribution, \( * \) represents the convolution, \( j \) denotes the imaginary component with a value \( j^2 = -1 \), \( \alpha \) represents the quadratic penalty-term, and \( \lambda \) denotes the Lagrangian multiplier (LM). Reconstructing the original signal \( f(t) \) is possible by combining all modes. Implementing the alternating direction approach of multiplying, Equation (2) can be successfully minimised. In the frequency response, the relevant updated centre frequency and assessed modes are displayed as follows:
Equations (3) and (4) describe the mathematical strategy for updating the modes and their centre frequencies based on the number of selected modes \( n \), where \( \hat{M}_{k}^{n+1}(w) \) and \( w_{k}^{n+1} \) demonstrate the magnitude and centre frequency of the next mode \( n+1 \) of the \( k \)th level of decomposition. \( \hat{f} \) denotes the Fourier transform of \( f \). \( \epsilon \) denotes an adequate tolerance rate. The LM is updated as,

\[
\hat{\lambda}_{n+1}(w) = \hat{\lambda}_{n}(w) + \tau \left( \hat{f}(w) - \sum_{k} \hat{M}_{k}^{n+1}(w) \right)
\]

where \( \tau \) represents the LM update rate. The iteration number is raised, and the modes are constantly being updated by the algorithm, centre frequencies, and LM until the convergence is attained. In this article, IMF mode-3 was employed because it fulfilled the design criteria.

2.2. Sliding-Window Mechanism

Before extracting fault features, one must preserve the real-time characteristics of the signals via a sliding window analysis technique. Therefore, VMD needs to specify the quantity of signal length to guarantee an accurate IMF output. Although time-consuming, the sliding window analysis technique reduces the negative effects of data processing, which minimises the impact induced by the data processing. With this process, a fixed-length sample can be made by sliding a fixed-length window over a specific time. The selection is continuously updated throughout the windowing process. This way, the VMD data can be accurately acquired and features extracted [5]. As a consequence, the data for VMD can be obtained quickly. Moreover, the functionality of the windowing mechanism can be enhanced by selecting the appropriate sliding window dimensions.

2.3. Ensemble Bagged-Trees Method

In general, the MLT can be classified into three types: supervised, semi-supervised, and unsupervised learning. Supervised machine learning is one of the most commonly used classification techniques. In the training function, training errors are used to achieve classification capability. This closed-loop feedback has the potential to increase MLT classification accuracy [33]. Thus, this project used supervised MLTs as the base learner for the EBTM.

The ensemble learning strategy requires three primary phases for implementation. The first phase is adjusting training datasets and developing models using various learning methods. The second phase is selecting the members, which only involves choosing models that can make predictions. In the third phase, defined as the member combining phase, the output from multiple classifiers is aggregated into one final prediction. Furthermore, three stages are required in the task, and every step requires numerous classifiers. Firstly, various perspectives are explored to integrate classifiers. Secondly, cooperating classifiers are integrated by utilising one or multiple perspectives. Thirdly, the selection of classifiers based on several criteria including the deployment of basic ensemble approaches. By combining the results of several classifiers, designers utilise several basic
ensemble approaches, such as the average, majority-voting, weighted-average, and weighted majority voting, to make a final accurate prediction. The three basic types of ensemble learning strategies for deploying machine learning classifiers are bagging, boosting, and random subspace [34]. The EBTM was proposed to address the classification problem in this project.

Bagging is a statistical method commonly known as bootstrap aggregation. There are two essential advantages of utilising bagging. First, by developing multiple classifiers with a fixed bias and averaging the outcomes, variance is dismantled, and model overfitting is minimised. This method is extremely effective when the input characteristics have a considerable variation and a minimal bias level. Second, bagging produces multiple bootstrap sets from the training data, trains the data with a classifier, and then combines the outputs of each model with a convenient method, e.g., majority voting [35]. The EBTM algorithm is conducted using the procedures mentioned in Algorithm 1.

Algorithm 1. Ensemble bagged trees Algorithm.

1. **Input**: Data Set \( TR = \{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\} \);
2. Base learning algorithm \( L \);
3. Number of learning rounds \( R \).
4. **Process**: for \( r = 1,\ldots,R \);
   5. \( TR_r = Bootstrap(TR) \);
   6. \( h_r = L(TR) \);
   7. end
8. **Outputs**: \( H(X) = \arg \max_{y \in \mathcal{Y}} \sum_{r=1}^{R} I(y = h_r(x)) \)

As shown in Algorithm 1, the algorithm starts by preparing the input using data set \( TR \), specifying the base learning algorithm \( L \) and the number of learning rounds \( R \). Then, the process phase involves generating a bootstrap sample from the data set and training a base learner for iteration from 1 to \( R \). Finally, the output can be determined by applying the \( \text{argmax} \) function where the value of \( l(a) \) is ‘1’ if it is true and ‘0’ otherwise.

2.4. Proposed Method for Fault Detection

The proposed fault detection technique using VMD and EBTM is demonstrated as in Figure 1. The sequence analyser is utilised to compute zero-, positive-, and negative-sequence components of a three-phase voltage signal. Next, IMFs (mode-3) are extracted using VMD to detect the fault. Hence, the signal analyser in MATLAB’s signal-processing toolbox was used to identify the IMF mode. Finally, all IMF (mode-3) are employed for fault detection in DN using EBTM. Figure 2 depicts the overall flow chart of the proposed technique for identifying faults in DN. The proposed method consists of four phases. In the first phase, the three-phase voltage signal is used to compute the zero-, positive-, and negative-sequence components. After manually selecting a sliding window size, the sliding window mechanism is applied on zero-, positive-, and negative-sequence components in the second phase. Next, VMD is employed to extract IMF (mode-3) in the feature extraction phase. Lastly, an MLT is used to distinguish between healthy and faulty DN conditions in the detection phase.
Three phase voltage

Detect

Zero-sequence component
Positive-sequence component
Negative-sequence component

IMF(mode-3)
IMF(mode-3)
IMF(mode-3)

Ensemble Bagged-Trees Method
Detect fault

Figure 1. Proposed fault detection structure.

First phase

Input three phase voltage signal

Compute zero-, positive-, and negative-sequence components

Sliding window size selection and execution

Second phase

Feature extraction phase

Extract IMF(mode-3) using VMD

Detection phase

Fault detection using Machine learning technique

DN condition: healthy, or faulty

Figure 2. The overall flow chart of the proposed technique.

2.5. Machine Learning Techniques Performance in Fault Detection

This section illustrates how the performance of an artificial intelligence (AI) technique is evaluated using VMD based on IMF mode-3. This framework examines and compares five AI techniques for fault detection: linear discriminant, linear SVM, cubic SVM, ensemble boosted tree and EBTM.

Indeed, accuracy is a statistical criterion employed to evaluate the efficiency of MLTs. Accuracy evaluates the reliability between expected and normal situations for both healthy and faulty events concurrently. It can be determined utilising Equation (6).

\[
\text{Accuracy} = \frac{\sum (\hat{X} + \hat{X})}{\sum (X + \bar{X})}
\]

(6)

where \( \hat{X} \) and \( \hat{X} \) are the predicted faulty and healthy events. Moreover, \( X \) and \( \bar{X} \) denote the actual faulty and healthy events, respectively. In addition to accuracy, the
suggested approach has been evaluated on various metrics, including recall, precision, and F1-score.

Recall: recall standard for classifying fault events is the percentage of fault events that the model classifies correctly \((T_p)\) to the total number of classification events in the testing set [36]. The total of classification fault events that have been classified incorrectly are considered false negatives \((F_n)\). When the recall metric has a high value, it means that only a small amount of data has been assigned to the inaccurate class. The recall is computed by using

\[
\text{recall} = \frac{T_p}{T_p + F_n}
\]  

(7)

Precision: this criterion, computed by Equation (8), measures the proportion of correctly classified events related to the sum of \(T_p\) and the total number of inaccurate fault events assigned to the class \(F_p\). A high accuracy criterion value denotes that there is minimal misclassification data in a certain class.

\[
\text{precision} = \frac{T_p}{T_p + F_p}
\]  

(8)

F1-score: This criterion can be achieved by using the recall and precision criterion. This score is found by using Equation (9). The F1-score provides a more feasible depiction of how the classifier model works on all classes in the data set compared to the accuracy criterion.

\[
F_1 = 2 \cdot \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = \frac{T_p}{T_p + 0.5(F_n + F_p)}
\]  

(9)

To ensure that all available data is used effectively, a re-substitution validation method is employed throughout the training and testing of all MLTs in this article. Furthermore, the training data consisted of 408 different instances of dynamic operating modes that were trained in MATLAB using supervised machine learning in classification learner application. The NCM and IsM modes of operation are both covered in the training data. Furthermore, all possible fault types, including single-line, double-line, double-to-ground, and three-line faults, were taken into account. Additionally, the HIFs and LIFs were investigated along all fault types in both modes of operation. The fault impedance is specified in 0.01, 10, 80, 100, 500, and 1000 \(\Omega\) for LIF and HIF, respectively. The fault is developed at 0.2 s, near bus-5, as depicted in Figure 3. Additionally, the condition of SNOP on the distribution grid and its effect on the suggested protection process is considered.
3. System Description and Distribution Network Models

A low-voltage distribution network with 11, and 0.4 kV voltage level is simulated and modelled in MATLAB Simulink environment, as shown in Figure 3. The distribution network is supplied by two inverters-interfaced distribution generations. Furthermore, the inverters-interfaced distribution generations are linked to the distribution network through bus six, and seven, respectively. The first step-down transformer 11/0.4 kV is coupled with the network between bus five and six. The second step-down transformer 11/0.4 kV is linked between bus four and seven.

The distribution grid is comprised of two photo-voltaic panels, two battery banks, two inverters, two step-up transformers, and five distributed loads. As depicted in Figure 3, the circuit breaker is utilised to transit a distribution grid topology from radial to mesh-SNOP. Table 2 tabulates the distributed grid component specifications.

Table 2. The distributed grid component specifications.

| No. | Component                          | Specification                                                                 |
|-----|------------------------------------|-------------------------------------------------------------------------------|
| 1   | Photo-voltaic panels (11 and 21)   | 2 Parallel strings, 28 series-connected modules per string, irradiance 1000 W/m² |
|     |                                    | 12.5 kVA, 0.4 kV, 50 Hz, switching frequency 5 kHz                             |
|     |                                    | Filter:                                                                      |
|     |                                    | Series-Inductance 4.6 mH                                                      |
|     |                                    | Series-Resistance 0.4596 Ω                                                   |
|     |                                    | Shunt-Capacitance 0.1102 µF                                                  |
| 2   | Inverters (1 and 2)                 |                                                                              |
|     |                                    | 12.5 kVA, 0.4 kV, 50 Hz, switching frequency 5 kHz                            |
| 3   | Battery banks (11 and 21)          | Lead-Acid, 980 V, 2.7 Ah                                                     |
| 4   | Transformers (TR1, and TR2)        | 11/0.4 kV, 24 kVA, 50 Hz, D11/Yn                                            |
| 5   | Loads (61 and 71)                  | 5 kW, 0.4 kV, 50 Hz                                                          |
| 6   | Load (21)                          | 10 kW, 11 kV, 50 Hz                                                          |
| 7   | Loads (11, and 12)                 | 8 kW, 11 kV, 50 Hz                                                           |
4. Results and Discussion

As indicated previously, the characteristic was extracted from three phase voltage signals by VMD. VMD can decompose the signal into multiple modes, although a high level of deconstruction is preferred to acquire the maximum number of dependable signal features. Additionally, a higher decomposition level has several negative sides. The higher number of modes increases the computing burden, which may cause in a longer reaction time for the relay. A rise in the relay response time could pose a serious hazard to the electrical grid. Consequently, the number of decomposed modes is limited to five. Unlike other VMD implementations, it is not mandatory to record the whole signal information in this article because VMD was performed within the MATLAB/ Simulink environment. Each signal must be adapted to a moving window-based signal to precisely operate VMD within the MATLAB Simulink environment. In addition to the sampling frequency, the sample size is a vital factor to consider. For the proposed moving window method to work effectively, it is necessary to precisely identify the size of the samples being used. As a consequence, a 500-samples-per-moving-window approach is being utilised to establish a balance between the need for minimal computing time and the requirement for the extraction of comprehensive features based on the investigation. In this study, the VMD algorithm was used to deconstruct five different IMFs, as shown in Figure 4. The sample time interval for the discrete-simulation type is stated as $2.5 \times 10^{-6}$ s. IMFs extracted by VMD technique during the single-line to the ground on phase A in ISM are shown in Figure 4. After a comprehensive investigation of all types of DN faults, IMF (mode-3) was used to extract features from zero-, positive-, and negative-sequence components of the voltage signal. Subsequently, the data is collected for the training process. As demonstrated in Figure 5 a,b, the IMF mode-3 feature of the negative-sequence component was variant from 27.8 to 26.95 during ISM-LIF whilst initiating the SNOP respectively. Likewise, the IMF mode-3 feature of the zero-sequence component changes from 13.24 to 13.17. While the IMF mode-3 feature of the positive-sequence component slightly changes from 0.001 to 0.0005. Figure 5 c,d shows how the IMF mode-3 feature of the negative-sequence component during HIF alters from 27.85 to 26.99 when SNOP is initiated. Similarly, the IMF mode-3 feature of the zero-sequence component varies from 13.35 to 13.28. Thus, the IMF mode-3 feature of sequence component signals during HIF surpasses LIF. Additionally, initiating the SNOP in the distribution grid reduces the IMF mode-3 feature of sequence component signals during both LIF and HIF. The functionality of MLTs was evaluated in three different operation modes: NCM, ISM, and dynamic operating mode. Moreover, there are four different topologies used in the training scenario: radial with NCM, mesh-SNOP with NCM, radial with ISM, and mesh-SNOP with ISM.
Five VMD IMFs were detected by a relay placed on bus-5 during the single-line to ground fault at phase A. (a) IMF obtained from the zero-sequence component of the voltage signal; (b) IMF obtained from the positive-sequence component of the voltage signal; (c) IMF obtained from the negative-sequence component of the voltage signal.
4.1. Normal-Connected Operation Mode

Due to the mismatch in fault current magnitude between NCM and ISM, it is challenging to maintain a protection solution with a constant threshold value. The outcomes of five MLTs for the radial topology with NCM are stated in Table 3. In total, 132 different cases were used in a distribution grid with radial topology to evaluate the protection functionalities of the MLTs. Furthermore, HIF and LIF are initiated in NCM at different fault locations. The findings showed that both linear discriminant and ensemble boosted tree method have an accuracy rate of 69.7%, which is the lowest accuracy of the five MLTs. While the accuracy rate of the linear SVM technique is found to be 78%. However, the cubic SVM performed slightly better and acquired an 83.3% accuracy rate. The best result is obtained from the EBTM classifier with an accuracy of one hundred percent.

Changing NCM-DN topology from radial to mesh-SNOP can influence the performance of five MLTs owing to the imbalanced power distribution in the feeder. The MLTs have been evaluated for distribution grid protection with 132 test cases in a mesh-SNOP topology. Additionally, HIF and LIF are initiated using a mesh-SNOP structure with NCM at different fault locations. Figure 6 depicts a comparative evaluation of the performance of five MLTs used in this research under radial and mesh-SNOP throughout NCM-DN. It has been observed that both linear discriminant and ensemble boosted tree technique provides the lowest rates of accuracy with 69.7%. While the linear SVM, and cubic SVM technique accuracy rates are found to be 78, and 83.3%, correspondingly. Conversely, it has been discovered that EBTM can achieve the highest possible accuracy of 100%.
Figure 6. Comparative evaluation of the performance of five MLTs used in this research under radial and mesh-SNOP throughout NCM-DN.

Table 3. Accuracies acquired by several MLTs.

| MLTs              | Accuracy %          | NCM | Radial | Mesh-SNOP | IsM | Radial | Mesh-SNOP | Dynamic |
|-------------------|---------------------|-----|--------|-----------|-----|--------|-----------|---------|
| Linear Discriminant | 69.7% | 69.7% | 97     | 97.7     | 90.2|
| Linear SVM        | 78     | 78    | 98.5   | 98.5     | 90.2|
| Cubic SVM         | 83.3   | 83.3  | 98.5   | 98.5     | 90.2|
| Ensemble Boosted tree | 69.7% | 69.7% | 69.7   | 69.7     | 100 |
| EBTM              | 100    | 100   | 100    | 100      | 100 |

4.2. Island Operation Mode

Conventional over-current relays could have significant relaying challenges because of the dynamic behaviour of the operating conditions. Table 3 presents the outcomes obtained by five MLTs in IsM with radial topology. In total, 132 cases in a radial topology were used to test the protection capabilities of the MLTs. It has been observed that the ensemble boosted tree method achieved a 69.7 percent accuracy, which is the lowest in comparison to the other four MLTs. The linear discriminant provided an accuracy rate of 97%. In contrast, the accuracy rate of the linear SVM, and cubic SVM techniques are observed to be 98.5%. Our investigative work demonstrated that the proposed EBTM can achieved the ideal accuracy of 100 percent.

Owing to the dynamic behaviour of the operating conditions, traditional over-current relays might have a considerable relaying issue. The performances of five MLTs in IsM are given for mesh-SNOP topology in Table 3. The DN protection abilities of the MLTs have been tested utilizing 132 cases in a mesh-SNOP topology. HIF and LIF are tested in mesh-SNOP topology with IsM in different fault locations. Figure 7 illustrates a comparative evaluation of the performance of five MLTs used in this investigation under radial and mesh-SNOP throughout the IsM-DN. It was discovered that the ensemble boosted tree only attained an accuracy of 69.7 percent, which is the lowest rate of accuracy achieved as compared with the other four MLTs. The linear discriminant provides an accuracy rate of 97.7%. While, the accuracy rate of both the linear SVM, and cubic SVM are found to be 98.5%. In contrast, the proposed EBTM reached the substantial ideal accuracy of 100% accuracy.
4.3. Dynamic Operation Mode

The efficiency of five MLTs is evaluated under the NCM and IsM in this section. In addition, radial and mesh-SNOP topologies are examined to determine the best approach for achieving high accuracy. The distribution grid protection capabilities of the MLTs have been tested over 408 scenarios. Both radial and mesh-SNOP topologies are used to investigate the HIF and LIF in different fault locations. Figure 8 shows the overall accuracy of the five MLTs used in this research throughout the dynamic operation mode in DN. The observation showed that linear discriminant, linear SVM, ensemble boosted tree, and cubic SVM methods have an accuracy rate of 90.2%. According to the results, EBTM has the highest degree of accuracy, achieving a rating of 100%.
Table 3 shows that the EBTM outperformed the linear discriminant, linear SVM, cubic SVM, and ensemble-boosted tree in terms of accuracy. Consequently, the EBTM was used to address the fault diagnosis problem.

The training data include 40 cases of healthy condition. The healthy data include both operating mode, radial topology, mesh-SNOP topology, and load change (among +10%, −10%, and −20%). Additionally, 368 faulty cases are included in the training data. The HIF data is incorporated with single-line to ground, double-line to ground, and three lines to ground defects. Moreover, diverse fault positions (between 0.2 and 0.4 km near bus-5) are also conducted. Figure 9 demonstrates the performance of all MLTs during NCM, IsM, and dynamic operation mode in terms of recall, precision, F1-score, and accuracy. It can be observed that EBTM outperformed the other four MLTs in terms of recall, precision, F1 score, and accuracy. The EBTM are then be implemented in MATLAB Simulink utilizing the classification ensemble-predict block. Lastly, VMD and EBTM techniques were examined in MATLAB Simulink to identify the detection time. Consequently, the proposed technique detected all kinds of faults, including LIF and HIF, within 1.25 milliseconds. Figure 10 demonstrates the output fault signal and detection time using the proposed strategy. The yellow line represents the observed fault signal by the proposed strategy. Once the fault signal reached 1, the trip signal was initiated, and the associated circuit breaker was opened. Moreover, the functionality of this technique is unaffected by the transition from NCM to IsM operating conditions. Furthermore, changing the distribution grid topology from radial to mesh-SNOP has no effect on the performance of this technique. Table 4 presents a comparison of the proposed technique with several existing strategies. The capability to identify HIF can be achieved by utilizing two of the existing techniques in [17,37]. Mesh-SNOP has not been considered by any other approach. It has no effect on the proposed method. The existing techniques require expensive communication link which is inadequate in a low-voltage DN. However, the proposed method uses only local information and does not need a communication link. The proposed technique produces considerably better outcomes in terms of accuracy and detection time when compared to the techniques listed in Table 4.

![Figure 9. The performance of all MLTs during NCM, ISM, dynamic operation mode in terms of recall, precision, F1-score, and accuracy.](image-url)
Figure 10. The output fault signal and detection time using the proposed strategy during single line to ground fault.

Table 4. Comparison of the proposed technique with several existing strategies.

| Ref. | Technique | Parameter Used | HIF | Mesh-SNOP | Operation Mode | Communication Link Required | Accuracy | Detection Time |
|------|-----------|----------------|-----|-----------|----------------|-----------------------------|----------|----------------|
| [14] | Phasor measurement unit and Tellegens theorem | Current and voltage | ×   | ×         | NCM            | Yes                         | 100%     | 20 ms          |
| [38] | Intelligent electronic devices, load flow, and sensitivity computations | Voltage | ×   | ×         | Both           | Yes                         | N/A      | N/A            |
| [3]  | Wavelet transform and SVM | Voltage | ×   | ×         | NCM            | -                           | 100%     | Less than 12 ms |
| [29] | CIGRE benchmark parameters and theoretical fundamentals | Current and voltage | ×   | ×         | Both           | Yes                         | N/A      | N/A            |
| [39] | S-transform | Voltage and current | ×   | ×         | Both           | Yes                         | 99.44%   | N/A            |
| [17] | Active power differential and sensitivity computations differential currents, and S-transform | Voltage | √   | ×         | Both           | Yes                         | 100%     | N/A            |
| [37] | Proposed Technique | Voltage | √   | √         | Both           | Yes                         | In NCM 100% and IsM 97.59% | Less than 17 ms |

Where the symbols √ and × indicate that the aspect is addressed and overlooked, respectively.

5. Conclusions

Increasing DN reliability is accomplished by fast and precise fault detection. Due to the unique characteristics and operations of DN, traditional protection strategies are insufficient to address DN challenges. These challenges include altering fault currents throughout the operation modes, a diversity of DN topology, and HIFs. The suggested method offers a cost-effective solution for operation mode transitions from NCM to IsM, diverse DN topologies (radial and mesh-SNOP), and HIFs. This article proposes a new
voltage-based protection strategy to detect faults quickly and precisely. The presented technique is developed using VMD, and EBTM. VMD is employed to extract prominent features from zero-, positive-, and negative-sequence components of a three-phase voltage signal. Afterward, the IMF mode-3 act as input signal of EBTM to detect fault events. The results indicated that the suggested technique can rapidly and accurately identify any type of fault without employing any form of communication channel. Additionally, the suggested method was evaluated in four different operational conditions (i.e., NCM with radial, NCM with mesh-SNOP, IsM with radial, and IsM with mesh-SNOP). Additionally, this technique can be employed to detect both HIF and LIF. Compared to the conventional machine learning techniques (i.e., linear discriminant, linear SVM, cubic SVM, ensemble boosted tree), the proposed EBTM provides the highest degree of detection accuracy. The detection accuracy of the proposed technique is 100 percent, and its detection time is 1.25 milliseconds. The EBTM is reasonable to implement because it relies on local information and does not have connection latency. Lastly, further research is needed to authenticate the proposed EBTM in real-time within a hardware-in-the-loop environment. Although EBTM has demonstrated excellent outcomes in detecting faults under different scenarios, the training operation can be more effective by considering all possible conditions such as operation mode changes, various topology, low impedance fault, and high impedance fault to avoid failure of the proposed technique.

Author Contributions: Conceptualization, Y.M.N.; methodology, Y.M.N.; software, Y.M.N.; validation, Y.M.N. and M.S.H.L.; writing—original draft preparation, Y.M.N.; writing—review and editing, M.S.H.L., A.H., A.A., Y.Y. and M.A.A.M.Z.; visualization, Y.M.N. and A.H.; supervision, M.S.H.L. and A.H.; project administration, A.H.; funding acquisition, A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the University Kebangsaan Malaysia under Grant Code GGPM-2020-006.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

| Abbreviation | Description |
|--------------|-------------|
| AI           | Artificial intelligence |
| DN           | Distribution network |
| EBTM-Ensemble bagged trees method |
| HIF          | High Impedance Fault |
| IMFs         | Intrinsic Mode Functions |
| IsM          | Island Mode |
| LIF          | Low Impedance Fault |
| MLTs         | Machine Learning Techniques |
| NCM          | Normal Connected Mode |
| SNOP         | Soft Normally Open Point |
| SVM          | Support Vector Mechanism |
| VMD          | Variational Mode Decomposition |

References

1. Nsaif, Y.M.; Hossain Lipu, M.S.; Ayob, A.; Yusof, Y.; Hussain, A. Fault Detection and Protection Schemes for Distributed Generation Integrated to Distribution Network: Challenges and Suggestions. *IEEE Access* **2021**, *9*, 142693–142717.
2. Rocha Junior, E.B.; Batista, O.E.; Simonetti, D.S.L. Differential Analysis of Fault Currents in a Power Distribution Feeder Using ABC, Aβ0, and Dq0 Reference Frames. *Energies* **2022**, *15*, 526. https://doi.org/10.3390/en15020526.
3. Ahmadipour, M.; Hizam, H.; Othman, M.L.; Mohd Radzi, M.A.; Chireh, N. A Fast Fault Identification in a Grid-Connected Photovoltaic System Using Wavelet Multi-Resolution Singular Spectrum Entropy and Support Vector Machine. *Energies* **2019**, *12*, 2508.
4. Mansouri, N.; Nasiri, S.; Lashab, A.; Guerrero, J.M.; Cherif, A. Innovative Grid-Connected Photovoltaic Systems Control Based on Complex-Vector-Filter. *Energies* **2022**, *15*, 6772. https://doi.org/10.3390/EN15186772.
5. Li, Y.; Lin, J.; Niu, G.; Wu, M.; Wei, X. A Hilbert–Huang Transform-Based Adaptive Fault Detection and Classification Method for Microgrids. *Energies* **2021**, *14*, 5040. https://doi.org/10.3390/en14165040.
32. Sharma, N.K.; Samantaray, S.R.; Bhende, C.N. VMD-Enabled Current-Based Fast Fault Detection Scheme for DC Microgrid. *IEEE Syst. J.* 2021, 16, 933–944. https://doi.org/10.1109/JSYST.2021.3057334.

33. Le, V.; Yao, X.; Miller, C.; Tsao, B.H. Series DC Arc Fault Detection Based on Ensemble Machine Learning. *IEEE Trans. Power Electron.* 2020, 35, 7826–7839. https://doi.org/10.1109/TPEL.2020.2969561.

34. Dhibi, K.; Mansouri, M.; Bouzrara, K.; Nounou, H.; Nounou, M. An Enhanced Ensemble Learning-Based Fault Detection and Diagnosis for Grid-Connected PV Systems. *IEEE Access* 2021, 9, 155622–155633. https://doi.org/10.1109/ACCESS.2021.3128749.

35. Zararsiz, G.; Akyildiz, H.Y.; Göksülük, D.; Korkmaz, S.; Öztürk, A. Statistical Learning Approaches in Diagnosing Patients with Nontraumatic Acute Abdomen. *Turkish J. Electr. Eng. Comput. Sci.* 2016, 24, 3685–3697. https://doi.org/10.3906/elk-1501-181.

36. Shadi, M.R.; Ameli, M.T.; Azad, S. A Real-Time Hierarchical Framework for Fault Detection, Classification, and Location in Power Systems Using PMUs Data and Deep Learning. *Int. J. Electr. Power Energy Syst.* 2022, 134, 107399. https://doi.org/10.1016/j.ijepes.2021.107399.

37. Langarizadeh, A.; Hasheminejad, S. A New Differential Algorithm Based on S-Transform for the Micro-Grid Protection. *Electr. Power Syst. Res.* 2022, 202, 107590. https://doi.org/10.1016/j.epsr.2021.107590.

38. Manditereza, P.T.; Bansal, R.C. Fault Detection and Location Algorithm for DG-integrated Distribution Systems. *J. Eng.* 2018, 2018, 1286–1290. https://doi.org/10.1049/joe.2018.0247.

39. Eslami, R.; Sadeghi, S.H.H.; Askarian Abyaneh, H. A Probabilistic Approach for the Evaluation of Fault Detection Schemes in Microgrids. *Eng. Technol. Appl. Sci. Res.* 2017, 7, 1967–1973.