ORIGINAL PAPER

LSTM-based low-impedance fault and high-impedance fault detection and classification

Maanvi Bhatnagar[1] · Anamika Yadav[1] · Aleena Swetapadma[2] · Almoataz Y. Abdelaziz[3]

Abstract

In this article, a long short-term memory based protection scheme for power transmission lines is presented. A fault detection framework is developed that uses voltage and current signals' RMS values as input. The proposed work is established for various shunt faults, both low-impedance and high-impedance faults which are tested on a standard IEEE 14 bus system and an existing real transmission network. Results confirm the detection and classification of faults with accuracy and precision higher than 99%. The impact of non-faulty events such as load switching, capacitor switching, noisy data, and load variation conditions is also studied to analyze the model's performance. The efficacy of the proposed method is confirmed by comparing it with different methods in the literature. Results indicate the aptness of the proposed scheme for the protection of power transmission lines.

Keywords LSTM · Shunt fault · High-impedance fault · Evolving fault · Fault detection · Fault classification · Power transmission lines

Abbreviations

LLLG	Three lines to ground fault
LLL	Three lines fault
LLG	Double line to ground
LG	Line to ground
LL	Line-to-line
HIF	High-impedance faults
LIF	Low-impedance faults
AI	Artificial intelligence

FFT	Fast Fourier transforms
WT	Wavelet transforms
ST	S-transforms
ANN	Artificial neural network
SVM	Support vector machines
k-NN	K-Nearest neighbors
CNNs	Convolution neural networks
KL	Kullback–Leibler divergence
DT	Decision tree
RF	Random forest
IED	Intelligent electronic device
PMU	Phasor measurement unit
DFRs	Digital fault recorders
LSTM	Long short-term memory
RNN	Recurrent neural network
RMS	Root mean square
FD	Fault detection
FC	Fault classification
BN	Batch normalization
TP	True positives
FP	False positives
TN	True negatives
FN	False negatives
CM	Confusion matrix
SNR	Signal–noise ratio

✉ Anamika Yadav
ayadav.ele@nitrr.ac.in

Maanvi Bhatnagar
maanvib93@gmail.com

Aleena Swetapadma
aleena.swetapadma@gmail.com

Almoataz Y. Abdelaziz
almoataz.abdelaziz@fue.edu.eg

[1] Department of Electrical Engineering, National Institute of Technology Raipur, G. E. Road, Raipur, CG 492010, India

[2] School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar 751024, India

[3] Faculty of Engineering & Technology, Future University in Egypt, Cairo 11835, Egypt

PSB   Power swing blocking
DCTL   Dual-circuit transmission lines
TCTL   Triple circuit transmission line

1 Introduction

Modern power system network has experienced a rapid increase in size and complexity with the integration of renewable energy resources [1]. As the system is exposed to different atmospheric conditions, the chances of faults are greater. A fault is a condition in which abnormal electric current flows through the system. The protection system's main functions are to detect the fault, and classify them into different types (LG, LL, LLLG, LLL, and LLLG) and to identify the faulty phase(s). There are however frequent and unavoidable faults due to a variety of random reasons, which severely affect the performance of the power system, interrupt the supply of energy, and compromise the efficiency and reliability of the system [2]. To meet the increasing energy demand and ensure continuous delivery, the impact of faults must be minimized. It is therefore crucial to detect these faults early to eradicate them as quickly as possible. Maintaining the faulted component allows for faster recovery of the main system function and makes components more reliable by restoring their reliability in a matter of minutes. To mitigate these faults and restore proper system operation, we need effective protection and maintenance schemes. There are several types of short-circuit faults that may produce short circuits in generation, transmission, and distribution systems, including generators, transformers, insulation, HVDC converters, feeder buses, and transmission lines. Electric short circuits adversely affect power system performance and threaten the key function of the system. Many of the faults occur on transmission and distribution lines, as mentioned in [3]. A short-circuit fault occurs commonly and is the most hazardous type of fault, posing high risks for the line, including reduced component life expectancy, increased power loss and heat, and damaged insulators. The short circuits encountered can be divided into symmetrical and asymmetrical faults. Symmetrical or balanced fault keeps the system balanced. It consists of three lines to ground (LLLG) and three lines (LLL) faults. They are relatively rare, but they cause the major harm to the system equipment because of the higher fault current magnitude involved. Double line-to-ground (LLG), line-to-ground (LG), and line-to-line (LL) faults are asymmetrical and unbalanced faults. The occurrence probability of single-line ground faults is 0.80, although less severe than balanced faults. Another type of fault showcasing a large impedance thereby leading to very small fault currents is known as high-impedance fault

(HIF). In addition to endangering power system equipment, HIFs can also endanger human safety. Designing a proper protection scheme for HIFs is not just a matter of ensuring detection of this hazardous fault but also to avoid any fatal accident which may happen due to fact that it involves current carrying conductor lying on ground with some high-impedance surfaces which can cause fatal damage to human beings. Further it also involves addressing the fact that HIF features depend on many aspects such as surface humidity, soil type, and weather condition [4]. HIF current and voltage waveforms possess random, asymmetric, nonlinear features, making it imperative to develop a robust detection scheme using effective signal processing methods.

The type, location, and duration of a fault affect the system's performance. By detecting, identifying, and localizing faults faster, system protection and upkeep strategies are improved and power systems are maintained with high quality and quantity. Researchers and practitioners study fault detection and classification extensively. Currently, most methods rely on digital sampling of voltage and/or current signals. Detection and classification tasks are then performed on the sampled data. As part of the processing phase, features are extracted and applied to classify and detect faults [2]. A data-driven approach to fault detection and classification will replace rule-based algorithms with machine learning and other artificial intelligence (AI) tools soon. Compared to rule-based algorithms, data-driven models are more effective and can help develop generic solutions. With state-of-the-art AI techniques, features are not required to be extracted using the proposed scheme.

Detecting and classifying LIFs usually involve two steps: (a) identifying features from input signals, and (b) calculating results based on features with the help of various AI tools. Several techniques are used in conjunction with these two parts. These include fast Fourier transforms (FFT) [5], wavelet transforms (WT) [6], S-Transforms (ST) [7], and other statistical methods. A feature extraction process requires repeated efforts and is usually specific to system configurations. For example, multiresolution wavelet analysis and statistical features-based techniques in conjugation with an artificial neural network (ANN) have been utilized [8]. This method is remarkably accurate but is susceptible to higher fault impedances and has not been tested for noise. In feature extraction, WT or DWT is commonly used algorithms [6]. Their applications, nevertheless, do not separate the feature extraction from working algorithms. ST can reveal joint time–frequency characteristics; some researchers have used it instead of WT [7]. ST generally has a better ability to reveal harmonics than DWT and is less susceptible to noise than DWT. However, the criteria for selecting features are alike, typically established on the standard deviation or signal energy. The method of feature extraction may be limited

by the lack of generalization, as well as the need to investigate constraints not directly recommended in the literature. For instance, using WT or DWT encompasses the choice of appropriate mother wavelet levels and decomposition levels which is not only computationally affluent but can affect algorithm performance as well [6]. Next, the features must be processed to detect and classify faults. To accomplish this, numerous researchers have used a variety of methods. Due to the capability to learn from patterns, neural networks (NNs) have attracted a lot of attention [9]. Data-driven tools like decision trees (DT) [10] and k-nearest neighbors (k-NN) [11] are also useful for fault analysis and decision making. Literature also reports on contemporary deep learning methods such as convolution neural networks (CNNs) [2]. However, these techniques are quite inefficient, because they rely on two distinct parts to extract features and work the algorithm. Feature extraction should be eliminated, and operational data should be worked on directly. There is no generalized set of rules for the process of feature extraction. Owing to this unpredictability, the process of feature extraction could be time-consuming and can disturb the system's performance.

HIF detection and classification can also be achieved by different signal processing techniques as discussed in [4, 12]. The HIF detection schemes can be categorized according to the domain for feature extraction, such as the time domain or the frequency/time–frequency domain. Using time-domain methods, HIF is typically detected by measuring voltages and currents and analyzing their unique properties. Mathematical morphology technology [13], fractal geometry techniques [14], and Kullback–Leibler (KL) divergence [15] are some of the time-domain-based feature extraction techniques available for HIF detection. HIF voltage and current signals are analyzed by frequency-domain methods. FFT technology detects HIF in [16] by calculating distance changes between the harmonic components of fault currents.

The consideration of low-impedance faults along with HIF (high-impedance faults) is indeed a crucial aspect that merits attention in this paper. As the system may be subjected to both kinds of faults, a protection scheme must be able to detect the presence of both kinds of faults then only the protection scheme is said to be reliable and robust. Thus, the proposed scheme which is designed to cater both kinds of faults provides reliability and robustness against both LIF and HIF. While HIFs are extensively studied for their damaging effects, the inclusion of low-impedance faults is equally important for a comprehensive understanding of the overall fault landscape. Low-impedance faults, often characterized by reduced resistance, can pose distinct challenges and have different implications compared to their high-impedance counterparts. They may trigger different protective responses in the system and exhibit unique fault signatures. By incorporating an analysis of both LIF and HIF, the paper provides a more holistic view of fault scenarios, leading to a more effective and versatile protective framework. Addressing LIF and HIF in this paper also contributes to a more practical and applicable model, ensuring that the fault detection system is well equipped to handle a wide range of potential issues in real-world electrical systems. Data sources for the modern power grid include intelligent electronic devices (IEDs), phasor measurement units (PMUs), digital fault recorders (DFRs), along with many other devices [17]. According to the literature review, there is a requirement to develop automated fault detection and classification methods whose parameters are flexible in terms of working conditions and data sources. With the help of long short-term memory (LSTM) units [18], this paper presents an innovative technique for automating feature extraction that avoids the need for distinct feature extraction tasks and merges it with the working technique. Learning process parts such as feature extraction and working algorithms are unified, which makes deployment more feasible. The major rewards offered by the method are that it does not need communication links, requires low sampling frequency, is easy to implement, does not get affected by system noise, unavoidable transients, or operating conditions, and is overall robust and reliable for operation. The paper is organized as follows—Sect. 2 describes the LSTM methods used, Sect. 3 contains the proposed method, Sects. 4,5, and 6 cover the results, Sect. 8 contains a comparison with other schemes and is followed by the conclusion of the work in Sect. 9.

2 Long short-term memory method

As a type of artificial neural network, LSTM [18] is a component of artificial intelligence and deep learning. An LSTM has feedback connections, unlike a feed-forward neural network. In addition to processing single data points, this type of recurrent neural network (RNN) can also process whole sequences of data. LSTM has both long-term and short-term memories, like a standard RNN. In the network, weights and biases are altered in each iteration, just as synaptic strength changes physiologically to accumulate long-term memories; activation patterns in the network alter once per time step, similarly to how short-term memories are stored in the brain through moment-to-moment electrical firing patterns. By providing long short-term memory to RNNs, the LSTM architecture offers long short-term memory lasting thousands of time steps. Three gates control the flow of information into and out of a common LSTM cell unit. These gates are a forget gate, an input gate, and an output gate [18]. With three gates controlling information flow, the cell can remember values for arbitrary periods. A time series may have significant lags between significant events in the series, so LSTM networks are appropriate to classify, process, and prediction of events.
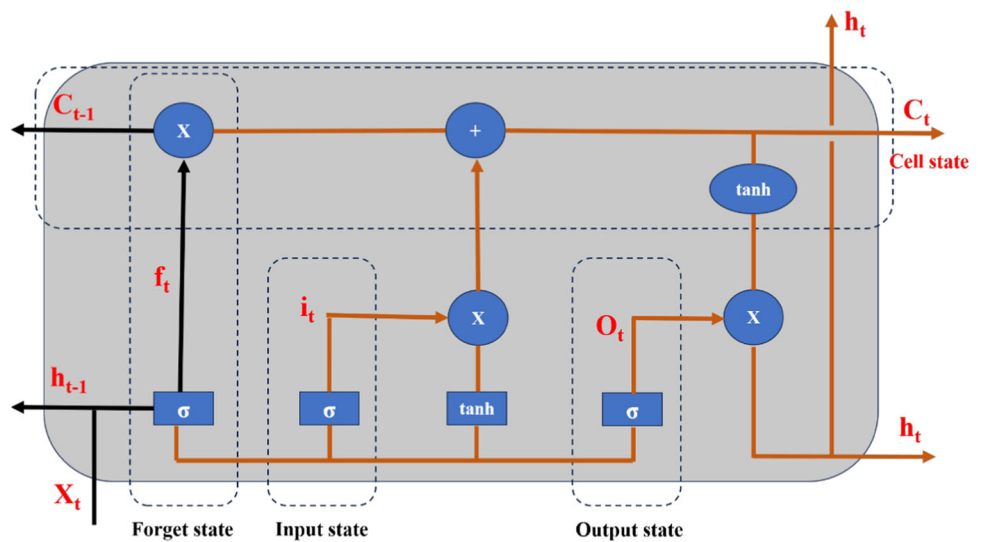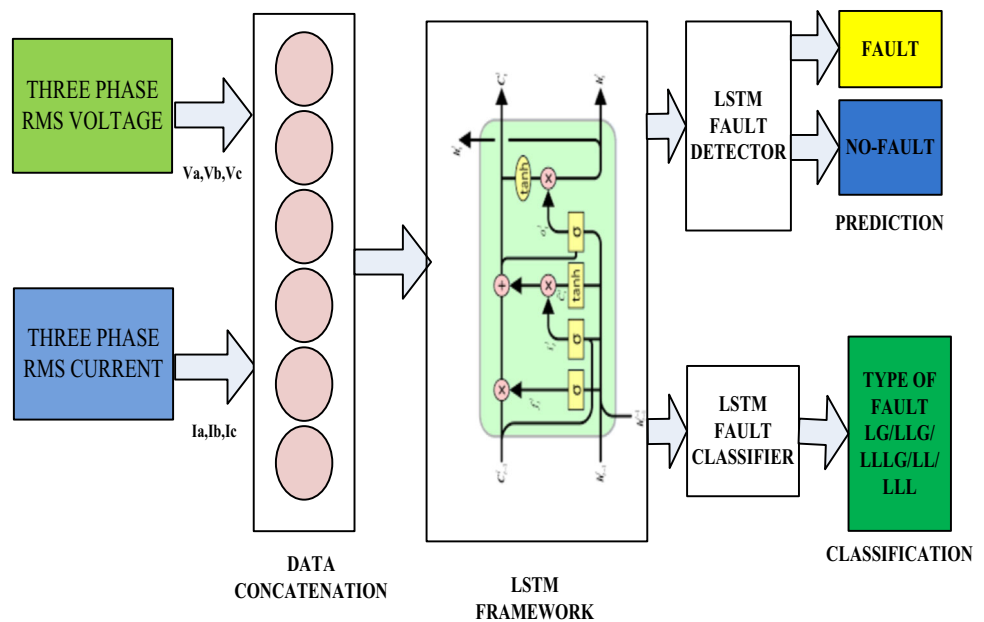
Fig. 1 LSTM Architecture

Fig. 2 LSTM architecture for the proposed scheme

To overcome the problem of vanishing gradients encountered in training conventional RNNs, LSTMs were developed. In various applications, LSTMs are advantageous over RNNs and other sequence learning approaches. This is because they are relatively insensitive to gap length. The basic block diagram of a conventional LSTM unit is illustrated in Fig. 1.

LSTM begins by deciding how much information about the cell state (C_t) it is willing to discard. Sigmoid layers known as "forget gate layers" make these decisions. It decides by looking at h_{t-1} (previous time step hidden state vector) and x_t (input vector to LSTM unit) and returns a value between 0 and 1 for each number in the cell state C_{t-1}. In general, "1" indicates that the information should be kept permanently, while "0" indicates that it should be thrown away

completely. This step is shown mathematically in Eq. (1).

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right) \tag{1}$$

Here f_t is the activation vector of the forget gate, W_f is the corresponding weight matrix, and b_f is the bias function. σ represents the sigmoid activation function. Choosing the new data to store in the cell state is the next step. It consists of two parts. As a first step, a sigmoid layer known as the "input gate layer" decides which values to appraise. Afterward, a $tanh$ layer creates a vector of new candidate values, C'_t (cell input activation vector), which could be added to the state. In the next step, these two are combined to create an update to the
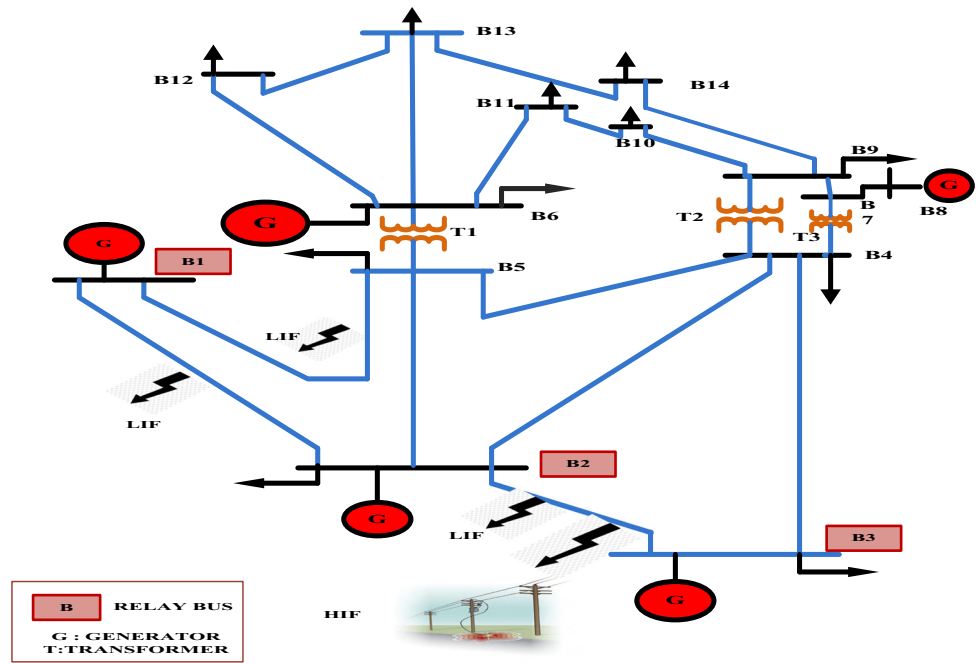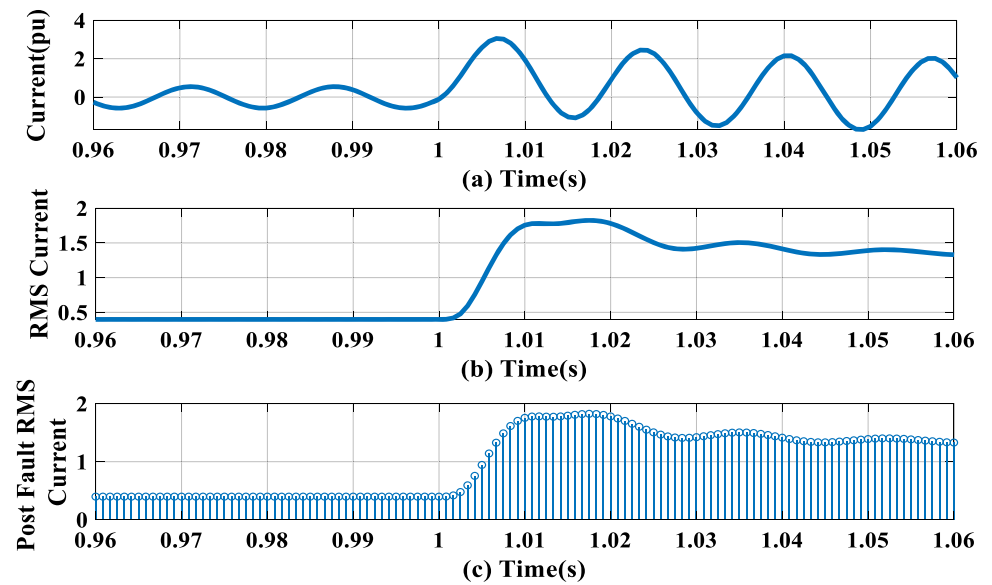
Fig. 3 Single line diagram of IEEE 14 bus system

Fig. 4 **a** Phase A current. **b** RMS current. **c** RMS current of two cycles post-fault

state [24].

$$i_t = \sigma\left(W_i[h_{t-1}, x_t] + b_i\right) \qquad (2)$$

$$C'_t = tanh\left(W_C[h_{t-1}, x_t] + b_c\right) \qquad (3)$$

Here i_t is the activation vector of the input gate, W_i and W_C are the corresponding weight matrix, and b_f and b_c are the bias functions. Next the old cell state (C_{t-1}) is updated to cell state (C_t). In the previous steps what information needs to be added was decided upon. This step involves the execution of the same. The old state is multiplied by f_t which leads to

forgetting the new and maxima of the signal the information that was not needed. Then, $i_t C'_t$ is added to it. The new candidate value is scaled before updating every state value. Mathematically this step is represented by Eq. (4)

$$C_t = \left(f_t C_{t-1} + i_t C'_t\right) \qquad (4)$$

The last stage is the output stage. The output depends on the filtered cell state. Firstly, the information passes through the sigmoid layer that decides what parts of the cell state will be output. Afterward, the cell state passes through a layer to produce outputs between − 1 and 1. It later gets multiplied with sigmoid layers output so that only the outputs decided

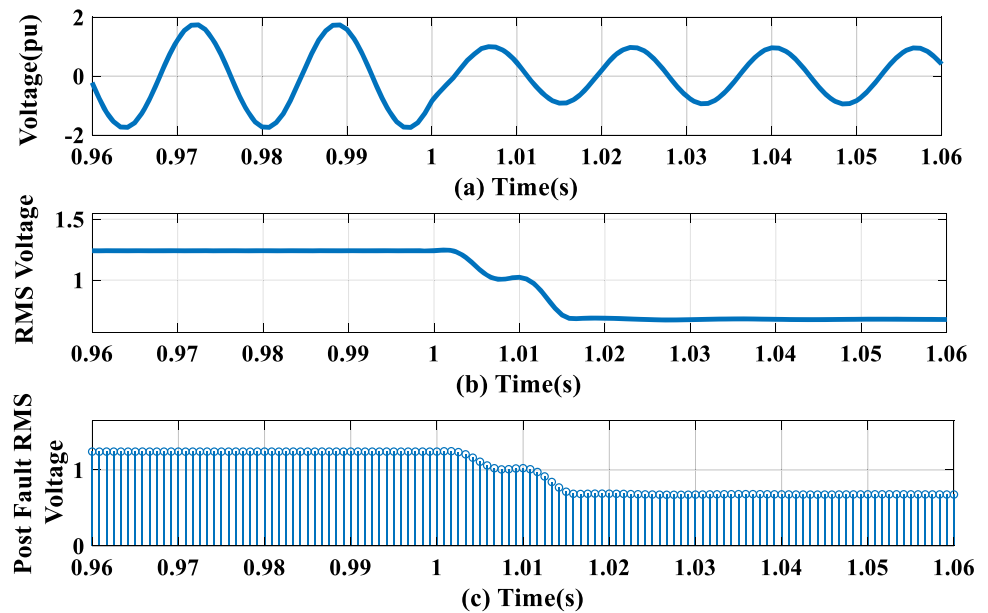**Fig. 5 a** Phase A voltage. **b** RMS voltage. **c** RMS voltage of two cycles post-fault

Table 1 Fault cases used for validation of the proposed method

Parameter variation	Train/Test dataset	No. of cases
HIF Fault Type (LG, LLG, LLLG)	AG, BG, CG, ABG, ACG, BCG, ABCG	7
HIF Fault Resistance	Fault resistance varied between 150 Ω and 500 Ω	10
LIF Fault Type (LG, LLG, LLLG, LL)	AG, BG, CG, ABG, ACG, BCG, ABCG, AB, BC, AC, ABC	11
LIF Fault Resistance	Fault resistance varied randomly between 0 Ω and 50 Ω	10
Fault inception angle	Angles between $0°$ and $90°$ with steps of $15°$	6
Fault location	Between 5–60 km with steps of 5 km	12

Table 2 Non-fault cases used for validation of the proposed method

Parameter variation	Train/Test dataset	No. of cases
Capacitor switching (CS) (In, Out)	Capacitor bank rating: 5,10,15,20 MVAR	8
Generator tripping (Out-In)	Bus 1, 2, 3, 6, and 8	10
Load change (All 11 loads)	± (10, 20, 30, 40, 50, 60)% change in original loading	12
Voltage sag	Bus 1, 2, 3, and 8	4
Nonlinear load (NLL) switching	10–100 kW with steps of 10 kW	10
Motor load switching (MLS)	200–600 kW with steps of 100 kW	4
Distorted grid (CS, MLS, NLLS)	3rd and 5th harmonics up to 4.5%	5
Miscellaneous cases	Evolving fault, LIF, HIF, simultaneous faults	10
Load switch in capacitor out	At bus 8	1
Load Tripping	at 11 buses	11

upon earlier are obtained. The mathematical representation of the last stage is given by Eqs. (5) and (6).

$$O_t = \sigma\left(W_O\left[h_{t-1}, x_t\right] + b_O\right) \tag{5}$$

$$h_t = (o_t \tanh(C_t) \tag{6}$$

Here O_t is the output state activation vector and W_O and b_O represent the corresponding weight and bias matrix.

The LSTM network used in the approach proposed in this paper classifies various types of symmetrical and asymmetrical faults as well as detects their presence. The data are transferred from the input nodes to the LSTM hidden layer, which is composed of several LSTM units. A fully connected dense layer receives the LSTM layer output. As the output of this layer provides the probability of class labels, the fully connected layer is responsible for the high-level reasoning required for classification. Furthermore, a fully connected layer optimizes the objective by learning nonlinear amalgamations present among designated features. In contrast to pooling layers, fully connected layers contain weights and
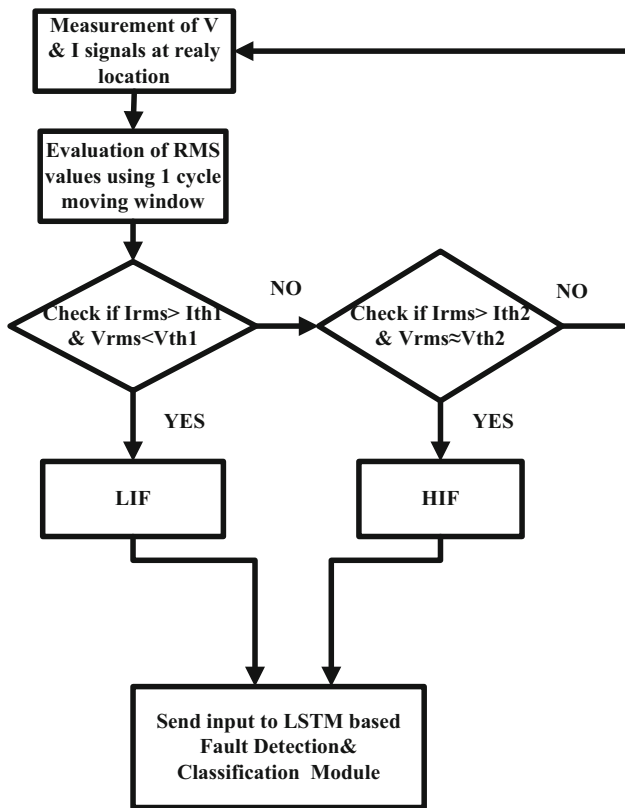
Fig. 6 Flowchart of the proposed method

Table 3 LSTM parameters used for different modules

Module	Sub modules	Values of parameters
Fault Detection	–	Epochs: 140, Batch Size: 150, Dropout: 0.2, Optimizer: Adam, Learning Rate: 0.001
Fault Classification	LIF	Epochs: 200, Batch Size: 90, Dropout: 0.2, Optimizer: Adam, Learning Rate: 0.001
	HIF	Epochs: 100, Batch Size: 40, Dropout: 0.1, Optimizer: Adam, Learning Rate: 0.001

intercepts that multiply trainable weights by the input features. In addition, they contain an additional bias that can be selectively applied. Finally, a multiclass classifier employs a softmax activation function in the last dense layer. A sigmoid activates this layer during binary classification. CNNs are considered one of the most favored deep learning practices, but they suffer from overfitting and vanishing gradient problems and training large-scale networks requires a lot of processing power. LSTM uses additive gradient structures to solve vanishing gradient problems. The network can activate

the forget gate directly and update the gate frequently at every time step, thereby achieving the desired performance from error gradients. With LSTMs, patterns can be remembered over long periods, giving them an advantage over other deep learning methods [18]. In LSTMs, information flows through cell states, allowing some information to be retained while some information is forgotten. An LSTM only adjusts by multiplying and adding, unlike other methods. In addition, LSTM networks preserve constant backward propagation error rates. Even if the time steps are large, the network can learn dependencies.

3 Proposed methodology

The relaying scheme proposed for fault detection and classification in the transmission system consists of three parts: input preparation, fault detection, and fault classification. The detailed architecture of the proposed methodology is shown in Fig. 2. The various steps incorporated in the relaying scheme design are discussed below.

1. Simulation of power system module: The IEEE 14 bus system has been simulated in MATLAB/Simulink environment [19]. All the simulations are executed on a system having an Intel CORE i5 processor, 3.4 GHz CPU speed, and 8 GB RAM. Various fault and no-fault cases by varying different system parameters such as fault location and impedances have been simulated to test the efficacy of the proposed scheme.
2. Input preparation: The two-cycle post-fault three-phase voltage and current signals are collected from buses 1, 2, and 3. The root mean square (RMS) value of these signals is calculated using a one-cycle-long moving window. The recursive RMS voltage and current obtained are then utilized as input for the fault detection and classification modules.
3. Training Modules: Two modules have been created, each for fault detection (FD) and fault type classification (FC) with the help of input RMS voltage and current values and their corresponding targets. The modules are trained using suitable LSTM parameters.
4. Testing Modules: 20% of the entire dataset has been reserved for testing both modules. The testing set was randomly selected. The testing data determines the efficacy of the trained module in terms of various parameters such as accuracy.
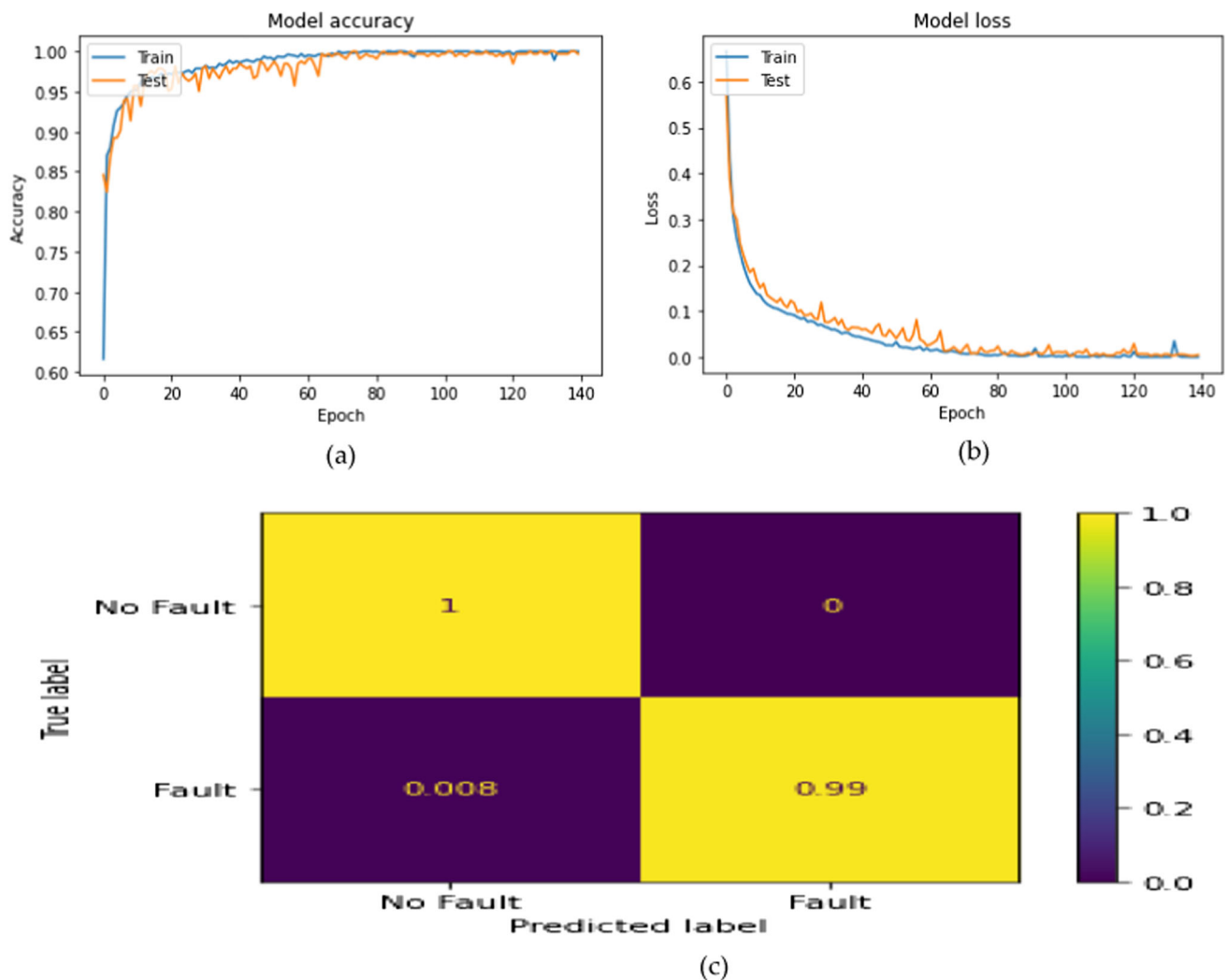
(a)

(b)

(c)

Fig. 7 FD model (**a**) accuracy curve, (**b**) loss curve, (**c**) confusion matrix for the test dataset

Table 4 Class-wise result of the proposed model for FD

Class	Accuracy	Precision	Recall	Specificity	F1-score
No-fault	1.00	0.99	1.00	0.99	1.00
Fault	0.99	1	0.99	1	1.00

3.1 Power system model under study

In this paper, an IEEE 14 bus transmission system [20] as shown in Fig. 3 has been utilized for fault detection and classification purposes. It consists of a 220/132 kV, 60 Hz system, five synchronous generators located on buses 1, 2, 3, 6, and 8, fourteen buses, twenty transmission lines, eleven loads, and various transformers. Out of the five synchronous generators, three are synchronous condensers located on buses 3, 6, and 8. The 11 loads combined have a total power of 259 MW and 81.3 MVar. All the necessary system parameters such as line lengths are discussed in detail in [20]. The

dataset has been generated by simulating the different types of symmetrical and unsymmetrical faults along with some no-fault cases such as load variation and switching events. The fault simulations are done between lines connecting bus 1 and 2, bus 2 and 3, and bus 1 and 5, and data from buses 1.2 and 3 are recorded.

3.2 Preprocessing and data collection

The proposed scheme uses processed three-phase voltage and current signals collected from different buses. The sampling frequency used for data collection is 1.2 kHz. Therefore,

Fig. 8 (**a**) and (**b**) Phase B and C current and voltage signals obtained at bus 3 for BC fault of 5 ohms resistance (**c**) Output of FD module
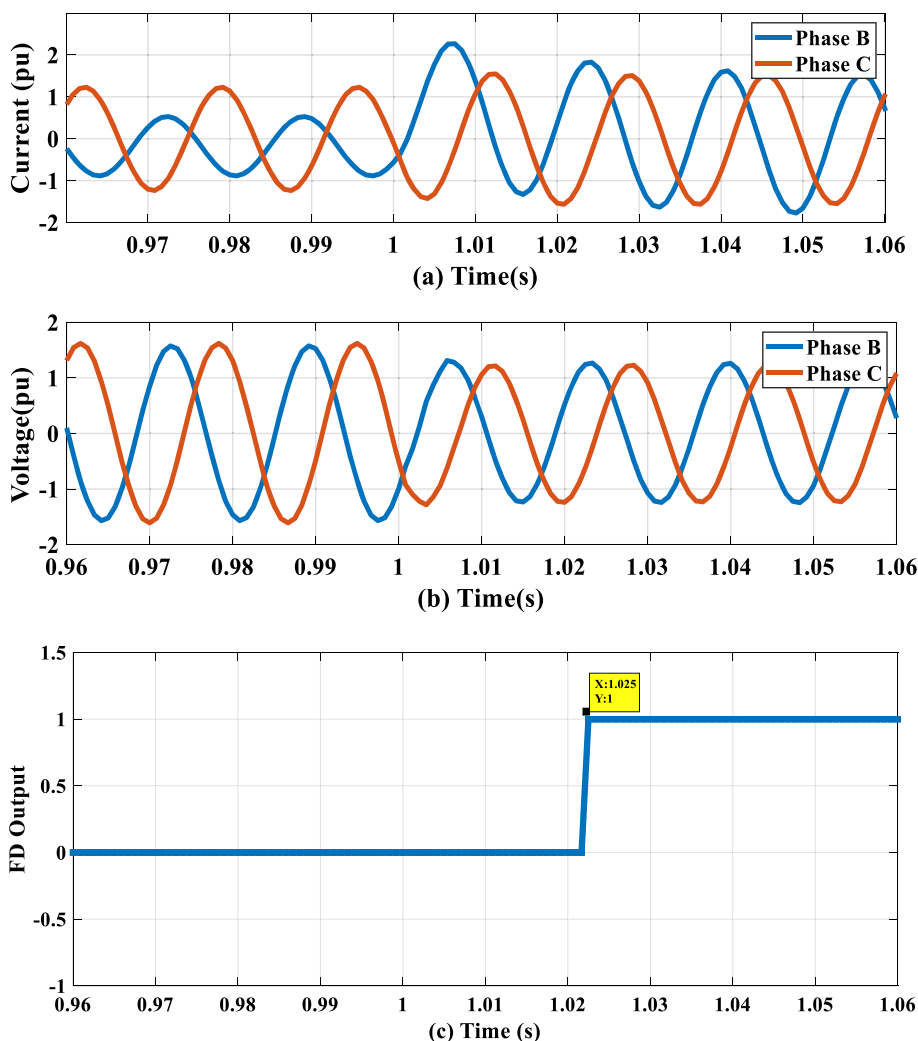

Table 5 FD time for different fault resistance values

Line section	Type of fault	Resistance value (ohm)	FD time (ms)
1–2	AG	0.01	19.4
	ABG	10	20
2–3	CG	15	20.8
	BC	20	22
1–5	BG	100	33
	ACG	250	49.5

the total number of samples in each cycle is ($\frac{1200}{60} = 20 samples$). These signals are further processed by an RMS filter with a moving window of one cycle. The data from two cycles post-fault are then utilized for further processing. Figures 4 and 5 show how pre-processing is performed. The process is shown in detail for an AB-type fault occurring between lines 1 and 5 at 20 km simulated at t = 1 s. Bus 1

voltage and current signals are obtained, and RMS is calculated. Finally, the RMS voltage and current signals of two cycles post-fault are used as input for the fault detection and classification modules.

To train and check the versatility of the proposed model, a variety of fault and no-fault cases are simulated. To achieve this, the modeled system is configured, and its parameters are adjusted. Tables 1 and 2 depict the possible configurations for fault and non-fault cases. All possible fault resistance configurations and fault inception angles are evaluated at different lines (1–2, 2–3, and 1–5) and locations of the system to test all different types of faults. Both LIF and HIF have been simulated.

Also, there are different characteristics of faults occurring at angles from $0°$ to $90°$. A first approach is demonstrated in this paper where fault and non-fault events are assessed on a pure sine voltage main grid and a distorted main grid, where individual harmonics vary from 1 to 4.5%. Power quality standards recommend a maximum of 5% harmonic distortion.

Table 6 FD time for different fault locations

| Line section	Type of fault	Fault location (km)	FD time (ms)
1–2	AG	5	16.7
	AC	10	19.4
2–3	CG	15	20
	ABC	20	20.8
1–5	BC	30	25
	ACG	40	33
	BG	45	50
	BCG	60	55

Table 7 FD time for different fault inception angles

| Line section	Type of fault	Fault inception angle (in °)	FD time (ms)
1–2	BG	0	20.04
	AB	15	19.4
2–3	AC	30	20
	BCG	45	20.8
1–5	ABG	60	33
	ABC	75	50

The main signal is somewhat distorted because of harmonics, even though it is presumed to be of a pure sine nature. It is typically significant in fault detection. Or it may lead to fault detection failure in distorted grid cases. The literature references [21] consider capacitor and load switching as examples of no-fault events. In addition to these events, this paper covers motor switches, nonlinear loads, and a brief generator shutdown. Several capacitances and local load values are considered when simulating the events of capacitor, motor, and nonlinear load switching. Table 2 describes each case in detail. Each event is applied 1 s after the simulation starts, and the total simulation time is 2 s.

Fig. 9 a FD output and RMS current at different locations for an AG fault simulated at t = 1 s between bus 1 and 2 having a fault resistance of 5 ohms. **b** FD output and RMS voltage for different types of faults simulated between bus 2 and 3 having a fault resistance of 10 ohms
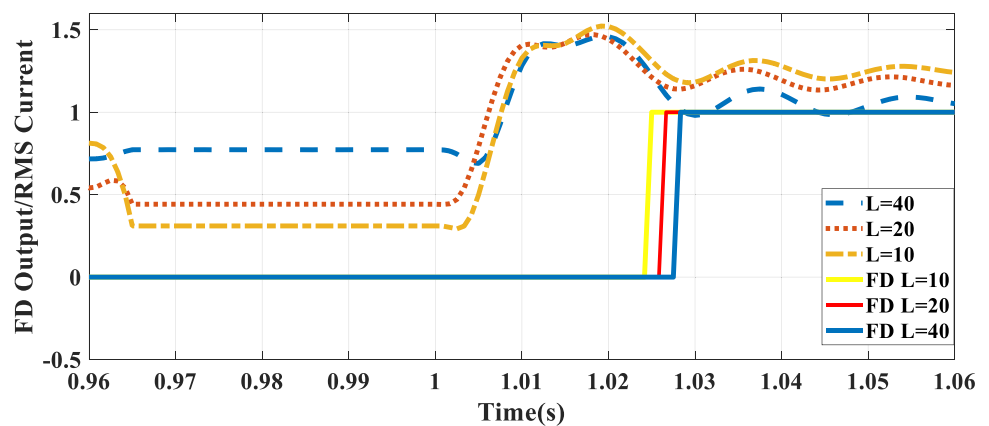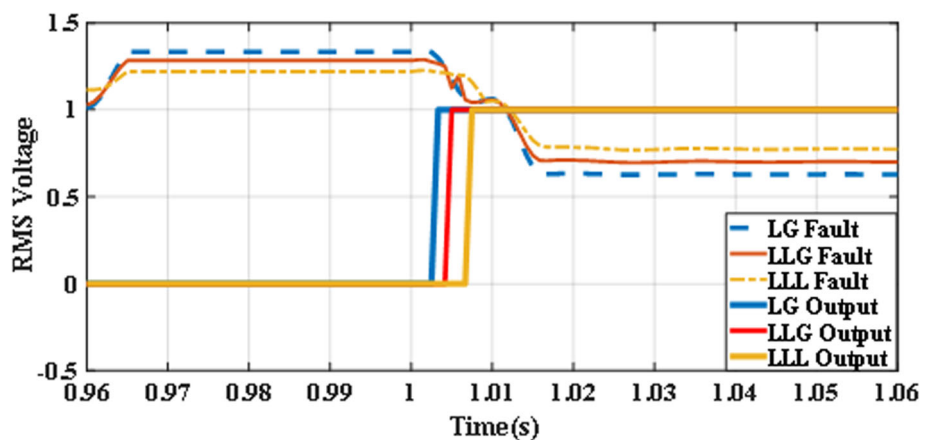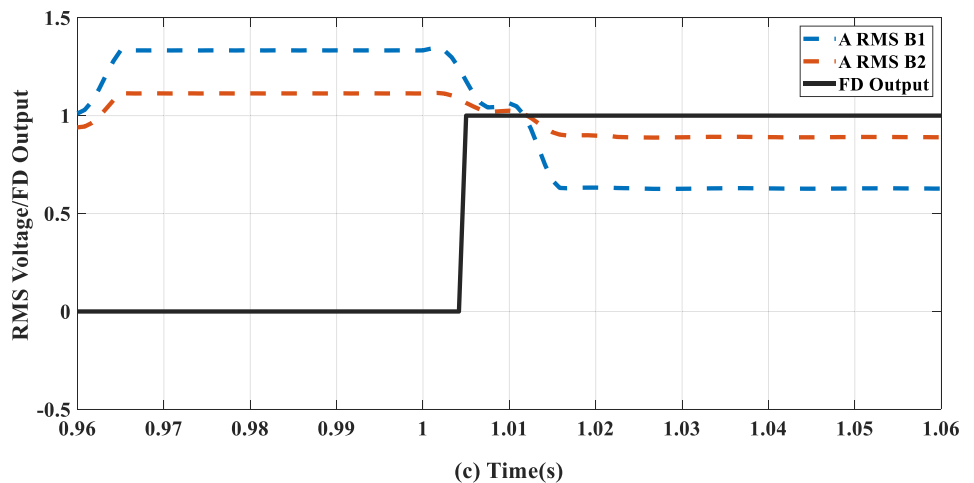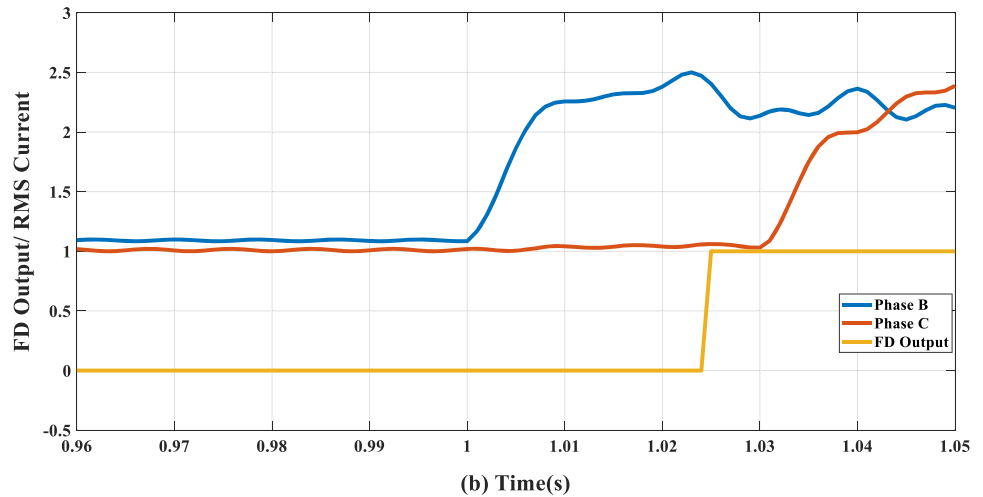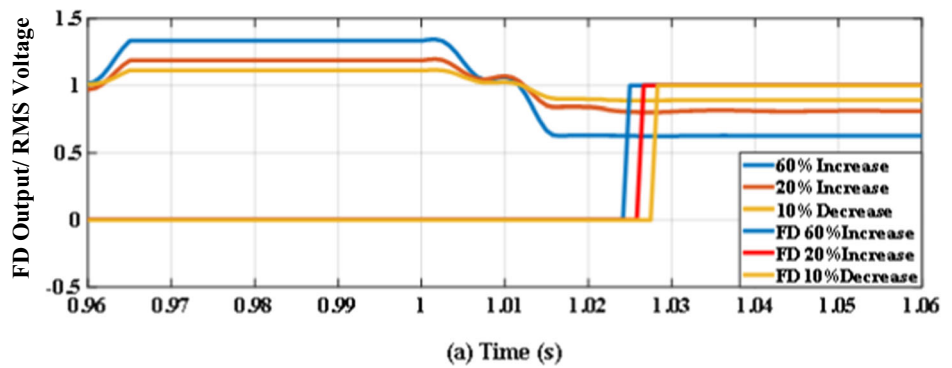
Fig. 10 **a** FD output for variation in load followed by a CG fault incepted between bus 1 and 5 having a fault resistance of 15 ohms. **b** FD output for an evolving fault (BG to BCG) incepted between bus 1 and 2, having a fault resistance of 5 ohms. **c** FD output for a simultaneous fault ABG with 0.01 Ω and ABG with 50 ohms incepted between bus 1 and 2 and bus 2 and 3, respectively

3.3 Design of the fault detection module

An LSTM-based approach has been used in the proposed work for detecting faults in the transmission network. The use of both V and I signals in the presented approach holds significant promise for practical applications, in particular to differentiate between the LIF and HIF faults. During the LIF, the current magnitude increases considerably whereas there is simultaneous decrease in the magnitude of voltage signal as well. On the other hand, during HIF fault, the current magnitude is increased by small magnitude and there is no change in the voltage magnitude. Thus, considering both V and I signal provides a benefit of detecting both LIF and HIF faults using the proposed methodology. The selection of the suitable FD module for LIF and HIF has been determined through threshold analysis. The RMS current and voltage, acquired from the relaying point, are

Table 8 Data description for no-fault and fault data generation

Parameter variation	Train/Test dataset	No. of cases
Capacitor switching in (CS)	Capacitor bank rating: 5,10,15,20 MVAR	4
Generator tripping (Out-In)	Bus 1 and 8	4
Load change (All 11 loads)	$\pm$ (10, 20, 30)% change in original loading	6
Voltage sag	Bus 1 and 3	2
Nonlinear load (NLL) switching	10 and 20 kW	2
Motor load switching (MLS)	200 kW	1
Load tripping	Load at B1 and B2	2
Fault type: AG, BG CG, ABCG	B1, B2, and B3	3

compared with a predefined threshold settings. In the case of LIF detection, the current and voltage thresholds are set at $I_{th1} = 1.5I_{Nominal}$ and $V_{th1} = 0.9V_{rms}$, respectively. Similarly, for HIF detection, the current and voltage thresholds are set at $I_{th2} = 1.16I_{Nominal}$ and $V_{th2} \approx V_{rms}$, respectively. Figure 6 provides a detailed depiction of the process involved in the proposed method.

The LSTM module works in two stages, the training stage where it learns from the input dataset. The testing stage is in which the model's efficiency for testing unknown data is evaluated. All the fault (LIF and HIF) and no-fault current and voltage RMS signals obtained from the three-relay bus are used as input to the FD unit. The input matrix X for the proposed scheme consists of 18 indices. The X matrix is of $18 \times m$ dimensions where m represents the total number of fault cases and 18 is the three-phase RMS voltage and current signals obtained from three buses. The input matrix X is given in Eq. (7). The FD module distinguishes between fault and no-fault conditions. Therefore, a single output is selected at a time, i.e., 0 in case of no-fault and 1 in case of fault. The output matrix for the FD unit is given in Eq. (8).

The LSTM model has been implemented with Keras library with Tensorflow at the backend in Python. A variety of LSTM parameters are applied to design LSTM-FD modules based on inputs and targets. The LSTM network structure used in this experiment consisted of 128 memory units in the LSTM layer with tanh as an activation function. This layer is configured with a dropout. To prevent overfitting in deep learning, dropout is utilized as a regularization technique [22]. If a model is overfitted, it only works on a single dataset and cannot be applied to other datasets. For the construction of these units, a variety of gates are used. These are the forget gate, input gate, and output gate. By removing unnecessary information from the LSTM network, the forget gate increases its efficiency. A tanh function is utilized by the

input gate to add data. In contrast, a sigmoid function is utilized by the regulatory filter to multiply information and add it to the cell's state. Analyzing meaningful data is the output gate's responsibility. After the LSTM layer, a batch normalization (BN) layer is applied. With BN, gradient saturation is reduced throughout the covariate shift, and nonlinearity is added to the convolution output to speed up the learning process. After the LSTM layer, three fully connected, dense hidden layers are added, each with 64, 32 ad 1 memory unit each. Instead of using ReLU, which retains both positive and negative values between 1 and -1, the tanh activation function is implemented in the model for the intended task. The last layer employs a sigmoid activation function because of the binary classification task. Binary cross entropy is used for loss function evaluation. The model is evaluated for 140 epochs for a batch size of 90. Table 3 gives details of model parameters.

$$X = \left\langle \begin{matrix} RMSV_{a1}^{1} \\ RMSV_{b1}^{1} \\ RMSV_{c1}^{1} \\ RMSI_{a1}^{1} \\ RMSI_{b1}^{1} \\ RMSI_{c1}^{1} \\ .. \\ .. \\ .. \\ RMSV_{a3}^{1} \\ RMSV_{b3}^{1} \\ RMSV_{c3}^{1} \\ RMSI_{a3}^{1} \\ RMSI_{b3}^{1} \\ .. \end{matrix} \quad \middle| \ldots\ldots\ldots\ldots \middle| \quad \begin{matrix} RMSV_{a1}^{n} \\ RMSV_{b1}^{n} \\ RMSV_{c1}^{n} \\ RMSI_{a1}^{n} \\ RMSI_{b1}^{n} \\ RMSI_{c1}^{n} \\ .. \\ .. \\ .. \\ .. \\ RMSV_{a3}^{n} \\ RMSV_{b3}^{n} \\ RMSV_{c3}^{n} \\ RMSI_{a3}^{n} \\ RMSI_{b3}^{n} \\ .. \end{matrix} \right\rangle \tag{7}$$

$$Y_1 = \langle 0 \, or \, 1 \rangle \tag{8}$$

3.4 Design of fault classification module

After the detection of faults, the next task is fault classification. There are ten fault events in the case of LIF and six fault scenarios in HIF cases. Two fault classification modules, for LIF and HIF, were created. For fault classification models as well one LSTM layer with 128 memory units and three dense layers with 64, 32, and 10 or 6 memory units have been designed. The first layer is coupled with a dropout layer. BN is performed after every hidden layer. The activation function used at the output dense layer is softmax, and categorical cross entropy is evaluated for obtaining the loss function. Table 3 describes the modules in detail.

Fig. 11 During capacitor switching and CG fault event at bus 8 with fault resistance of 5 ohms occurring consequently (**a**) Phase C current, (**b**) Phase C voltage signals, (**c**) output of FD module
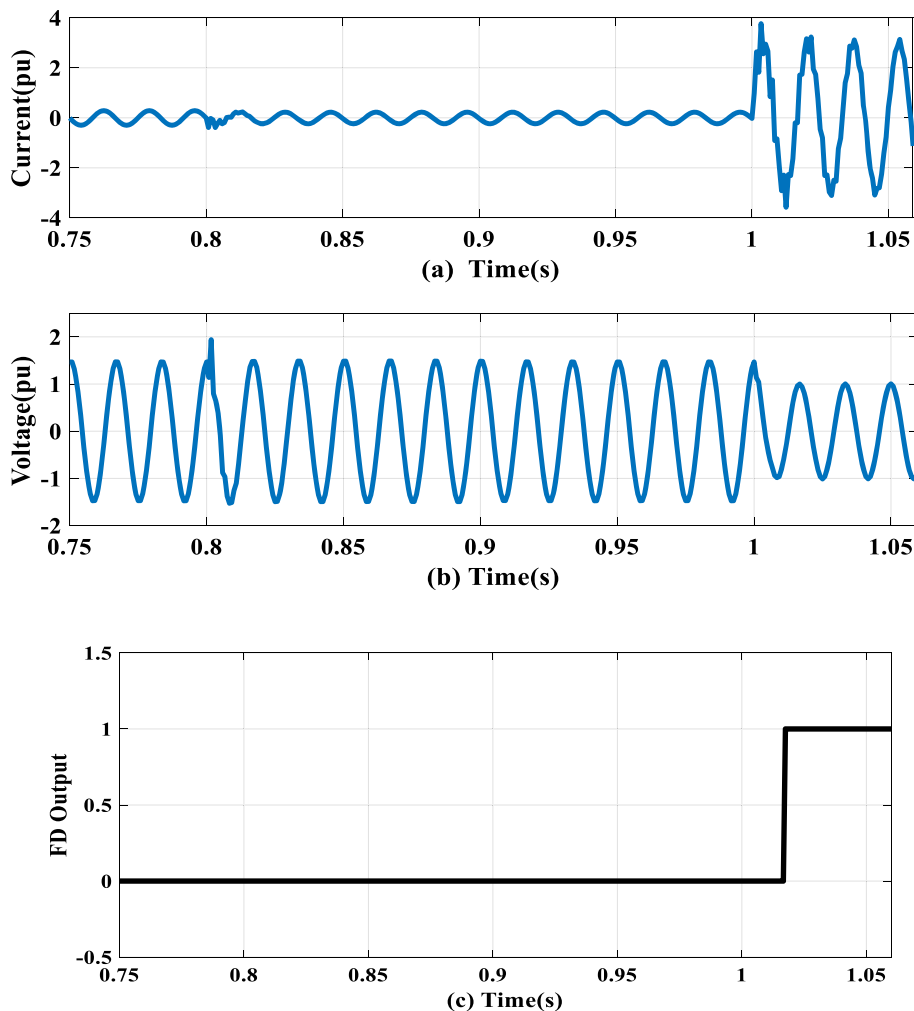
4 Performance evaluation parameters

The post-fault two-cycle RMS voltage and current signals are obtained from buses 1, 2, and 3 and used for further processing. The proposed LSTM-based fault detection and classification have been evaluated under a wide variety of faulty and non-faulty cases. The testing dataset was obtained by splitting the training dataset into different ratios. It has been observed that a train–test split of 80%-20% yields the most desirable results. The module's performance has been verified using accuracy, recall, specificity, precision, and F1-score. Machine learning model accuracy indicates how many times it was correct overall. Precision refers to how well the model predicts a specific class. Recall is the number of times the model has identified an explicit category. Precision and recall yield an F1-score. A higher F1-score indicates greater precision and recall. F1-scores range from 0 to 1. A model is better if it is closer to 1. An algorithm or model's specificity is its capability to anticipate true negatives. These matrices are obtained using a confusion matrix (CM). A CM is a table

consisting of actual and predicted labels. Furthermore, this table maps the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) of the class labels. The fault detection time is calculated with the help of Eq. (9). The time to detect fault is obtained by taking the difference between the time taken to identify the fault and the time of inception of the fault.

$$FDTime = T_{identification\ time} - T_{fault\ inception\ time} \qquad (9)$$

4.1 Fault detection results

The role of the fault detection module is to process the RMS values of both voltage and current signals of two cycles, post-fault data obtained from different buses, and determine whether a fault is present or not present in the system. The performance parameters evaluated for this module have been discussed in the previous section. The performance of the FD model is characterized by various types of symmetrical and unsymmetrical HIF and LIF faults and non-fault cases such
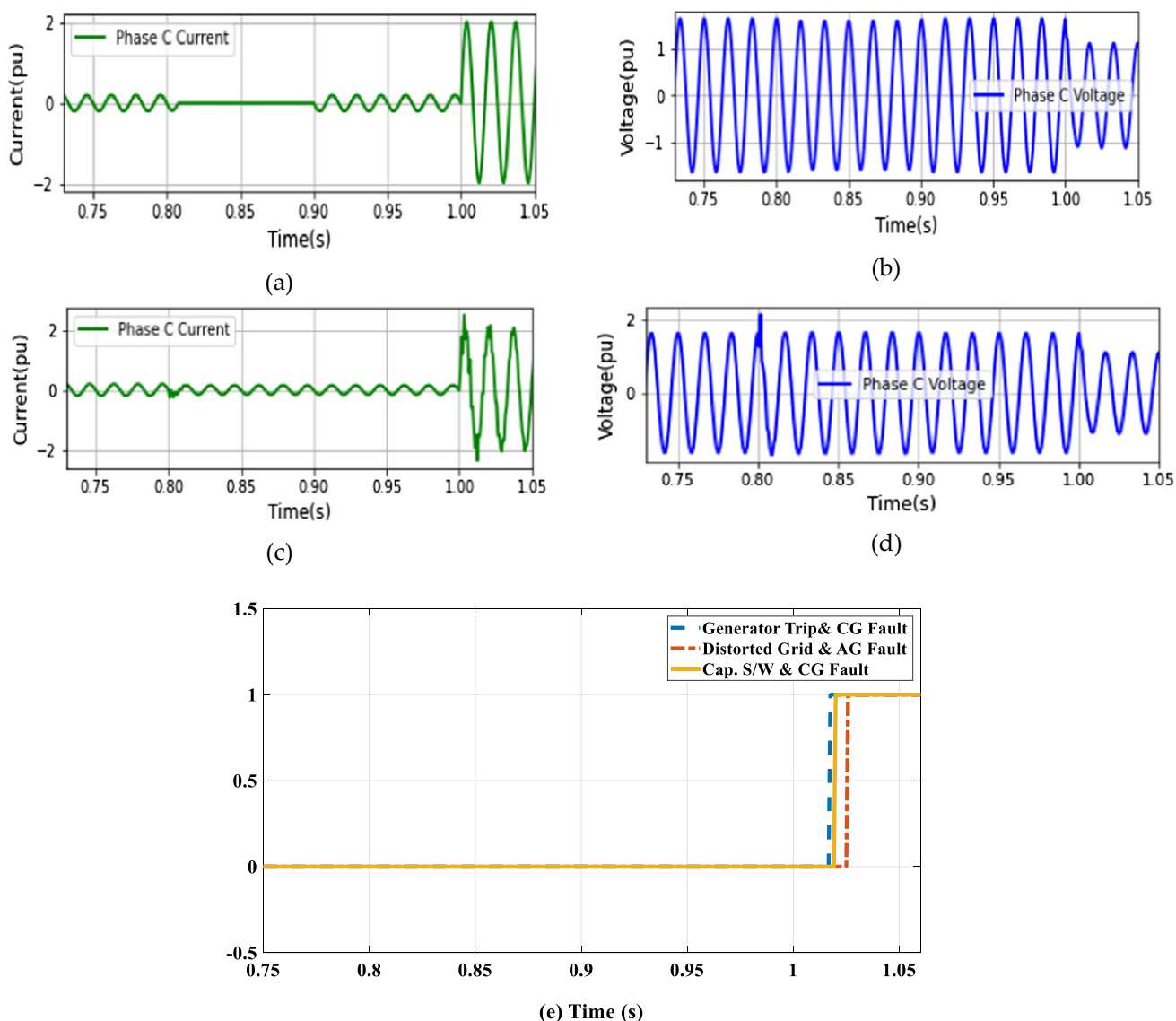
Fig. 12 Case 1—During generator tripping event followed by a CG fault: (**a**) Phase C current (**b**) voltage, Case 2—During capacitor switching event followed by a CG fault: (**c**) Phase C current and (**d**) voltage (**e**) output of FD module in two cases
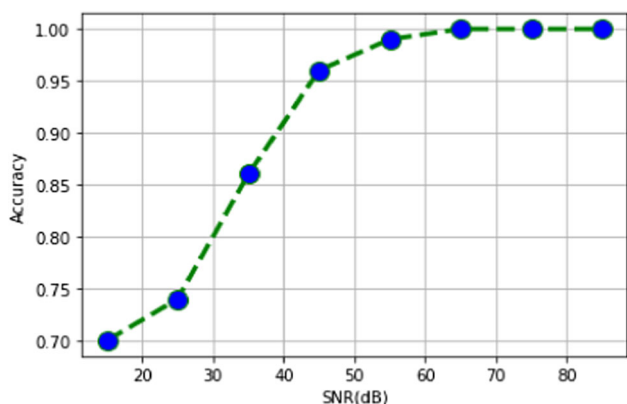

Fig. 13 Accuracy curve of model for fault detection for different noise levels

as changes in load conditions, capacitor and load switching, and generator switching among many other non-faulty conditions as discussed below. Figure 7a demonstrates the model's accuracy curve for both training and testing datasets. It can be observed that the model has a 99.99% accuracy rate during training and testing. Figure 7b shows the loss curve obtained for the model. It can be observed that at the end of 140 epochs, a loss of about 5.1927e-04 is achieved by both training and testing data thereby showing the model's efficiency. The normalized CM obtained for both no-fault and fault classes is shown in Fig. 7c. The output of CM is shown in Table 4. The accuracy for the no-fault class is 100% and that for fault detection is 99%.

Table 9 FD time for different types of faults

Fault type	Line section	SNR (Db)	FD time (ms)
AG	1–2	35	25.04
BCG	2–3	45	19.4
ABCG	2–5	65	20
AB	2–3	75	20.8

4.2 Output of FD for faulty events

The efficiency of the proposed scheme has also been tested for fault detection. A line-to-line BC type fault of resistance 5 ohms has been applied near bus 3 at t = 1 s, and the corresponding current and voltage signals obtained at bus 3 are shown in Fig. 8a, b, respectively. The FD unit issues a trip signal at t = 1.02505 s as depicted in Fig. 8c. Hence it can be seen that the FD module can identify faults in 25.05 ms.

4.2.1 Effect of fault resistance on FD

Often the fault resistance varies with the occurrence of fault in the system. Therefore, irrespective of the fault resistance value the relaying method proposed should work correctly. A wide range of fault resistances as discussed previously has been tested for the proposed LSTM-FD module. The results of the time taken for the detection of faults for different resistance values occurring on different line sections are given in Table 5. Table 5 reflects that the time taken for detection of faults is between 1.5–2.5 cycles in most of the cases.

4.2.2 Effect of fault location on FD

The distance of fault location from the relaying point plays a crucial role in determining the time taken for fault detection. The faults have been simulated at various locations as mentioned in Table 1. The results for the time taken for the detection of faults simulated at various locations on different line lengths are shown in Table 6. The faulty resistance for all the cases demonstrated is kept constant at 5 ohms. Various fault locations have been analyzed, but only a few results are summarized in Table 6 for demonstration purposes. The detection time is considered very small based on the fault locations. In Fig. 9a for an AG fault occurring between bus 15 at various locations, the FD model output has been displayed. Similarly, as the location of the fault increases from the relaying bus, the time to detect the fault also increases.

4.2.3 Effect of type of fault on FD

The type of symmetrical and asymmetrical fault also determines how long it may take to develop these types of faults.

The result for the detection of different types of faults is shown in Fig. 9b. The RMS voltage and the output are shown for simulated faults between lines 2 and 3. This is for a fault resistance value of 10 ohms at a location of 15 km distance. It is observed that the model can identify all types of faults easily.

4.2.4 Effect of variation of fault inception angle on FD

Transmission line faults cannot be predicted exactly in time or at a specific angle. It is therefore imperative that any fault occurrence can be identified by the protection system without fail. The detection scheme should be able to determine faults occurring at different instances of their inception. The results for varying fault inception angles are shown in Table 7. The results suggest the model is capable of determining faults effectively.

4.2.5 Effect of variation of system load on FD

The transmission system is susceptible to load changes due to various reasons. Therefore, the system has been tested for various load increase and decrease cases and its effect on fault determination has been evaluated. Figure 10a showcases the occurrence of a CG type with varying loading on the system. The fault is applied between buses 1 and 5 for a fault resistance of 5 ohms at 20 km. It can be observed that even with an increase or decrease in the loads the fault has been easily detected.

4.2.6 Evolving fault

Line faults can start as one type but eventually change to another type. For example, an LG fault may convert to an LLG fault later. Various such cases have been studied in this research work, and the output of FD for one such case is shown in Fig. 10b. In the presented case, a BG type fault occurring near bus 1 at 1 s later evolves into a BCG type fault at 1.03 s. The fault has been effectively determined in more than a cycle time as seen from the results obtained.

4.2.7 Simultaneous fault

These types of faults occur simultaneously on two or more lines at the same time. To check if the system is robust to any such types of faults, an ABG type of fault is simulated between buses 1 and 2 and also between buses 2 and 3 at 1 s. The fault occurring between buses 1 and 2 has a fault resistance of 0.01 Ω and is located at a distance of 5 km from the relaying bus. Similarly, the other LLG fault is simulated at a distance of 10 km from the relaying bus with a fault resistance of 50 ohms. The RMS current of Phase A along with the FD model output is shown in Fig. 10c. Hence it can
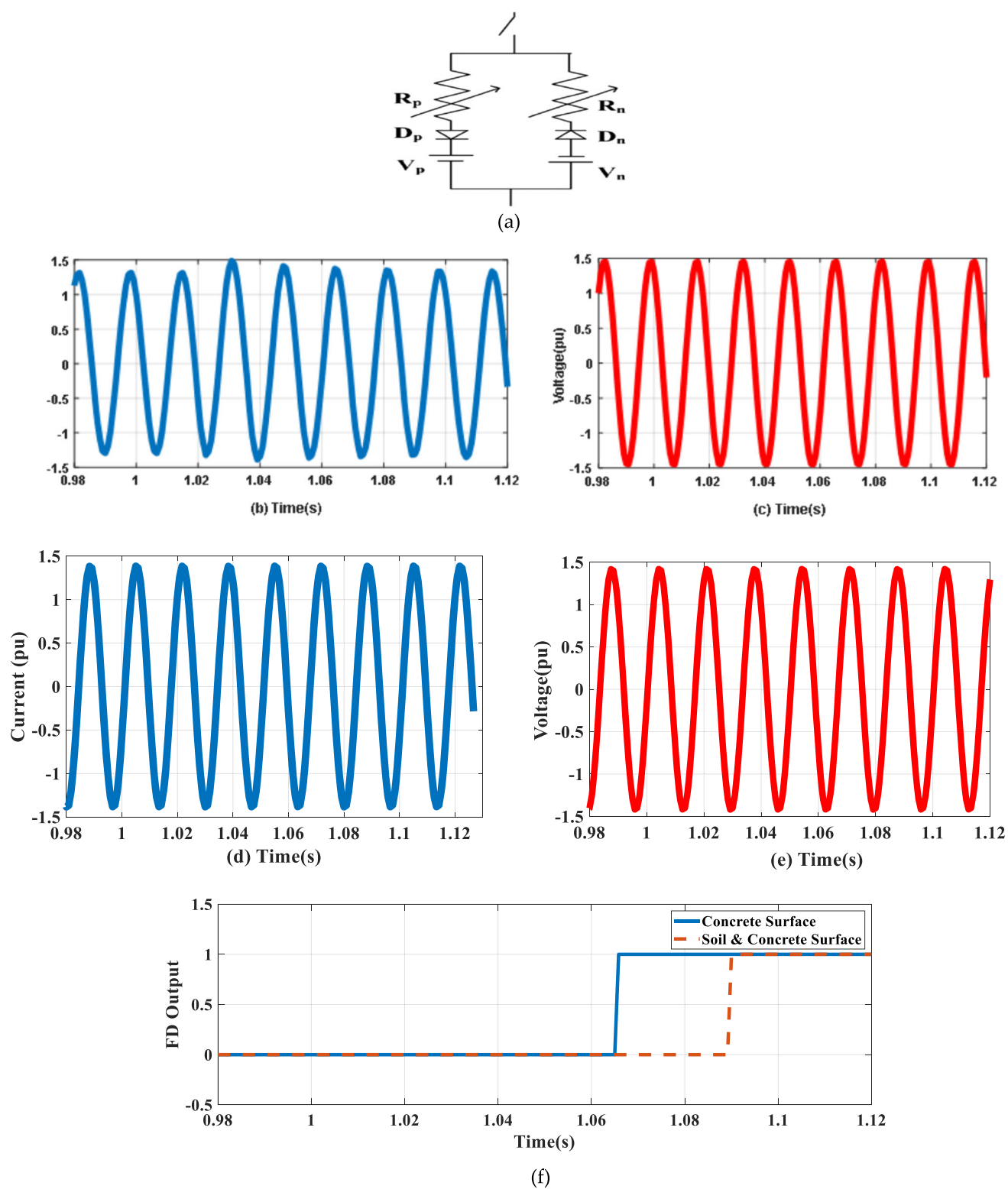
Fig. 14 a Emanuel model for HIF simulation. **b** Case 1: HIF at 1.0 s occurring on concrete surface of 10 cm thickness current signals for Phase A. **c** Voltage signals for Phase A (**d**) Case 2: current signals for Phase A with HIF occurring on a concrete and soil surface of 10 cm thickness (**e**) voltage signals for Phase A (**f**) FD output for Case 1 and 2

Table 10 Simulation parameters for HIF for different surfaces

Type of surface	R_p (Ω)	R_n (Ω)	V_p (V)	V_n (V)	Resistance variations
Concrete surface with 10 cm thickness	150–500	150–500	4500–5500	7500–8500	In steps of 20 ohms every half-cycle
Soil and concrete surface with 10 cm thickness	150–500	150–500	3500–4500	3500–4500	In steps of 5 ohms every half-cycle

Table 11 FD time for HIF for different surfaces

HIF surface	Type of fault	FD time (ms)
Concrete surface	AG	50
	BG	58.35
	BCG	66
Soil and concrete surface	AG	58.35
	ACG	83.3
	ABCG	116.6

be seen that these types of faults are also rapidly detected by the model.

4.3 Fault detection in case of a non-fault event followed by a faulty event

The robustness of the system's FD capability is studied in typical situations. For example, a non-fault event is preceded by a fault event. In the case of capacitor switching followed by a fault in the line to ground, for example, Table 8 discusses all the non-faulty and faulty case studies simulated. To exemplify such a situation, a case of capacitor switching in of 15MVAR rating at bus 9 followed by a CG fault occurring near the same bus has been studied. The voltage and current signals for Phase C captured at bus 8 for capacitor switching in at t = 0.8 s and a CG fault appearing at t = 1 s are illustrated. Figure 11a, b shows that upon switching a capacitor at t = 0.8 s a momentary distortion in the signals occurs. Current and voltage signals increase and decrease in magnitude, respectively, after the fault happens at t = 1 s. Figure 11c corresponds to the FD unit trip signal. It remains low even while capacitor switching occurs and only goes up in case of a fault. Hence this proves the competence of the scheme in such complex scenarios as well.

Figure 12 shows a few more cases considering non-faulty events followed by faulty events. Figure 12a, b demonstrates Phase C current and voltage signals for a generator tripping at bus 1 at t = 0.8 s and a CG fault of resistance 10 ohms at t = 1 s occurring in line between bus 1 and 5. Figure 12c, d corresponds to a CG fault of 20-Ω fault resistance preceded by a 20MVAR capacitor switching event at t = 0.8 s on buses 1 and 2 and bus 9, respectively. All fault events are simulated at t = 1 s. The output in Fig. 11e justifies the

LSTM-FD's ability to distinguish between faulty and non-faulty situations.

4.4 Influence of noise

Power system signals are corrupted by noise. This distorts the signals. To replicate the same white Gaussian noise has been added to the signals to check the model's efficiency in the detection of faults. Noisy signals ranging between 15–85 dB signal–noise ratio (SNR) with a step size of 5 dB have been added to the voltage and current signals. In Fig. 13, we see the average accuracy of the proposed model for FD at different levels of SNR. For noise levels greater than 50 dB, the presented model attains greater than 95% fault detection accuracy. Further, the FD time for different types of fault with varying noise levels is shown in Table 9. In this investigational study, the projected model of fault detection was validated to be robust to noise levels between 40–60 decibels.

4.5 High-impedance fault detection

The HIF occurs when a highly resistive surface such as ground, wet sand, concrete surface, asphalt surface, or grass makes contact with an open live conductor. The HIF exhibits conglomerates consisting of nonlinearity, randomness, asymmetry, shouldering, buildup, and intermittent behavior [21]. These types of faults are highly dangerous because they usually remain undetected by conventional relays because of low fault current values. Additionally, faulted surfaces and humidity conditions play a major role in HIF current characteristics. Consequently, conventional protection mechanisms, such as an overcurrent relay, cannot detect most HIFs. Fires can be caused or personnel safety may be jeopardized by undetected HIFs. Based on antiparallel diodes, DC sources, and nonlinear resistance for each phase, a realistic HIF model, also known as the Emanuel model, utilized by many researchers is modeled to simulate a realistic HIF model. This is shown in Fig. 14a. To make the study more relatable to actual conditions, two different types of HIFs occurring on a concrete surface of 10 cm thickness and a soil and concrete surface of 10 cm thickness have been replicated with the help of the HIF model. Table 10 shows the values of R_p, R_n, V_p, V_n that are changed for every half-cycle to replicate the HIFs of different contact surfaces. The
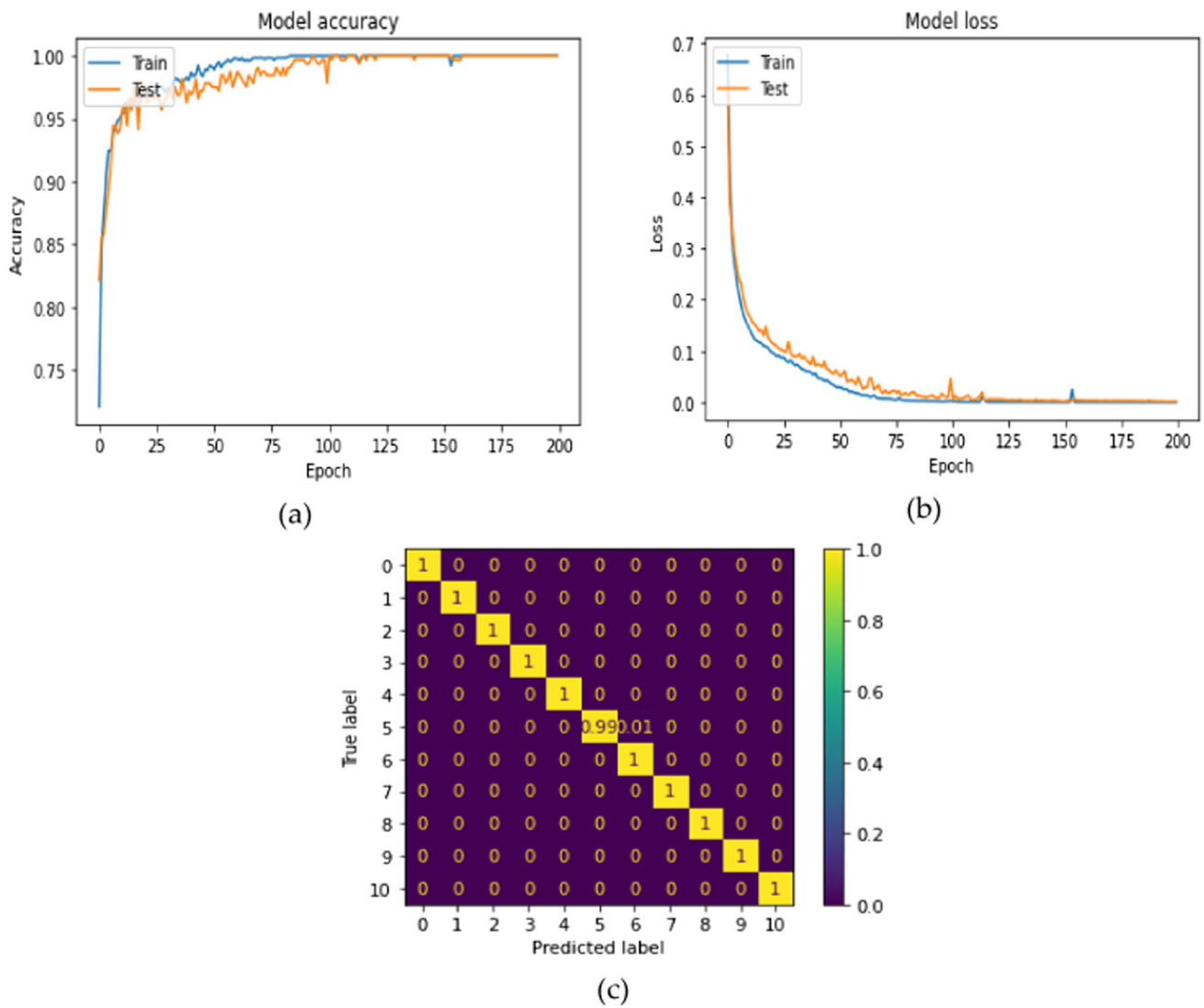
Fig. 15 LIF classification model (**a**) accuracy curve, (**b**) loss curve, (**c**) class-wise confusion matrix

Table 12 Class-wise result of the proposed model for fault classification of LIF

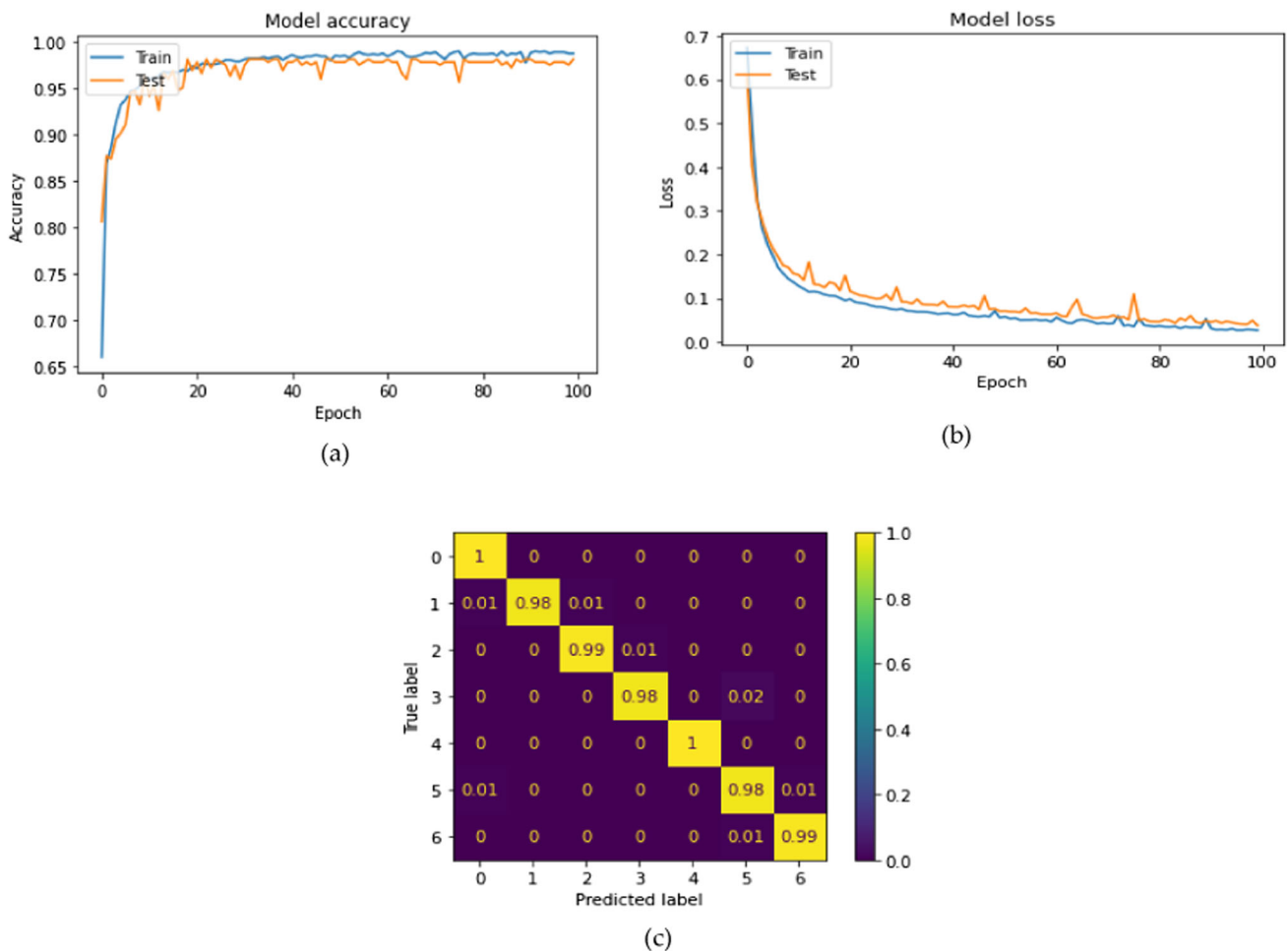| Class	Accuracy	Precision	Recall	Specificity	F1-score
0	1.00	1.00	1.00	1.00	1.00
1	1.00	1.00	1.00	1.00	1.00
2	1.00	1.00	1.00	1.00	1.00
3	1.00	1.00	1.00	1.00	1.00
4	1.00	1.00	1.00	1.00	1.00
5	1.00	1.00	0.99	0.99	0.99
6	0.99	0.99	1.00	1.00	1.00
7	1.00	1.00	1.00	1.00	1.00
8	1.00	1.00	1.00	1.00	1.00
9	1.00	1.00	1.00	1.00	1.00
10	1.00	1.00	1.00	1.00	1.00

Fig. 16 HIF classification model (**a**) accuracy curve, (**b**) loss curve, (**c**) class-wise Confusion Matrix

Table 13 Class-wise result of the proposed model for fault classification of HIF

Class	Accuracy	Precision	Recall	Specificity	F1-score
0	1.00	0.98	1.00	0.99	0.99
1	0.98	1.00	0.98	1.00	0.99
2	0.99	0.99	0.99	0.99	0.99
3	0.98	0.99	0.98	0.98	0.98
4	1.00	1.00	1.00	1.00	1.00
5	0.98	0.97	0.98	0.97	0.98
6	0.99	0.99	0.99	0.98	0.99

current and voltage waveforms of an HIF occurring on a concrete surface of 10 cm thickness as well as a soil and concrete surface of 10 cm thickness are shown in Fig. 14b–e. The FD model output for both cases is shown in Fig. 14f.

All HIF cases are simulated at different distances between bus 2 and bus 3. The dataset is generated by simulating various types of LG, LLG, and LLLG faults. There is a mixture of faults occurring on both types of surfaces in the detection dataset. Table 11 shows the fault detection time for various types of faults occurring on different surfaces. It can be observed that HIF detection time is between 3–7 cycles.

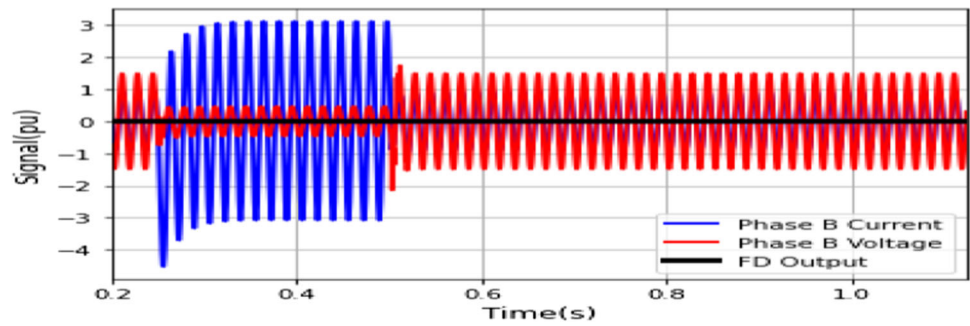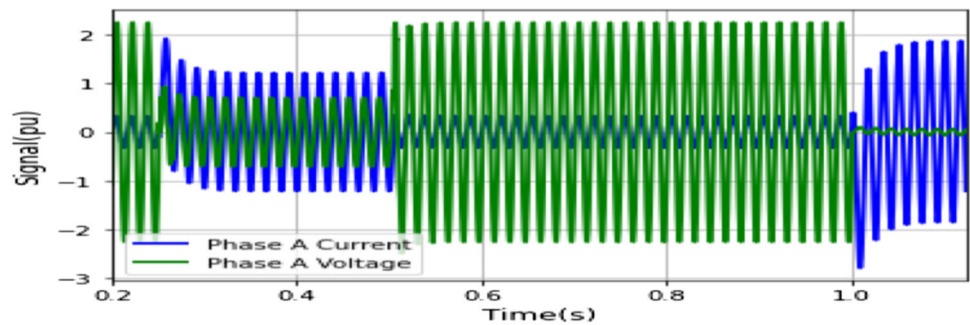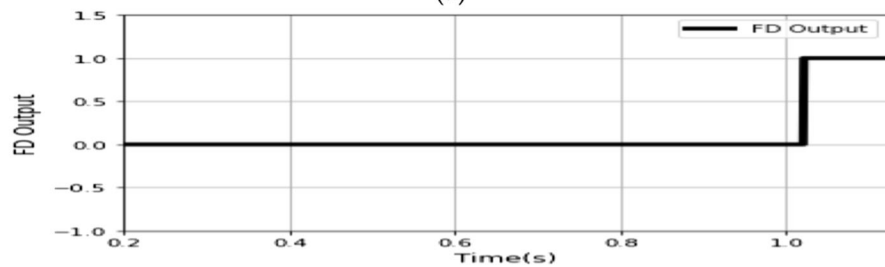**Fig. 17** Output of FD during a power swing situation

Fig. 18 a Phase A current and voltage waveforms during power swing followed by a three-phase fault. **b** Output of FD during a power swing situation

(a)

(b)

Table 14 FD time for LIF occurring during power swing

Type of fault	Fault impedance (Ω)	FD time (s)
ABCG	1	19.4
ABG	15	20
BCG	20	20.04
BG	25	20.8
CG	45	25
AB	50	25

5 Result fault classification

5.1 Low-impedance fault (LIF)

LIFs are the most commonly occurring types of faults in the power system network. These faults can occur for a variety of reasons, such as the fall of trees on the transmission lines. These types of faults cause maximum damage to the

system if not cleared on time. This is because they offer a low-impedance pathway to the current which results in high current magnitude. This can damage the attached power system devices. The effect of symmetrical and unsymmetrical faults on varying fault resistances, inception angles, and locations has been evaluated for the generation of a dataset. The accuracy curve of the model for training and testing dataset split into 80%-20%, respectively, is illustrated in Fig. 15a. From the figure, it can be observed that at the end of 200 epochs, the model's accuracy curves for both train and test data converge and achieve 100% accuracy. As shown in Fig. 15b, the obtained loss curve for the specified model. Both the train and test curves converge near 200 epochs, showing no sign of overfitting and under-fitting. Figure 15c demonstrates the normalized overall class-wise CM obtained for the LIF model. The model correctly classifies all types of faults except for the BC type of fault where classification accuracy is 99.9%. This shows that the proposed model performs fault classification correctly in a multibus system. Table 12
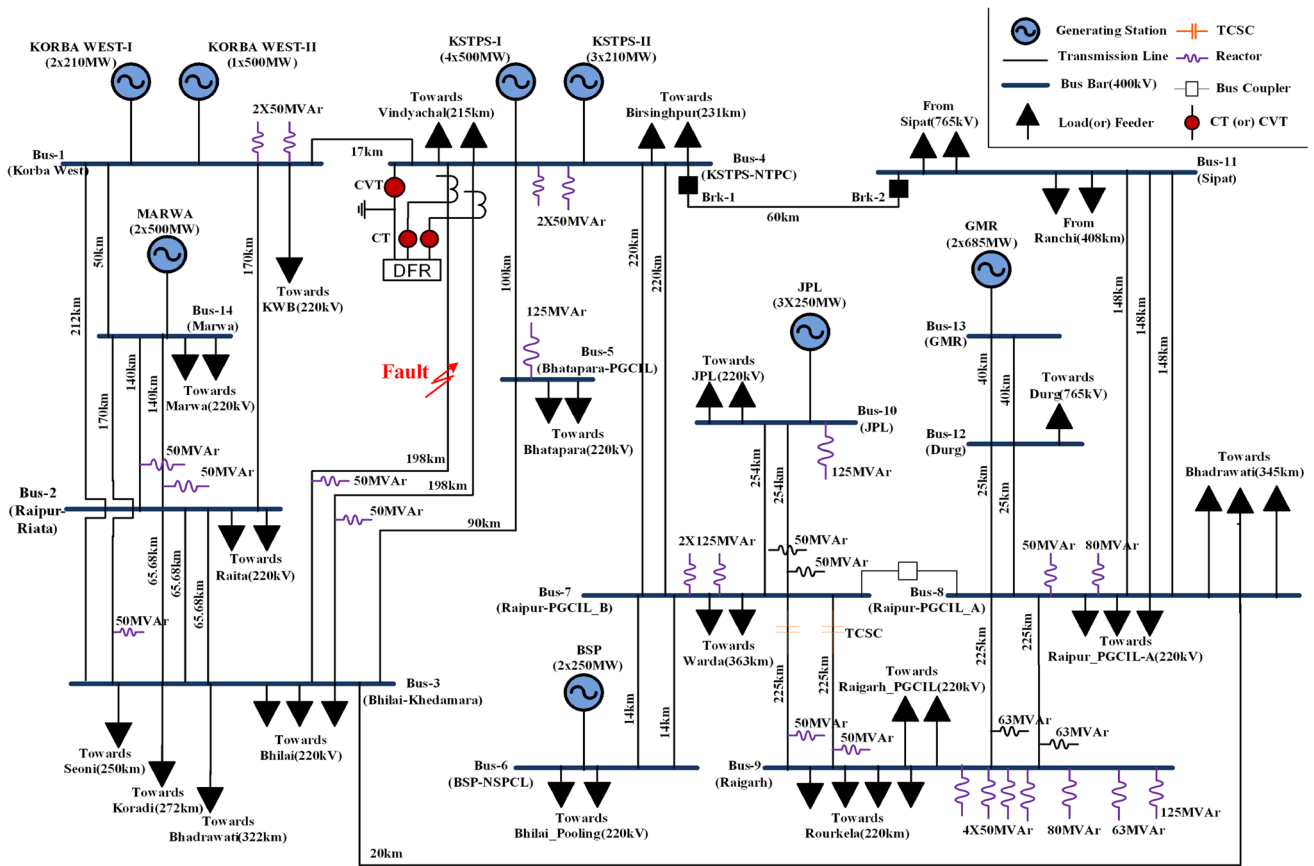
Fig. 19 Chhattisgarh state 400 kV power transmission network diagram

Table 15 LIF FD time for different fault resistance and location values

Type of fault	Resistance value (ohm)	Location (km)	FD time (ms)
CG	0.01	10	24
BCG	10	25	25
AG	15	50	26.1
AB	50	75	32.15
BC	100	100	36
ABCG	250	150	40

discusses the model's performance based on various parameters. Based on the results obtained it can be concluded that overall, the fault classification system performs accurately in all fault scenarios.

5.2 High-impedance fault (HIF)

Various HIF cases corresponding to LG, LLG, and LLLG types have been used for the generation of training datasets. The train-to-test split ratio is 80%-20%, respectively. The accuracy and loss curve for the obtained fault classification model for HIF are shown in Fig. 16a, b, respectively. The average accuracy achieved in this case is 99%. The training is done for 100 epochs. The achieved accuracy is fairly high considering this is a HIF classification. This type of fault is difficult to detect and yet the proposed model has attained an excellent accuracy score with minimal loss. The CM of the HIF model is shown in Fig. 16c shows that labels 0 and 4 for AG and BCG types have been classified with maximum efficiency. Table 13 discusses the model's performance based on various parameters. Based on the results obtained, it can be concluded that overall, the fault classification system performs optimally for all HIF fault scenarios.

6 Power swing

In power systems, power swing is a phenomenon of large fluctuations in power flow from two zones to more, frequently triggered by unpredictable synchronous generators. Power swings can occur as a result of system maintenance, such as replacing a line switch or disconnecting a generator. In the presence of these disturbances, the generator's rotor angle will oscillate, resulting in voltage and current oscillations.

Fig. 20 Phase A voltage, current, and FD output for an AB fault at 0.30 s with fault resistance of 50 ohms occurring at 75 km distance from bus 4
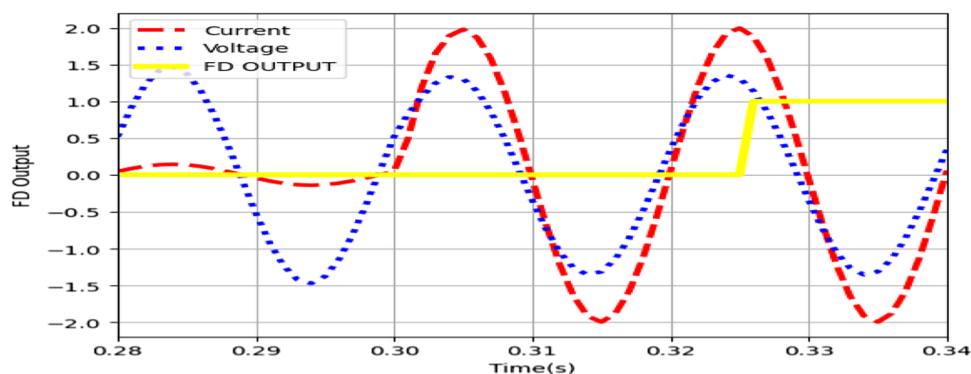
When voltage and current are changing simultaneously, it will cause the relay's measured apparent impedance to be too small, resulting in undesired operation of the relay. To prevent unwanted distance relay operations during power swings, modern distance relays are fortified with power swing-blocking (PSB) schemes. By blocking operation during power swings and unblocking operation when a genuine fault occurs, the power swing protection scheme aims to prevent blocking. Several data surveys and studies have been conducted on power swing-blocking schemes in recent years. The surveys show that these ideas have progressed from early concepts to research in wide-area protection systems.

The proposed method detects power swings, and faults can be detected during them. The test system is set to simulate a power swing by executing a three-phase-to-ground fault at lines 1–2 at t = 0.25 s with an impedance of 1Ω. It clears after 0.5 s. The voltage and current waveform and the corresponding output of the FD unit are shown in Fig. 16. Figure 17 shows that the scheme's proficiency in treating power swing as a fault condition. In addition to simulating power swing, three-phase faults can also be examined during power swing. This is accomplished by configuring a three-phase fault with impedance 1 Ω at lines 2–3 at t = 1 s during power swing. During a three-phase fault and a power swing, Fig. 18a shows the voltage and current waveforms. The output of the FD unit is shown in Fig. 18b. Several cases have been simulated to verify the model's competence by varying the fault impedance, the results of which are presented in Table 14. The scheme has shown its capability to detect faults even during power swing conditions.

7 Evaluation of scheme on an Indian power system network of Chhattisgarh State

To validate the efficacy of the scheme for an existing real transmission network, an analysis has been done in this section. In Fig. 19, an Indian transmission network of 400 kV, 50 Hz is exemplified for Chhattisgarh state. This is based on the power system network data provided in [23]. In this

transmission network, there are eight power generation units, KORBA WEST-I with 2X210MW, KORBA WEST-II with 1 × 500 MW at bus 1, KSTPS-I with 4 × 500 MW and KSTPS-II with 3 × 210 MW at bus 4, MARWA with 2 × 500 MW at bus 14, BSP with 2 × 250 MW at bus 6, JPL with 3 × 250 MW at bus 10, and GMR with 2 × 685 MW at bus 13. The transmission network also consists of four dual-circuit transmission lines (DCTL) (to Vindyachal, Raipur/PGCIL-B, Birsinghpur, Bhilai/Khedamara, 198 km) and three single-circuit transmission lines (to Korba-West, 100 km, and Sipat, 60 km). The bus 3 (Bhilai/Khedamara) has a triple circuit transmission line (TCTL) linking Raipur/Raita 65.68 km to Bhilai 220 kV substation, three feeders connected by stepdown transformer to Bhilai substation, one DCTL (198 km from KSTPS/NTPC), and seven SCTLs (212 km between Korba-West and Bhatpara, 90 km toward Raipur/PGCIL-A, 20 km between Seoni and Koradi, 250 km, Koradi, 272, Bhadrawati, 322 km and Marwa a distance between them). We consider the DCTL connecting buses 4 and 3 (i.e., between KSTPS/NTPC and Bhilai/Khedamara) in this study. At bus 4 (KSTPS/NTPC), we have used the proposed scheme.

7.1 LIF and HIF detection

Different high and low-impedance fault cases have been simulated between buses 3 and 4. Various fault scenarios are simulated at various fault locations and resistances. The output of different case studies for LIF is shown in Table 15. The fault in all these cases is applied at t = 0.3 s. It can be observed that as fault resistance and location increase, the time to detect the fault increases as well. Figure 20 demonstrates the results for an AB type of fault occurring 75 km from the relaying point and having a fault resistance of 50 ohms. It takes approximately 32 ms to detect the fault.

*S*imilarly, different HIF cases have also been simulated. The results are shown in Table 16. The time taken to detect a fault is between 5 and 8 cycles. Figure 21 demonstrates the test results for CG HIF faults occurring at a 10 km distance

Table 16 HIF FD time for different fault resistance and location values

HIF surface	Type of fault	Location (km)	FD time (ms)
Concrete surface	CG	10	85
	ABG	25	100
	BCG	50	124
Soil and concrete surface	CG	75	126.6
	ABCG	100	150
	ACG	150	166.6

on the concrete surface and 75 km at the soil and concrete surface.

8 Comparison with other schemes

In this section, a brief comparison between the proposed LSTM-based fault detection and classification scheme and existing state-of-the-art models is presented. An assessment is made of the accuracy, input, feature extraction, and sturdiness of the model concerning operating conditions, etc., with several machine learning techniques. Table 17 reports the corresponding findings. For instance, the authors in [25] do not demonstrate the efficacy of their LSTM schemes for a range of situations such as evolving fault situations, or effects associated with no-fault situations such as capacitor switching, generator tripping, and load tripping. Moreover, the analysis was performed on a simple 2 machine transmission system. Although the effect of high impedances on fault detection has been demonstrated a detailed study given in this work such as the effect of different fault surfaces has not been shown. The work conducted by authors in [24, 24], and [2] has only presented either a classification or detection task. Also, a detailed analysis regarding the impact of HIF and evolving faults has not been discussed. The presented technique is flexible enough to be used in a wide range of

operating conditions and momentary conditions while maintaining acceptable precision. It is also possible to implement the suggested method in real time since the proposed method does not entail feature extraction. The main strengths of the proposed scheme can be summarized as follows:

- Versatility in fault detection for both LIF and HIF faults.
- Fault classification of both LIF and HIF faults.
- Performing well even in noisy environments.
- Robustness to non-faulty events
- A data-driven approach that automatically learns relevant patterns from the data.
- Proposed scheme has been tested on IEEE 14 bus system and real Indian electrical transmission network.

The LSTM-based approach has scaled well to larger power systems with multiple transmission lines without significant degradation in performance. This scalability advantage is vital for applications in complex and extensive power grids.

9 Conclusion

For fault detection and classification tasks in electrical power distribution networks, an end-to-end learning method is presented in this paper that utilizes LSTMs. By utilizing operational data obtained from measurement nodes, the proposed method avoids a multifaceted feature extraction procedure. LSTM exploits temporal dependencies to process sequential data. Faults, fault types, and no faults are distinguished by temporal insights. Backpropagation is used to train LSTM models on categorized datasets. Current and voltage signals have been used in the designed models. The windowing approach to obtaining RMS voltage and current lowers the computational burden by reconfiguring operational data. The IEEE 14 bus system is extensively tested for fault detection and classification. Performance is robust in various fault settings. This includes fault impedance, line loading conditions, distance from the measurement node, system parameters,

Fig. 21 Output for a CG-type fault occurring on both concrete surface and soil and concrete surface
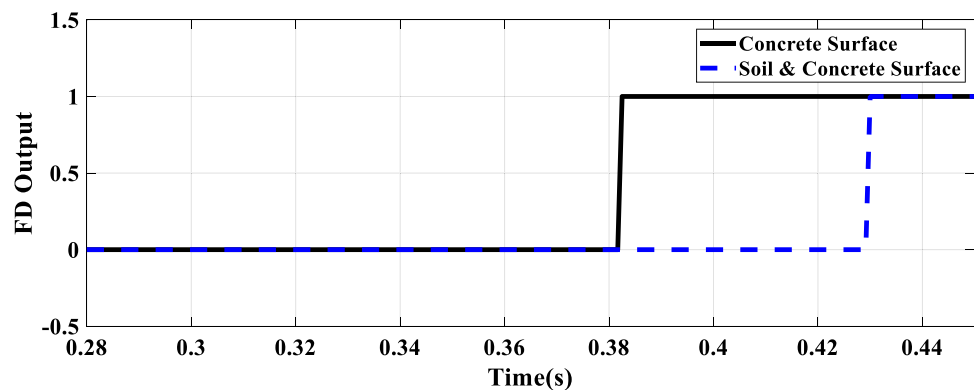
Table 17 Comparison with other schemes

Technique used	Signals recorded	Protective relaying functions provided		HIF	Evolving faults	Simultaneous faults	Impact of noise	Feature extraction	Accuracy (%)
		Detection	Classification						
FCN [17]	V and I	No	Yes	No	No	No	Yes	No	99.57
ANN [24]	I	Yes	Yes	No	No	No	Yes	DWT	99.33
CNN [2]	V and I	No	Yes	No	No	No	Yes	No	99.52
LSTM [25]	V and I	Yes	Yes	Yes	No	No	Yes	No	99.45
Proposed Scheme	V and I	Yes	Yes	Yes	Yes	Yes	Yes	No	99.99

fault inception angles, and measurement device availability. Power swing, evolving faults, simultaneous faults, and high-impedance fault conditions can similarly be detected and worked under by the intended model as well. To further advocate the effectiveness of the method, a real Indian electrical transmission network has also been tested. The scheme has proved efficient for fault detection in this network as well as HIF and LIF. It also illustrates the potential of using end-to-end learning for modern power system actions that include a variety of conservative and non-conventional data sources. However, the scheme poses certain limitations such as it is not suitable for the detection of faults corrupted with noise lower than 30 dB and may not work with a very limited amount of data. In the future authors will work on overcoming these limitations.

Author contributions A. Swetapadma was involved in conceptualization, and A. Yadav took part in methodology. M. Bhatnagar participated in software, validation, and writing—original draft. A. Yadav, A. Swetapadma, and A.Y. Abdelaziz were responsible for writing—review and editing and supervision. All authors have read and agreed to the published version of the manuscript.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. Belagoune S, Bali N, Bakdi A, Baadji B, Atif K (2021) Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems. Measurement 177:109330
2. Rai P, Londhe ND, Raj R (2021) Fault classification in power system distribution network integrated with distributed generators using CNN. Electr Power Syst Res 192:106914
3. Moradzadeh A, Teimourzadeh H, Mohammadi-Ivatloo B, Pourhossein K (2022) Hybrid CNN-LSTM approaches for identification of type and locations of transmission line faults. Int J Electr Power Energy Syst 135:107563
4. Gadanayak DA, Mallick RK (2019) Inter-harmonics based high impedance fault detection in distribution systems using maximum overlap wavelet packet transform and a modified empirical mode decomposition. Int J Electr Power & Energy Syst 112:282–293
5. Aboshady FM, Thomas DWP, Sumner M (2019) A new single end wideband impedance-based fault location scheme for distribution systems. Electr Power Syst Res 173:263–270
6. Lin C, Gao W, Guo M-F (2019) Discrete wavelet transform-based triggering method for single-phase earth fault in power distribution systems. IEEE Trans Power Delivery 34(5):2058–2068
7. Kar S, Samantaray SR (2014) Time-frequency transform-based differential scheme for microgrid protection. IET Gener Transm & Distrib 8(2):310–320
8. Raval PD, Pandya AS (2022) A hybrid PSO-ANN-based fault classification system for EHV transmission lines. IETE J Res 68(4):3086–3099
9. Akhikpemelo A, Evbogbai MJE, Okundamiya MS (2019) Fault detection on a 132kV transmission line using artificial neural network. Int. Rev. Electr. Eng 14(3):220–225
10. Godse R, Bhat S (2020) Mathematical morphology-based feature-extraction technique for detection and classification of faults on power transmission line. IEEE Access 8:38459–38471
11. Wasnik PP, Phadkule NJ, Thakur KD (2020) Fault detection and classification in transmission line by using KNN and DT technique. Int Res J Eng Technol 7(4):335–340
12. Ghaderi A, Ginn HL III, Mohammadpour HA (2017) High impedance fault detection: a review. Electr Power Syst Res 143:376–388
13. Kavi M, Mishra Y, Vilathgamuwa MD (2018) High-impedance fault detection and classification in power system distribution networks using morphological fault detector algorithm. IET Gener Transm Distrib 12(15):3699–3710
14. Wang B, Geng J, Dong X (2016) High-impedance fault detection based on nonlinear voltage–current characteristic profile identification. IEEE Trans Smart Grid 9(4):3783–3791
15. Nezamzadeh-Ejieh S, Sadeghkhani I (2020) HIF detection in distribution networks based on Kullback-Leibler divergence. IET Gener Transm Distrib 14(1):29–36
16. Soheili A, Sadeh J, Bakhshi R (2018) Modified FFT based high impedance fault detection technique considering distribution non-linear loads: simulation and experimental data analysis. Int J Elect Power Energy Syst 94:124–140
17. Tikariha A, Londhe ND, Bag B, Raj R (2021) Classification of faults in an IEEE 30 bus transmission system using fully convolutional network. Int Trans Electr Energy Syst 31(11):e13134

18. Staudemeyer RC, Morris ER (2019) Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. arXiv preprint arXiv:1909.09586

19. MATLAB (2019) The MathWorks, Inc., Natick, Massachusetts. United States

20. Boudreaux JA (2018) Design, simulation, and construction of an IEEE 14-bus power system. LSU Master's Theses. 4801

21. Chaitanya BK, Yadav A, Pazoki M (2019) An intelligent detection of high-impedance faults for distribution lines integrated with distributed generators. IEEE Syst J 14(1):870–879

22. Cheng G, Peddinti V, Povey D, Manohar V, Khudanpur S, Yan Y (August 2017). An exploration of dropout with LSTMs. In: Interspeech (pp. 1586–1590)

23. Ashok V, Yadav A, Abdelaziz AY (2019) MODWT-based fault detection and classification scheme for cross-country and evolving faults. Electr Power Syst Res 175:105897

24. Abdullah A (2017) Ultrafast transmission line fault detection using a DWT-based ANN. IEEE Trans Ind Appl 54(2):1182–1193

25. Rafique F, Fu L, Mai R (2021) End to end machine learning for fault detection and classification in power transmission lines. Electric Power Systems Research 199:107430